A Unified Approach for Solving Sequential Selection Problems

Alexander Goldenshluger*
University of Haifa

Yaakov Malinovsky †
University of Maryland
Baltimore County

Assaf Zeevi‡
Columbia University

Abstract

In this paper we develop a unified approach for solving a wide class of sequential selection problems. This class includes, but is not limited to, selection problems with no–information, rank–dependent rewards, and considers both fixed as well as random problem horizons. The proposed framework is based on a reduction of the original selection problem to one of *optimal stopping* for a sequence of judiciously constructed independent random variables. We demonstrate that our approach allows exact and efficient computation of optimal policies and various performance metrics thereof for a variety of sequential selection problems, several of which have not been solved to date.

Keywords: sequential selection, optimal stopping, secretary problems, relative ranks, full information problems, no–information problems.

2000 AMS Subject Classification: 60G40, 62L15

1 Introduction

In sequential selection problems a decision maker examines a sequence of observations which appear in random order over some horizon. Each observation can be either accepted or rejected, and these decisions are irrevocable. The objective is to select an element in this sequence to optimize a given criterion. A classical example is the so-called *secretary problem* in which the objective is to maximize the probability of selecting the element of the sequence that ranks highest. The existing literature contains numerous settings and formulations of such problems, see, e.g., Gilbert and Mosteller (1966), Freeman (1983), Berezovsky & Gnedin (1984), Ferguson (1989), Samuels (1991) and Ferguson (2008); to make more concrete connections we defer further references to the subsequent section where we formulate the class of problems more precisely.

Sequential selection problems are typically solved using the principles of dynamic programming, relying heavily on structure that is problem-specific, and focusing on theoretical properties of the optimal solution; cf. Gilbert and Mosteller (1966), Berezovsky & Gnedin

*Department of Statistics, University of Haifa, Haifa 31905, Israel. e-mail: *goldensh@stat.haifa.ac.il*. The research was supported by the ISF grant No. 361/15.

†Department of Mathematics and Statistics, University of Maryland, Baltimore County, MD 21250, USA. e-mail: *yaakovm@umbc.edu*

‡Graduate School of Business, Columbia University, New York, NY 10027, USA. e-mail: *assaf@gsb.columbia.edu*

(1984) and Ferguson (2008). Consequently, it has become increasingly difficult to discern commonalities among the multitude of problem variants and their solutions. Moreover, the resulting optimal policies are often viewed as difficult to implement, and focus is placed on deriving sub–optimal policies and various asymptotic approximations; see, e.g., Mucci (1973), Frank & Samuels (1980), Krieger & Samuel-Cahn (2009), and Arlotto & Gurvich (2018), among many others.

In this paper we demonstrate that a wide class of such problems can be solved optimally and in a unified manner. This class includes, but is not limited to, sequential selection problems with *no–information*, *rank–dependent rewards* and allows for fixed or random horizons. The proposed solution methodology covers both problems that have been worked out in the literature, albeit in an instance-specific manner, as well as several problems whose solution to the best of our knowledge is not known to date. We refer to Section 2 for details. The unified framework we develop is based on the fact that various sequential selection problems can be reduced, via a conditioning argument, to a problem of optimal stopping for a sequence of independent random variables that are constructed in a special way. The latter is an instance of a more general class of problems, referred to as *sequential stochastic assignments*, first formulated and solved by Derman, Lieberman & Ross (1972) (some extensions are given in Albright (1972)). The main idea of the proposed framework was briefly sketched in Goldenshluger and Zeevi (2018, Section 4); in this paper it is fully fleshed and adapted to the range of problems alluded to above.

The approach we take is operational, insofar as it supports exact and efficient computation of the optimal policies and corresponding optimal values, as well as various other performance metrics. In the words of Robbins (1970), we "put the problem on a computer." Optimal stopping rules that result from our approach belong to the class of memoryless threshold policies and hence have a relatively simple structure. In particular, the proposed reduction constructs a new sequence of *independent* random variables, and the optimal rule is to stop the first time instant when the current "observation" exceeds a given threshold. The threshold computation is predicated on the structure of the policy in sequential stochastic assignment problems a lá Derman, Lieberman & Ross (1972) and Albright (1972) (as part of the so pursued unification, these problems are also extended in the present paper to the case of a random time horizon). The structure of the optimal stopping rule we derive allows us to explicitly compute probabilistic characteristics and various performance metrics of the stopping time, which, outside of special cases, are completely absent from the literature.

The rest of the paper is structured as follows. Section 2 provides the formulation for the various problem instances that are covered by the proposed unified framework. Section 3 describes the class of stochastic sequential selection problems first formulated in Derman, Lieberman & Ross (1972) that are central to our solution approach. Section 4 formulates the auxiliary stopping problem, and explains its solution via the mapping to a stochastic assignment problem. It then explains the details of the reduction and the structure of the algorithm that implements our proposed stopping rule. Section 5 presents the implementation of said algorithm to the various sequential selection problems surveyed in Secton 2. We close with a few concluding remarks in Section 6.

2 Sequential selection problems

Let us introduce some notation and terminology. Let $X_1, X_2, \dots$ be an infinite sequence of independent identically distributed continuous random variables defined on a probability space $(\Omega, \mathscr{F}, \mathrm{P})$. Let R_t be the relative rank of X_t and $A_{t,n}$ be the absolute rank of X_t among the first n observations (which we also refer to as the *problem horizon*):

$$R_t := \sum_{j=1}^{t} \mathbf{1}(X_t \le X_j), \quad A_{t,n} := \sum_{j=1}^{n} \mathbf{1}(X_t \le X_j), \quad t = 1, \dots, n.$$

Note that with this notation the largest observation has the absolute rank one, and $R_t = A_{t,t}$ for any t. Let $\mathscr{R}_t := \sigma(R_1, \dots, R_t)$ and $\mathscr{X}_t := \sigma(X_1, \dots, X_t)$ denote the σ–fields generated by $R_1, \dots, R_t$ and $X_1, \dots, X_t$, respectively; $\mathscr{R} = (\mathscr{R}_t, 1 \le t \le n)$ and $\mathscr{X} = (\mathscr{X}_t, 1 \le t \le n)$ are the corresponding filtrations. In general, the class of all stopping times of a filtration $\mathscr{Y} = (\mathscr{Y}_t, 1 \le t \le n)$ will be denoted $\mathscr{T}(\mathscr{Y})$; i.e., $\tau \in \mathscr{T}(\mathscr{Y})$ if $\{\tau = t\} \in \mathscr{Y}_t$ for all $1 \le t \le n$.

Sequential selection problems are classified according to the information available to the decision maker and the structure of the reward function. The settings in which only relative ranks $\{R_t\}$ are observed are usually referred to as *no–information problems*, whereas *full information* refers to the case when random variables $\{X_t\}$ are available.

In this paper we mainly consider the class of problems with *no–information* and *rank–dependent reward*. The prototypical sequential selection problem with no–information and rank–dependent reward is formulated as follows; see, e.g., Gnedin & Krengel (1996).

> Problem (A1). Let n be a fixed positive integer, and let $q : \{1, 2, \dots, n\} \to \mathbb{R}$ be a reward function. The average reward of a stopping rule $\tau \in \mathscr{T}(\mathscr{R})$ is
>
> $$V_n(q; \tau) := \mathrm{E} q\big(A_{\tau,n}\big),$$
>
> and we want to find the rule $\tau_* \in \mathscr{T}(\mathscr{R})$ such that
>
> $$V_n^*(q) := \max_{\tau \in \mathscr{T}(\mathscr{R})} V_n(q; \tau) = \mathrm{E} q\big(A_{\tau_*,n}\big).$$
>
> We are naturally also interested in the computation of the optimal value $V_n^*(q)$.

Depending on the reward function q we distinguish among the following types of sequential selection problems.

Best–choice problems. The settings in which the reward function is an indicator are usually referred to as *best–choice stopping problems*. Of special note are the following.

(P1). *Classical secretary problem* corresponds to the case $q(a) = q_{\mathrm{csp}}(a) := \mathbf{1}\{a = 1\}$. Here we want to maximize the probability $\mathrm{P}\{A_{\tau,n} = 1\}$ of selecting the best alternative over all stopping times τ from $\mathscr{T}(\mathscr{R})$. It is well known that the optimal policy will pass on approximately the first n/e observations and select the first subsequent to that which is superior than all previous ones, if such an observation exists; otherwise the last element in the sequence is selected. The limiting optimal value is $\lim_{n\to\infty} V_n^*(q_{\mathrm{csp}}) = 1/e$ (Dynkin 1963, Gilbert and Mosteller 1966). We refer to Ferguson (1989) where the history of this problem is reviewed in detail.

(P2). *Selecting one of the k best values.* The problem is usually referred to as *the Gusein–Zade stopping problem* (Gusein–Zade 1966, Frank & Samuels 1980). Here $q(a) = q_{\rm gz}^{(k)}(a) := \mathbf{1}\{a \le k\}$, and the problem is to maximize $\mathrm{P}\{A_{\tau,n} \le k\}$ with respect to $\tau \in \mathscr{T}(\mathscr{R})$. The optimal policy was characterized in Gusein–Zade (1966). It is determined by k natural numbers $1 \le \pi_1 \le \cdots \le \pi_k$ and proceeds as follows: pass the first $\pi_1 - 1$ observations and among the subsequent $\pi_1, \pi_1 + 1, \ldots, \pi_2 - 1$ choose the first best observation; if it does not exists then among the set of observations $\pi_2, \pi_2+1, \ldots, \pi_3-1$ choose one of the two best, etc. Gusein–Zade (1966) studied the limiting behavior of the numbers $\pi_1, \ldots, \pi_k$ as the problem horizon grows large, and showed that $\lim_{n\to\infty} V_n^*(q_{\rm gz}^{(2)}) \approx 0.574$. Exact results for the case $k = 3$ are given in Quine & Law (1996). The above optimal policy requires determination of $\pi_1, \ldots, \pi_k$ which is computationally challenging for general $k > 3$ and general n; exact values of $V_n^*\big(q_{\rm gz}^{(k)}\big)$ are not reported in the literature. Based on general asymptotic results of Mucci (1973), Frank & Samuels (1980) computed numerically $\lim_{n\to\infty} V_n^*\big(q_{\rm gz}^{(k)}\big)$ for a range of different values of k. The recent paper Dietz et al. (2011) studies some approximate policies.

(P3). *Selecting the kth best alternative.* In this problem $q(a) = q_{\rm pd}^{(k)}(a) := \mathbf{1}\{a = k\}$, i.e. we want to maximize the probability of selecting the kth best candidate. The problem was explicitly solved for $k = 2$ by Rose (1982a) and Vanderbei (2012); the last paper coined the name the *postdoc problem* for this setting. The optimal policy for $k = 2$ is to reject first $\lfloor n/2 \rfloor$ observations and then select the one which is the second best relative to this previous observation set, if it exits; otherwise the last element in the sequence is selected. The optimal value is $V_n^*(q_{\rm pd}^{(2)}) = (n + 1)/4n$. An optimal stopping rule for the case $k = 3$ and some results on the optimal value were reported recently in Lin et al. (2016). We are not aware of results on the optimal policy and exact computation of the optimal values for general n and k. Recently approximate policies were developed in Bruss & Louchard (2016). The problem of selecting the median value $k = (n + 1)/2$, where n is odd, was considered in Rose (1982b). It is shown there that $\lim_{n\to\infty} V_n^*(q_{\rm pd}^{((n+1)/2)}) = 0$.

Expected rank type problems. To this category we attribute problems with reward q which is not an indicator function.

(P4). *Minimization of the expected rank.* In this problem the goal is to minimize $\mathrm{E}A_{\tau,n}$ with respect to $\tau \in \mathscr{T}(\mathscr{R})$. If we put $q(a) = q_{\rm er}(a) := -a$ then

$$\min_{\tau\in\mathscr{T}(\mathscr{R})} \mathrm{E}A_{\tau,n} = - \max_{\tau\in\mathscr{T}(\mathscr{R})} \mathrm{E}\, q_{\rm er}(A_{\tau,n}). \tag{1}$$

This problem was discussed heuristically by Lindley (1961) and solved by Chow et al. (1964). It was shown there that $\lim_{n\to\infty} \min_{\tau\in\mathscr{T}(\mathscr{R})} \mathrm{E}A_{\tau,n} = \prod_{j=1}^{\infty}(1 + \frac{2}{j})^{1/(j+1)} \approx 3.8695$. The corresponding optimal stopping rule is given by backward induction relations. A simple suboptimal stopping rule which is close to the optimal one was proposed in Krieger & Samuel-Cahn (2009).

(P5). *Minimization of the expected squared rank.* Based on Chow et al. (1964), Robbins (1991) developed the optimal policy and computed the asymptotic optimal value in the problem of minimization of $\mathrm{E}[A_{\tau,n}(A_{\tau,n} + 1)\cdots(A_{\tau,n} + k - 1)]$ with respect to $\tau \in \mathscr{T}(\mathscr{R})$.

In particular, he showed that for the optimal stopping rule τ_*

$$\lim_{n\to\infty} \mathrm{E}[A_{\tau_*,n}(A_{\tau_*,n}+1)\cdots(A_{\tau_*,n}+k-1)] = k!\bigg\{\prod_{j=1}^{\infty}\bigg(1+\frac{k+1}{j}\bigg)^{1/(k+j)}\bigg\}^k.$$

Robbins (1991) also discussed the problem of minimization of $\mathrm{E}A^2_{\tau,n}$ over $\tau \in \mathscr{T}(\mathscr{R})$ and mentioned that the optimal stopping rule and optimal value are unknown. As we will demonstrate below, optimal policies for any problem of this type can be easily derived, and the corresponding optimal values are straightforwardly calculated for any fixed n.

Problems with a random horizon. The standard assumption in sequential selection problems is that the problem horizon n is fixed beforehand, and optimal policies depend critically on this assumption. However, in practical situations n may be unknown. This fact motivates settings in which n is assumed to be a random variable independent of the observations. A general sequential selection problem with no–information, rank–dependent reward and random horizon can be formulated as follows.

Problem (A2). Let N be a positive integer random variable with distribution $\gamma = \{\gamma_k\}$, $\gamma_k = \mathrm{P}(N = k)$, $k = 1, 2, \ldots, N_{\max}$, where $N_{\max}$ may be infinite. Assume that N is independent of the sequence $\{X_t, t \geq 1\}$. Let $q : \{1, 2, \ldots, N_{\max}\} \to \mathbb{R}$ be a reward function, and let the reward for stopping at time t be $q(A_{t,N})$ provided that $N \geq t$. The performance of a stopping rule $\tau \in \mathscr{T}(\mathscr{R})$ is measured by

$$V_\gamma(q;\tau) := \mathrm{E}\big[q(A_{\tau,N})\mathbf{1}(\tau \leq N)\big].$$

We want to find the stopping rule $\tau_* \in \mathscr{T}(\mathscr{R})$ such that

$$V^*_\gamma(q) := \max_{\tau\in\mathscr{T}(\mathscr{R})} V_\gamma(q;\tau) = V_\gamma(q;\tau_*).$$

We are also interested in computation of the optimal value $V^*_\gamma(q)$.

The problems (P1)–(P5) discussed above can be all considered under the assumption that the observation horizon is random. Below we discuss the following two problem instances.

(P6). *Classical secretary problem with random horizon.* The classical secretary problem with random horizon N was studied in Presman and Sonin (1972). In Problem (P1) where n is fixed, the stopping region is an interval of the form $\{k_n, \ldots, n\}$ for some integer k_n. In contrast to (P1), Presman and Sonin (1972) show that for general distributions of N the optimal policy can involve "islands," i.e., the stopping region can be a union of several disjoint intervals ("islands"). The paper derives some sufficient conditions under which the stopping region is a single interval and presents specific examples satisfying these conditions. In particular, it is shown that in the case of the uniform distribution on $\{1, \ldots, N_{\max}\}$, i.e., $\gamma_k = 1/N_{\max}$, $k = 1, \ldots, N_{\max}$, the stopping region is of the form $\{k_{N_{\max}}, \ldots, N_{\max}\}$ with $k_{N_{\max}}/N_{\max} \to 2e^{-2}$, $V_\gamma(q_{\mathrm{csp}}) \to 2e^{-2}$ as $N_{\max} \to \infty$. The characterization of optimal policies for general distributions of N is not available in the existing literature.

(P7). *Minimization of the expected rank over a random horizon.* Consider a variant of Problem (P4) under the assumption that the horizon is a random variable N with known distribution. In this setting the loss (the negative reward) for stopping at time t is the absolute rank $A_{t,N}$ on the event $\{N \geq t\}$; otherwise, the absolute rank of the last available observation $A_{N,N} = R_N$ is received. We want to minimize the expected loss over all stopping rules $\tau \in \mathscr{T}(\mathscr{R})$. This problem has been considered in Gianini-Pettitt (1979). In particular, it was shown there that if N is uniformly distributed over $\{1, \ldots, N_{\max}\}$ then the expected loss tends to infinity as $N_{\max} \to \infty$. On the other hand, for distributions which are more "concentrated" around $N_{\max}$, the optimal value coincides asymptotically with the one for Problem (P4). Below we demonstrate that this problem can be naturally formulated and solved for general distributions of N using our proposed unifying framework; the details are given in Section 5.

Multiple choice problems. The proposed framework is also applicable for some multiple choice problems. We review some of these settings below.

(P8). *Maximizing the probability of selecting the best observation with k choices.* Assume that one can make k selections, and the reward function equals one if the best observation belongs to the selected subset and zero otherwise. Formally, the problem is to maximize the probability $\mathrm{P}(\cup_{j=1}^{k}\{A_{\tau_j} = 1\})$ over stopping times $\tau_1 < \cdots < \tau_k$ from $\mathscr{T}(\mathscr{R})$. This problem has been considered in Gilbert and Mosteller (1966) who gave numerical results for up to $k = 8$; see also Haggstrom (1967) for theoretical results for $k = 2$.

(P9). *Minimization of the expected average rank.* Assume that k choices are possible, and the goal is to minimize the expected average rank of the selected subset. Formally, the problem is to minimize $\frac{1}{k}\mathrm{E}\sum_{j=1}^{k} A_{\tau_j}$ over stopping times $\tau_1 < \cdots < \tau_k$ of $\mathscr{T}(\mathscr{R})$. For related results we refer to Ajtai et al. (2001), Krieger et al. (2008), Krieger et al. (2007) and Nikolaev & Sofronov (2007).

Miscellaneous problems. The proposed framework extends beyond problems with rank–dependent rewards and no–information. The next two problem instances demonstrate such extensions.

(P10). *Moser's problem with random horizon.* Let $\{X_t, t \geq 1\}$ is a sequence of independent identically distributed random variables with distribution G and expectation μ. Let N be a positive integer-valued random variable with distribution $\{\gamma_k\}$, $\gamma_k = \mathrm{P}(N = k)$, $k = 1, 2, \ldots, N_{\max}$, where $N_{\max} < \infty$.Assume that N is independent of the sequence $\{X_t, t \geq 1\}$. We observe the sequence $X_1, X_2, \ldots$, and the reward for stopping at time t is X_t provided that $t \leq N$; otherwise the reward is X_N. Formally, we want to maximize

$$\mathrm{E}\big[X_\tau \mathbf{1}(\tau \leq N) + X_N \mathbf{1}\{\tau > N\}\big].$$

with respect to all stopping times τ of filtration $\mathscr{X}$. The formulation with fixed $N = n$ and uniformly distributed X_t's on $[0, 1]$ corresponds to the classical problem of Moser (1956).

(P11). *Bruss' Odds–Theorem.* Bruss (2000) considered the following optimal stopping problem. Let $Z_1, \ldots, Z_n$ be independent Benoulli random variables with success probabilities $p_1, \ldots, p_n$ respectively. We observe $Z_1, Z_2, \ldots$ sequentially and want to stop at the time of the last success, i.e., the problem is to find a stopping time $\tau \in \mathscr{T}(\mathscr{Z})$ so as the probability $\mathrm{P}(Z_\tau = 1, Z_{\tau+1} = Z_{\tau+2} = \cdots = Z_n = 0)$ is maximized. Odds–Theorem (Bruss 2000, Theorem 1) states that it is optimal to stop at the first time instance t such that

$$Z_t = 1 \quad \text{and} \quad t \geq t_* := \sup\left\{1, \sup\left\{k = 1, \ldots, n : \sum_{j=k}^{n} \frac{p_j}{q_j} \geq 1\right\}\right\},$$

with $q_j := 1 - p_j$ and $\sup\{\emptyset\} = -\infty$. This statement has been used in various settings for finding optimal stopping policies. In what follows we will demonstrate that Bruss' Odds–Theorem can be easily derived using the proposed framework.

3 Sequential stochastic assignment problems

The unified framework we propose leverages the sequential assignment model toward the solution of the problems presented in Section 2. In this section we consider two formulations of the stochastic sequential assignment problem: the first is the classical formulation introduced by Derman, Lieberman & Ross (1972), while the second one is an extension for random horizon.

3.1 Sequential assignment problem with fixed horizon

The formulation below follows the terminology used by Derman, Lieberman & Ross (1972). Suppose that n jobs arrive sequentially in time, referring henceforth to the latter as the problem horizon. The tth job, $1 \leq t \leq n$, is identified with a random variable Y_t which is observed. The jobs must be assigned to n persons which have known "values" $p_1, \ldots, p_n$. Exactly one job should be assigned to each person, and after the assignment the person becomes unavailable for the next jobs. If the tth job is assigned to the jth person then a reward of $p_j Y_t$ is obtained. The goal is to maximize the expected total reward.

Formally, assume that $Y_1, \ldots, Y_n$ are integrable independent random variables defined on probability space $(\Omega, \mathscr{F}, \mathrm{P})$, and let F_t be the distribution function of Y_t for each t. Let $\mathscr{Y}_t$ denote the σ–field generated by $(Y_1, \ldots, Y_t)$: $\mathscr{Y}_t = \sigma(Y_1, \ldots, Y_t)$, $1 \leq t \leq n$. Suppose that $\pi = (\pi_1, \ldots, \pi_n)$ is a random permutation of $\{1, \ldots, n\}$ defined on $(\Omega, \mathscr{F})$. We say that π is an *assignment policy* (or simply *policy*) if $\{\pi_t = j\} \in \mathscr{Y}_t$ for every $1 \leq j \leq n$ and $1 \leq t \leq n$. That is, π is a policy if it is non–anticipating relative to the filtration $\mathscr{Y} = \{\mathscr{Y}_t, 1 \leq t \leq n\}$ so that tth job is assigned on the basis of information in $\mathscr{Y}_t$. Denote by $\Pi(\mathscr{Y})$ the set of all policies associated with the filtration $\mathscr{Y} = \{\mathscr{Y}_t, 1 \leq t \leq n\}$.

Now consider the following sequential assignment problem.

> Problem (AP1). Given a vector $p = (p_1, \ldots, p_n)$, with $p_1 \leq p_2 \leq \cdots \leq p_n$, we want to maximize *the total expected reward* $S_n(\pi) := \mathrm{E} \sum_{t=1}^{n} p_{\pi_t} Y_t$ with respect to $\pi \in \Pi(\mathscr{Y})$. The policy π^* is called *optimal* if $S_n(\pi^*) = \sup_{\pi \in \Pi(\mathscr{Y})} S_n(\pi)$.

In the sequel the following representation will be useful

$$\sum_{t=1}^{n} p_{\pi_t} Y_t = \sum_{t=1}^{n} \sum_{j=1}^{n} p_j Y_t \mathbf{1}\{\pi_t = j\} = \sum_{j=1}^{n} p_j Y_{\nu_j};$$

here the random variables $\nu_j \in \{1, \ldots, n\}$, $j = 1, \ldots, n$ are given by the one-to-one correspondence $\{\nu_j = t\} = \{\pi_t = j\}$, $1 \le t \le n$, $1 \le j \le n$. In words, ν_j denotes the index of the job to which the jth person is assigned.

The structure of the optimal policy is given by the following statement.

Theorem 1 (Derman, Lieberman & Ross (1972)) *Consider Problem (AP1) with horizon n. There exist real numbers $\{a_{j,n}\}_{j=0}^{n}$,*

$$-\infty \equiv a_{0,n} \le a_{1,n} \le \cdots \le a_{n-1,n} \le a_{n,n} \equiv \infty$$

such that on the first step, when random variable Y_1 distributed F_1 is observed, the optimal policy is $\pi_1^* = \sum_{j=1}^{n} j\mathbf{1}\{Y_1 \in (a_{j-1,n}, a_{j,n}]\}$. The numbers $\{a_{j,n}\}_{j=1}^{n}$ do not depend on $p_1, \ldots, p_n$ and are determined by the following recursive relationship

$$a_{j,n+1} = \int_{a_{j-1,n}}^{a_{j,n}} z \mathrm{d}F_1(z) + a_{j-1,n}F_1(a_{j-1,n}) + a_{j,n}[1 - F_1(a_{j,n})], \;\; j = 1, \ldots, n,$$

where $-\infty \cdot 0$ and $\infty \cdot 0$ are defined to be 0. At the end of the first stage the assigned p is removed from the feasible set and the process repeats with the next observation, where the above calculation is then performed relative to the distribution F_2 and so on. Note that $a_{j,n+1} = \mathrm{E}Y_{\nu_j}$, $\forall 1 \le j \le n$, i.e., $a_{j,n+1}$ is the expected value of the job which is assigned to the jth person.

3.2 Stochastic sequential assignment problems with random horizon

In practical situations the horizon, or number of available jobs, n is often unknown. Under these circumstances the optimal policy of Derman, Lieberman & Ross (1972) is not applicable. This fact provides motivation for the setting with random number of jobs. Nikolaev & Jacobson (2010) considered the sequential assignment problem with a random horizon. They show that the optimal solution to the problem with random horizon can be derived from the solution to an auxiliary assignment problem with dependent job sizes. Here we demonstrate that the problem with random horizon is in fact equivalent to a certain version of the sequential assignment problem with fixed horizon and independent job sizes.

> Problem (AP2). Let N be a positive integer-valued random variable with distribution $\gamma = \{\gamma_k\}$, $\gamma_k = \mathrm{P}(N = k)$, $k = 1, \ldots, N_{\max}$, where $N_{\max}$ can be infinite. Let $Y_1, Y_2, \ldots$ be an infinite sequence of integrable independent random variables with distributions $F_1, F_2, \ldots$, independent of N. Given real numbers $p_1 \le \ldots \le p_{N_{\max}}$ the objective is to maximize the expected total reward $S_\gamma(\pi) = \mathrm{E}\sum_{t=1}^{N} p_{\pi_t} Y_t$ over all policies $\pi \in \Pi(\mathscr{Y})$.

In the following statement we show that Problem (AP2) is equivalent to a version of the standard sequential assignment problem with fixed horizon.

Theorem 2 *In Problem (AP2) assume that $N_{\max} < \infty$ and let $\tilde{Y}_t := Y_t \sum_{k=t}^{N_{\max}} \gamma_k$, $t = 1, \ldots, N_{\max}$. For any $\pi \in \Pi(\mathscr{Y})$ one has $S_\gamma(\pi) = \mathrm{E}\sum_{t=1}^{N_{\max}} p_{\pi_t} \tilde{Y}_t$, and the optimal policy in Problem (AP2) coincides with the optimal policy in Problem (AP1) associated with fixed horizon $n = N_{\max}$ and job sizes $\tilde{Y}_1, \ldots, \tilde{Y}_{N_{\max}}$.*

Proof: For any $\pi \in \Pi(\mathscr{Y})$ we have $S_\gamma(\pi) = \mathrm{E}\sum_{t=1}^{N} p_{\pi_t} Y_t = \sum_{t=1}^{N_{\max}} \mathrm{E}[p_{\pi_t} Y_t \mathbf{1}(N \geq t)]$, and

$$\mathrm{E}[p_{\pi_t} Y_t \mathbf{1}(N \geq t)] = \mathrm{E} \sum_{k=t}^{N_{\max}} \mathrm{E}\Big\{[p_{\pi_t} Y_t \mathbf{1}(N = k)] \,|\, \mathscr{Y}_t\Big\} = \mathrm{E}\Big\{p_{\pi_t} Y_t \sum_{k=t}^{N_{\max}} \gamma_k\Big\},$$

where we have used the fact that π_t is $\mathscr{Y}_t$–measurable, and N is independent of $\mathscr{Y}_t$. Therefore $\mathrm{E}\sum_{t=1}^{N} p_{\pi_t} Y_t = \mathrm{E}\sum_{t=1}^{N_{\max}} p_{\pi_t} \tilde{Y}_t$. Note that $\tilde{Y}_t$ are independent random variables, and σ-fields $\tilde{\mathscr{Y}}_t$ and $\mathscr{Y}_t$ are identical. This implies the stated result. ∎

Remark 1 *To the best of our knowledge, the relation between Problems (AP2) and (AP1) established in Theorem 2 is new. It is worth noting that Nikolaev & Jacobson (2010) developed an optimal policy by reduction of the problem to an auxiliary one with dependent random variables. In contrast, Theorem 2 shows that the problem with random number of jobs is equivalent to the standard sequential assignment problem with independent random variables which is solved by the procedure of Derman, Lieberman & Ross (1972).*

Remark 2 *In Theorem 2 we assume that $N_{\max}$ is finite. Under suitable assumptions on the weights $\{p_j\}$ and jobs sizes $\{Y_t\}$ one can construct ϵ–optimal policies for the problem with infinite $N_{\max}$. However, we do not pursue this direction here.*

4 A unified approach for solving sequential selection problems

4.1 An auxiliary optimal stopping problem

Consider the following auxiliary problem of optimal stopping.

> Problem (B). Let $Y_1, \ldots, Y_n$ be a sequence of integrable independent real-valued random variables with corresponding distributions $F_1, \ldots, F_n$. For a stopping rule $\tau \in \mathscr{T}(\mathscr{Y})$ define $W_n(\tau) := \mathrm{E}Y_\tau$. The objective is to find the stopping rule $\tau_* \in \mathscr{T}(\mathscr{Y})$ such that
>
> $$W_n^* := \max_{\tau \in \mathscr{T}(\mathscr{Y})} \mathrm{E}Y_\tau = W_n(\tau_*) = \mathrm{E}Y_{\tau_*}.$$

Problem (B) is a specific case of the stochastic sequential assignment problem of Derman, Lieberman & Ross (1972), and Theorem 1 has immediate implications for Problem (B). The following statement is a straightforward consequence of Theorem 1.

Corollary 1 *Consider Problem (B). Let $\{b_t,\, t \geq 1\}$ be the sequence of real numbers defined recursively by*

$$b_1 = -\infty, \quad b_2 = \mathrm{E}Y_n,$$
$$b_{t+1} = \int_{b_t}^{\infty} z \mathrm{d}F_{n-t+1}(z) + b_t F_{n-t+1}(b_t), \quad t = 2, \ldots, n. \tag{2}$$

Let

$$\tau_* = \min\{1 \leq t \leq n : Y_t > b_{n-t+1}\}; \tag{3}$$

then

$$W_n^* = \mathrm{E}Y_{\tau_*} = \max_{\tau \in \mathscr{T}(\mathscr{Y})} \mathrm{E}Y_\tau = b_{n+1}.$$

Proof: The integral in (2) is finite because the random variables $Y_1, \ldots, Y_n$ are integrable. Consider Problem (AP1) with $p = (0, \ldots, 0, 1)$. By Theorem 1, at step t the optimal policy assigns value p_n to the job Y_t only if $Y_t > a_{n-t,n-t+1}$, $t = 1, \ldots, n$, and

$$a_{n-t,n-t+1} = \int_{a_{n-t-1,n-t}}^{\infty} z\mathrm{d}F_{t+1}(z) + a_{n-t-1,n-t}F_{t+1}(a_{n-t-1,n-t}).$$

Setting $b_t := a_{t-1,t}$, and noting that $b_1 = -\infty,\ b_2 = \int_{-\infty}^{\infty} z\mathrm{d}F_n(z)$, we come to the required statement. ∎

4.2 Reduction to the auxiliary stopping problem

Problems (A1) and (A2) of Section 2 can be reduced to the optimal stopping of a sequence of independent random variables [Problem (B)]. In order to demonstrate this relationship we use well known properties of the relative and absolute ranks. These properties are briefly recalled in the next paragraph.

Let $A_n := (A_{1,n}, \ldots, A_{n,n})$, and let $\mathcal{A}_n$ denote then set of all permutations of $\{1, \ldots, n\}$; then $\mathrm{P}(A_n = A) = 1/n!$ for all $A \in \mathcal{A}_n$ and all n. The random variables $\{R_t, t \geq 1\}$ are independent, and $\mathrm{P}(R_t = r) = 1/t$ for all $r = 1, \ldots, t$. For any n and $t = 1, \ldots, n$

$$\mathrm{P}(A_{t,n} = a | R_1 = r_1, \ldots, R_t = r_t) = \mathrm{P}(A_t = a | R_t = r_t), \tag{4}$$

and

$$\mathrm{P}(A_{t,n} = a | R_t = r) \ = \ \frac{\binom{a-1}{r-1}\binom{n-a}{t-r}}{\binom{n}{t}}, \quad r \leq a \leq n - t + r. \tag{5}$$

Now we are in a position to establish a relationship between Problems (A1) and (B).

Fixed horizon. Let

$$I_{t,n}(r) := \sum_{a=r}^{n-t+r} q(a) \frac{\binom{a-1}{r-1}\binom{n-a}{t-r}}{\binom{n}{t}}, \quad r = 1, \ldots, t. \tag{6}$$

It follows from (5) that $I_{t,n}(R_t) = \mathrm{E}\{q(A_{t,n}) \,|\, R_t\}$. Define

$$\mathrm{Y}_t := I_{t,n}(R_t), \quad t = 1, \ldots, n. \tag{7}$$

By independence of the relative ranks, $\{Y_t\}$ is a sequence of independent random variables.

The relationship between stopping problems (A1) and (B) is given in the next theorem.

Theorem 3 *The optimal stopping rule τ_* solving Problem (B) with random variables $\{Y_t\}$ given in (6)–(7) also solves Problem (A1):*

$$V_n(q; \tau_*) = \max_{\tau \in \mathscr{T}(\mathscr{R})} \mathrm{E}q(A_{\tau,n}) = \max_{\tau \in \mathscr{T}(\mathscr{Y})} \mathrm{E}Y_\tau = W_n(\tau_*).$$

Proof: First we note that for any stopping rule $\tau \in \mathscr{T}(\mathscr{R})$ one has $\mathrm{E}q(A_{\tau,n}) = \mathrm{E}Y_\tau$, where $Y_t := \mathrm{E}[q(A_{t,n})|\mathscr{R}_t]$. Indeed,

$$\begin{aligned}\mathrm{E}q(A_\tau) &= \sum_{k=1}^{n} \mathrm{E}q(A_\tau)\mathbf{1}\{\tau = k\} = \sum_{k=1}^{n} \mathrm{E}q(A_k)\mathbf{1}\{\tau = k\} \\ &= \sum_{k=1}^{n} \mathrm{E}\Big[\mathbf{1}\{\tau = k\}\mathrm{E}\{q(A_k)|\mathscr{R}_k\}\Big] = \sum_{k=1}^{n} \mathrm{E}[\mathbf{1}\{\tau = k\}Y_k] = \mathrm{E}Y_\tau,\end{aligned}$$

where we have used the fact that $\{\tau = k\} \in \mathscr{R}_k$. This implies that $\max_{\tau\in\mathscr{T}(\mathscr{R})} \mathrm{E}q(A_{\tau,n}) = \max_{\tau\in\mathscr{T}(\mathscr{R})} \mathrm{E}Y_\tau$. To prove the theorem it suffices to show only that

$$\max_{\tau\in\mathscr{T}(\mathscr{R})} \mathrm{E}Y_\tau = \max_{\tau\in\mathscr{T}(\mathscr{Y})} \mathrm{E}Y_\tau. \tag{8}$$

Clearly,

$$\mathscr{Y}_t \subset \mathscr{R}_t, \quad \forall 1 \le t \le n, \tag{9}$$

Because $R_1, \dots, R_n$ are independent random variables, and $Y_t = I_{t,n}(R_t)$, $\forall t$ we have that for any $s, t \in \{1, \dots, n\}$ with $s < t$

$$\mathrm{P}\{G_t \,|\, \mathscr{Y}_s\} = \mathrm{P}\{G_t \,|\, \mathscr{R}_s\}, \quad \forall G_t \in \mathscr{Y}_t. \tag{10}$$

The statement (8) follows from (9), (10) and Theorem 5.3 of Chow et al. (1971). In fact, (8) is a consequence of the well known fact that randomization does not increase rewards in stopping problems (Chow et al. 1971, Chapter 5). This concludes the proof. ∎

It follows from Theorem 3 that the optimal stopping rule in Problem (A1) is given by Corollary 1 with random variables $\{Y_t\}$ defined by (7). To implement the rule we need to compute the distributions $\{F_t\}$ of the random variables $\{Y_t\}$ and to apply formulas (2) and (3).

Random horizon. Next, we establish a correspondence between Problems (A2) and (B). Let

$$J_t(r) := \sum_{k=t}^{N_{\max}} \gamma_k I_{t,k}(r), \quad r = 1, \dots, t. \tag{11}$$

where $I_{t,k}(\cdot)$ is given in (6), and $\gamma_k = \mathrm{P}(N = k)$. Below in the proof of Theorem 4 we show that

$$J_t(r) = \mathrm{E}\big\{q(A_{t,N})\mathbf{1}\{N \ge t\}|R_1 = r_1, \dots, R_{t-1} = r_{t-1}, R_t = r\big\}.$$

Define also

$$Y_t := J_t(R_t) = \sum_{k=t}^{N_{\max}} \gamma_k I_{t,k}(R_t), \quad t = 1, \dots, N_{\max}. \tag{12}$$

Theorem 4

(i) *Let $N_{\max} < \infty$; then the optimal stopping rule τ_* solving Problem (B) with fixed horizon $N_{\max}$ and random variables $\{Y_t\}$ given in (11)–(12) provides the optimal solution to Problem (A2):*

$$V_\gamma^*(q) = \max_{\tau\in\mathscr{T}(\mathscr{R})} V_\gamma(q;\tau) = \max_{\tau\in\mathscr{T}(\mathscr{Y})} \mathrm{E}Y_\tau = W_{N_{\max}}(\tau^*).$$

(ii) *Let $N_{\max} = \infty$ and assume that*

$$\sup_t \max_{1\le r\le t} \sum_{k=t}^{\infty} \gamma_k |I_{t,k}(r)| < \infty. \qquad (13)$$

Let $\epsilon > 0$ be arbitrary; then there exists $\tilde{N}_{\max} = \tilde{N}_{\max}(\epsilon)$ such that for any stopping rule $\tau \in \mathscr{T}(\mathscr{R})$ one has

$$W_{\tilde{N}_{\max}}(\tau) - \epsilon \le V_\gamma(q;\tau) \le W_{\tilde{N}_{\max}}(\tau) + \epsilon. \qquad (14)$$

In particular, the optimal stopping rule τ_* solving Problem (B) with fixed horizon $\tilde{N}_{\max} = \tilde{N}_{\max}(\epsilon)$ and $\{Y_t\}$ given (11)–(12) is an ϵ–optimal stopping rule for Problem (A2):

$$W_{\tilde{N}_{\max}}(\tau_*) - \epsilon \le V^*_\gamma(q) \le W_{\tilde{N}_{\max}}(\tau_*) + \epsilon \qquad (15)$$

Proof: (i). In Problem (A2) the reward for stopping at time t is $\tilde{q}(A_{t,N}) = q(A_{t,N})\mathbf{1}\{N \ge t\}$. The expectation of the reward conditional on the observations $R_1, \dots, R_t$ until time t is

$$\begin{aligned}
&\mathrm{E}\{q(A_{t,N})\mathbf{1}\{N \ge t\} \,|\, R_1 = r_1, \dots, R_{t-1} = r_{t-1}, R_t = r\} \\
&= \sum_{k=t}^{N_{\max}} \mathrm{E}\big\{q(A_{t,N})\mathbf{1}\{N = k\} \,|\, R_1 = r_1, \dots, R_{t-1} = r_{t-1}, R_t = r\big\} \\
&= \sum_{k=t}^{N_{\max}} \mathrm{E}\big\{\mathbf{1}\{N = k)\}\mathrm{E}\big[q(A_{t,k}) \,|\, N = k, R_t = r\big]\big\} \\
&= \sum_{k=t}^{N_{\max}} \gamma_k \sum_{a=r}^{k-t+r} q(a) \frac{\binom{a-1}{r-1}\binom{k-a}{t-r}}{\binom{k}{t}} = \sum_{k=t}^{N_{\max}} \gamma_k I_{t,k}(r) =: J_t(r). \qquad (16)
\end{aligned}$$

where we have used (4) and (5) with fixed horizon $N = k$, and independence of N and $\{R_t\}$. Together with (12) this implies that $\mathrm{E}\tilde{q}(A_{\tau,N}) = \mathrm{E}J_\tau(R_\tau) = \mathrm{E}Y_\tau$ for any $\tau \in \mathscr{T}(\mathscr{R})$. The remainder of the proof proceeds along the lines of the proof of Theorem 3.

(ii). Let $\tilde{N}_{\max} = \tilde{N}_{\max}(\epsilon)$ be the minimal integer number such that

$$\sup_t \max_{1\le r\le t} \sum_{k=\tilde{N}_{\max}+1}^{\infty} \gamma_k |I_{t,k}(r)| \le \epsilon. \qquad (17)$$

The existence of $\tilde{N}_{\max}(\epsilon)$ follows from (13). In view of (16) and (17) for any stopping rule $\tau \in \mathscr{T}(\mathscr{R})$ we have $V_\gamma(q;\tau) = \mathrm{E}\sum_{k=\tau}^{\infty} \gamma_k I_{\tau,k}(R_\tau)$, and

$$\mathrm{E}\sum_{k=\tau}^{\tilde{N}_{\max}} \gamma_k I_{\tau,k}(R_\tau) - \epsilon \;\le\; V_\gamma(q;\tau) \le \mathrm{E}\sum_{k=\tau}^{\tilde{N}_{\max}} \gamma_k I_{\tau,k}(R_\tau) + \epsilon.$$

This implies (14). In order to prove (15) we note that if $\tilde{\tau}$ is the optimal stopping rule in Problem (A2) then by (14) and definition of τ_*

$$V_\gamma(q;\tilde{\tau}) = V^*_\gamma(q) \le W_{\tilde{N}_{\max}}(\tilde{\tau}) + \epsilon \le W_{\tilde{N}_{\max}}(\tau_*) + \epsilon,$$

which proves the upper bound in (15). On the other hand, in view of (14)

$$V_\gamma^*(q) = V_\gamma(q;\tilde{\tau}) \geq V_\gamma(q;\tau_*) \geq W_{\tilde{N}_{\max}}(\tau_*) - \epsilon.$$

This concludes the proof. ∎

Remark 3 *Condition (13) imposes restrictions on the tail of the distribution of N. It can be easily verified in any concrete setting; for details see Section 5.*

Remark 4 *Theorems 3 and 4 imply that solution of Problems (A1) and (A2) can be obtained by solving Problem (B) with a suitably defined horizon and random variables $\{Y_t\}$ given by (6)–(7) and (11)-(12) respectively. The latter problem is solved by the recursive procedure given in Corollary 1.*

4.3 Specification of the optimal stopping rule for Problems (A1) and (A2)

Now, using Theorems 3 and 4, we specialize the result of Corollary 1 for solution of Problems (A1) and (A2). For this purpose we require the following notation:

$$\nu := \begin{cases} n, & \text{Problem (A1)}, \\ N_{\max} \text{ or } \tilde{N}_{\max}, & \text{Problem (A2)}, \end{cases} \qquad U_t(r) := \begin{cases} I_{t,n}(r), & \text{Problem (A1)}, \\ J_t(r), & \text{Problem (A2)}. \end{cases}$$

Note that in Problem (A2) we put $\nu = N_{\max}$ for distributions with the finite right endpoint $N_{\max} < \infty$; otherwise $\nu = \tilde{N}_{\max}$, where $\tilde{N}_{\max}$ is defined in the proof of Theorem 4. With this notation Problem (B) is associated with independent random variables $Y_t = U_t(R_t)$ for $t = 1, \ldots, \nu$.

Let $y_t(1), \ldots, y_t(\ell_t)$ denote distinct points of the set $\{U_t(1), \ldots, U_t(t)\}$, $t = 1, \ldots, \nu$. The distribution of the random variable Y_t is supported on the set $\{y_t(1), \ldots, y_t(\ell_t)\}$ and given by

$$f_t(j) \;:=\; \mathrm{P}\{Y_t = y_t(j)\} = \frac{1}{t}\sum_{r=1}^{t} \mathbf{1}\big\{U_t(r) = y_t(j)\big\}, \quad j = 1, \ldots, \ell_t, \tag{18}$$

$$F_t(z) \;=\; \sum_{j=1}^{\ell_t} f_t(j)\mathbf{1}\{y_t(j) \leq z\}, \quad z \in \mathbb{R}. \tag{19}$$

The following statement is an immediate consequence of Corollary 1 and formulas (18)–(19).

Corollary 2 *Let $\tau_* = \min\{1 \leq t \leq \nu : Y_t > b_{\nu-t+1}\}$, where the sequence $\{b_t\}$ is given by*

$$b_1 = -\infty, \;\; b_2 = \sum_{j=1}^{\ell_\nu} y_\nu(j) f_\nu(j), \tag{20}$$

$$b_{t+1} = \sum_{j=1}^{\ell_{\nu-t+1}} \big[\, b_t \vee y_{\nu-t+1}(j)\,\big]\, f_{\nu-t+1}(j), \quad t = 2, \ldots, \nu. \tag{21}$$

Then

$$\mathrm{E}Y_{\tau_*} = \sup_{\tau \in \mathscr{T}(\mathscr{R})} \mathrm{E}Y_\tau = b_{\nu+1}.$$

Proof: In view of (7) and (12) , $Y_1, \dots, Y_\nu$ are independent random variables; therefore Corollary 1 is applicable. We have

$$\begin{aligned}\int_{b_t}^{\infty} z \mathrm{d}F_{\nu-t+1}(z) &= \sum_{j=1}^{\ell_{\nu-t+1}} y_{\nu-t+1}(j)\mathbf{1}\{y_{\nu-t+1}(j) > b_t\} f_{\nu-t+1}(j),\\ b_t F_{\nu-t+1}(b_t) &= b_t \sum_{j=1}^{\ell_{\nu-t+1}} f_{\nu-t+1}(j)\mathbf{1}\{y_{\nu-t+1}(j) \le b_t\}.\end{aligned}$$

Summing up these expressions we come to (21). ∎

As we have already mentioned, in the considered problems the optimal stopping rule belongs to the class of memoryless threshold policies. This facilitates derivation of the distributions of the corresponding stopping times, and calculation of their probabilistic characteristics. One of the important characteristics is the expected time elapsed before stopping. In problems with fixed horizon $\nu = n$ it is given by the following formula

$$\begin{aligned}\mathrm{E}\,(\tau_*) &= \sum_{i=0}^{n-1} \mathrm{P}(\tau_* > i) = 1 + \sum_{i=1}^{n-1} \mathrm{P}(\tau_* > i)\\ &= 1 + \sum_{i=1}^{n-1}\prod_{t=1}^{i} \mathrm{P}\left(Y_t \le b_{n-t+1}\right) = 1 + \sum_{i=1}^{n-1}\prod_{t=1}^{i} F_t\left(b_{n-t+1}\right),\end{aligned} \tag{22}$$

where $\{F_t\}$ and $\{b_t\}$ are defined in (19) and (20)–(21).

In the problems where the horizon N is random, the time until stopping is $\tau_* \wedge N$. In this case

$$\mathrm{E}(\tau_* \wedge N) = \mathrm{E}\tau_*\mathbf{1}\{\tau_* \le N\} + \mathrm{E}N\mathbf{1}\{\tau_* > N\}, \tag{23}$$

where

$$\begin{aligned}\mathrm{E}[\tau_*\mathbf{1}\{\tau_* \le N\}] &= \mathrm{E}\Big(\tau_* \sum_{k=\tau_*}^{N_{\max}} \gamma_k\Big) = \sum_{j=1}^{N_{\max}} j \sum_{k=j}^{N_{\max}} \gamma_k \mathrm{P}(\tau_* = j)\\ &= \sum_{k=2}^{N_{\max}} \gamma_k(1 - F_1(b_{N_{\max}})) + \sum_{k=2}^{N_{\max}} \gamma_k \sum_{j=2}^{k} j(1 - F_j(b_{N_{\max}-j+1})) \prod_{t=1}^{j-1} F_t(b_{N_{\max}-t+1})\end{aligned} \tag{24}$$

and

$$\mathrm{E}[N\mathbf{1}(N < \tau_*)] = \sum_{k=1}^{N_{\max}} k\gamma_k \prod_{t=1}^{k} F_t(b_{N_{\max}-t+1}). \tag{25}$$

4.4 Implementation

In this section we present an efficient algorithm implementing the optimal stopping rule described earlier. In order to implement (20)–(21) we need to find the sets $\{y_t(j), j = 1, \dots, \ell_t\}$ in which random variables Y_t, $t = 1, \dots, \nu$ take values, and to compute the corresponding probabilities $\{f_t(j),\, j = 1, \dots, \ell_t\}$.

The following algorithm implements the optimal policy.

Algorithm 1.

1. Compute

$$U_t(r) = \sum_{a=r}^{\nu-t+r} q(a) \frac{\binom{a-1}{r-1}\binom{\nu-a}{t-r}}{\binom{\nu}{t}}, \quad r=1,\ldots,t; \quad t=1,\ldots,\nu.$$

   We note that the computations can be efficiently performed using the following recursive formula: for any reward function q

$$U_t(r) = \frac{r}{t+1} U_{t+1}(r+1) + \Big(1 - \frac{r}{t+1}\Big) U_{t+1}(r), \quad r=1,\ldots,t; \tag{26}$$

   see Gusein–Zade (1966) and Mucci (1973, Proposition 2.1).

2. Find the distinct values $(y_t(1),\ldots,y_t(\ell_t))$ of the vector $(U_t(1),\ldots,U_t(t))$, $t=1,\ldots,\nu$; here ℓ_t is a number of the distinct points.

3. Compute

$$f_t(j) = \frac{1}{t}\sum_{r=1}^{t} \mathbf{1}\big\{U_t(r) = y_t(j)\big\}, \quad j=1,\ldots,\ell_t; \quad t=1,\ldots,\nu.$$

4. Let $b_1 = -\infty$, $b_2 = \sum_{j=1}^{\ell_\nu} y_\nu(j) f_\nu(j)$.

   For $t=2,\ldots,\nu$ compute

$$b_{t+1} = \sum_{j=1}^{\ell_{\nu-t+1}} \big[\, b_t \vee y_{\nu-t+1}(j) \,\big]\, f_{\nu-t+1}(j). \tag{27}$$

5. Output $b_{\nu+1}$ and $\tau_* = \min\{t \in \{1,\ldots,\nu\} : U_t(R_t) > b_{\nu-t+1}\}$.

5 Solution of the sequential selection problems

In this section we revisit problems (P1)–(P11) discussed earlier from the viewpoint of the proposed framework. We refer to Section 2 for detailed description of these problems and related literature.

5.1 Problems with fixed horizon

First we consider problems (P1)-(P6) with fixed horizon; in all these problems $\nu = n$.

5.1.1 Classical secretary problem

For description of this problem and related references see Problem (P1) in Section 2. Here $q(a) = \mathbf{1}\{a=1\}$, and

$$U_t(r) = I_{t,n}(r) = \frac{t}{n}\mathbf{1}\{r=1\}, \quad r=1,\ldots,t; \quad \ell_t = 2, \quad t=1,\ldots,n.$$

The random variable $Y_t = (t/n)\mathbf{1}\{R_t = 1\} = \mathrm{P}(A_{t,n} = 1|R_t)$ takes two different values $y_t(1) = t/n$, $y_t(2) = 0$ with probabilities $f_t(1) = 1/t$ and $f_t(2) = 1 - (1/t)$. Then Step 4 of the Algorithm 1 takes the form: $b_1 = -\infty$, $b_2 = 1/n$,

$$b_{t+1} = b_t + \Big(\frac{1}{n} - \frac{b_t}{n-t+1}\Big)\mathbf{1}\Big\{b_t < \frac{n-t+1}{n}\Big\},\ t = 2, \dots, n.$$

The optimal policy is to stop the first time instance t such that $Y_t > b_{n-t+1}$, i.e.,

$$\tau_* = \min\Big\{1 \le t \le n : \frac{t}{n}\mathbf{1}\{R_t = 1\} > b_{n-t+1}\Big\},$$

which coincides with well known results.

5.1.2 Selecting one of k best alternatives

This setting is stated as Problem (P2) in Section 2. In this problem $q(a) = \mathbf{1}\{a \le k\}$ with some $k \le n$. We will assume here that $k \ge 2$; the case $k = 1$ was treated above.

We have

$$U_t(r) = \begin{cases} 0, & k+1 \le r \le t, \\ \sum_{a=r}^{(n-t+r)\wedge k} \frac{\binom{a-1}{r-1}\binom{n-a}{t-r}}{\binom{n}{t}}, & 1 \le r \le k, \end{cases} \qquad t = 1, \dots, n. \tag{28}$$

It is easily checked that for $q(a) = \mathbf{1}\{a \le k\}$ one has

$$U_n(r) = \begin{cases} 1, & r = 1, \dots, k \\ 0, & r = k+1, \dots, n. \end{cases} \tag{29}$$

Using this formula together with the recursive relationship (26) we can determine the structure of vector $U_t := (U_t(1), \dots, U_t(t))$ for each $t = 1, \dots, n$, and compute $\{y_t(j)\}$ and $\{f_t(j)\}$. Specifically, the following facts are easily verified.

(a) Let $n - k + 2 \le t \le n$. Here vector U_t has the following structure: the first $t + k - n$ components are ones, the next $n - t$ components are distinct numbers in $(0, 1)$ which are given in (28), and the last $t - k$ components are zeros. Formally, if $n - k + 2 \le t \le n - 1$ and $k > 2$ then we have

$$U_t(j) = \begin{cases} 1, & j = 1, \dots, k-n+t, \\ \in (0,1), & j = k-n+t+1, \dots, k, \\ 0, & j = k+1, \dots, t, \end{cases}$$

Note that if $k = 2$ the regime reduces to $t = n$; therefore if $k = 2$ or $t = n$ then U_n is given by (29). These facts imply the following expressions for $\{y_t(j)\}$ and $\{f_t(j)\}$:

$$\ell_t = n - t + 2; \quad y_t(j) = \begin{cases} 1, & j = 1, \\ U_t(k-n+t+j), & j = 2, \dots, n-t+1, \\ 0, & j = n-t+2, \end{cases} \tag{30}$$

and

$$f_t(j) = \begin{cases} 1 - (n-k)/t, & j = 1, \\ 1/t, & j = 2, \dots, n-t+1, \\ 1 - k/t, & j = n-t+2. \end{cases} \tag{31}$$

n	k	$P(n,k)$	$E(n,k)/n$	n	k	$P(n,k)$	$E(n,k)/n$	n	k	$P(n,k)$	$E(n,k)/n$
100	2	0.57956	0.68645	500	2	0.57477	0.68886	1,000	2	0.57417	0.68966
	5	0.86917	0.60871		5	0.86211	0.60921		5	0.86123	0.60988
	10	0.98140	0.54236		10	0.97754	0.54454		10	0.97703	0.54434
	15	0.99755	0.50428		15	0.99627	0.50845		15	0.99609	0.50893
5,000	2	0.57369	0.68931	10,000	2	0.57363	0.68927	50,000	2	0.57358	0.68923
	5	0.86052	0.61015		5	0.86043	0.61014		5	0.86036	0.61018
	10	0.97663	0.54499		10	0.97658	0.54496		10	0.97654	0.54500
	15	0.99594	0.50943		15	0.99592	0.50947		15	0.99591	0.50950

Table 1: Optimal probabilities $P(n,k)$ and the normalized expected time elapsed until stopping $E(n,k)/n$ for selecting one of the k best values.

If $t = n$ then

$$\ell_t = 2,\ y_n(1) = 1,\ y_n(2) = 0,\ f_n(1) = k/n,\ f_n(2) = 1 - k/n.$$

(b) If $k+2 \le t \le n-k+1$ then the set $\{U_t(1), \ldots, U_t(t)\}$ contains $k+1$ distinct values: $U_t(1), \ldots, U_t(k)$ are positive distinct, and $U_t(k+1) = \cdots = U_t(t) = 0$. Therefore

$$\ell_t = k+1; \quad y_t(j) = \begin{cases} U_t(j), & j = 1, \ldots, k \\ 0, & j = k+1; \end{cases} \qquad f_t(j) = \begin{cases} 1/t, & j = 1, \ldots, k, \\ 1 - k/t, & j = k+1. \end{cases} \tag{32}$$

(c) If $1 \le t \le k+1$ then all the values $U_t(1), \ldots, U_t(t)$ are positive and distinct. Thus

$$\ell_t = t; \quad y_t(j) = U_t(j),\ j = 1, \ldots, t; \quad f_t(j) = \frac{1}{t},\ \ j = 1, \ldots, t. \tag{33}$$

In our implementation we compute $U_t(j)$ for $t = 1, \ldots, n$ and $j = 1, \ldots, t$ using (29) and (26). Then $\{y_t(j)\}$, $\{f_t(j)\}$ and the sequence $\{b_t\}$ are easily calculated from (30)–(33) and (27) respectively.

Table 1 presents exact values of the optimal probability $P(n,k) = b_{n+1}$ and the expected time until stopping $E(n,k) = \mathrm{E}(\tau_*)$ normalized by n for different values of k and n. We are not aware of works that report exact results for general k and n as presented in Table 1. It is worth noting that the optimal policy developed by Gusein–Zade (1966) is expressed in terms of of relative ranks. In contrast, our policy is expressed via the random variables $Y_t = U_t(R_t)$, and it is memoryless threshold in terms of $\{Y_t\}$. This allows to efficiently compute the distribution of the optimal stopping time, and, in particular, the expected time until stopping. The value of $E(n,k)$ is computed using formula (22) combined with (18) and (28)–(33). The presented numbers agree with asymptotic results of Yeo (1997) proved for $k = 2, 3$ and 5.

5.1.3 Selecting the k-th best alternative

This setting is discussed in Section 2 as problem (P3). In this problem $q(a) = \mathbf{1}\{a = k\}$, $k \ge 2$. Similarly to the Gusein–Zade stopping problem, here we have three different regimes that define explicit relations for $\{U_t(r)\}$, $\{y_t(j)\}$ and $\{f_t(j)\}$.

(a) Let $1 \le t \le k$; then

$$U_t(r) = \frac{\binom{k-1}{r-1}\binom{n-k}{t-r}}{\binom{n}{t}}, \; r = 1, \dots, t.$$

All values of $U_t(1), \dots, U_t(t)$ are positive and distinct. Thus

$$\ell_t = t, \;\; y_t(j) = U_t(j), \;\; f_t(j) = \frac{1}{t}, \; 1 \le j \le t. \tag{34}$$

(b) If $k+1 \le t \le n-k+1$ then

$$U_t(r) = \begin{cases} \frac{\binom{k-1}{r-1}\binom{n-k}{t-r}}{\binom{n}{t}}, & 1 \le r \le k, \\ 0, & k+1 \le r \le t. \end{cases}$$

The set $\{U_t(1), \dots, U_t(t)\}$ contains $k+1$ distinct values: $U_t(1), \dots, U_t(k)$ are positive distinct, and $U_t(k+1) = \dots = U_t(t) = 0$. Therefore,

$$\ell_t = k+1; \quad y_t(j) = \begin{cases} U_t(j), & j = 1, \dots, k \\ 0, & j = k+1; \end{cases} \qquad f_t(j) = \begin{cases} 1/t, & j = 1, \dots, k, \\ 1 - k/t, & j = k+1. \end{cases} \tag{35}$$

(c) Let $n-k+2 \le t \le n$; then the sequence $\{U_t(r)\}$ takes the following values

$$U_t(r) = \begin{cases} 0, & r = 1, \dots, t-n+k-1, \\ \frac{\binom{k-1}{r-1}\binom{n-k}{t-r}}{\binom{n}{t}}, & r = t-n+k, \dots, k, \\ 0, & r = k+1, \dots, t. \end{cases}$$

Therefore,

$$\ell_t = n-t+2; \quad y_t(j) = \begin{cases} 0, & j = 1 \\ U_t(t-(n-k)-2+j), & j = 2, \dots, n-t+2. \end{cases} \tag{36}$$

and, correspondingly,

$$f_t(j) = \begin{cases} (2t-n-1)/t, & j = 1, \\ 1/t, & j = 2, \dots, n-t+2. \end{cases} \tag{37}$$

Table 2 presents optimal probabilities of selecting kth best alternative for a range of k and n. In the specific case of $k = 2$ Rose (1982a) showed that the optimal stopping rule is

$$\tau_* = \min\Big\{\{t \ge \lceil n/2 \rceil : R_t = 2\} \cup \{n\}\Big\},$$

and the optimal probability is $P(n,2) = \frac{n+1}{4n}$. The results for $k = 2$ in Table 2 are in full agreement with this formula. The table also presents numerical computation of optimal values in the problem of selecting the median value; see Rose (1982b) who proved that $\lim_{n\to\infty} V_n^*(q_{\rm pd}^{((n+1)/2)}) = 0$.

n	k	$P(n,k)$	$E(n,k)/n$	n	k	$P(n,k)$	$E(n,k)/n$	n	k	$P(n,k)$	$E(n,k)/n$
101	2	0.25247	0.82995	501	2	0.25050	0.75466	1,001	2	0.25025	0.74984
	5	0.19602	0.78968		5	0.19281	0.78890		5	0.19241	0.78896
	10	0.15962	0.84827		10	0.15506	0.84508		10	0.15451	0.84517
	50	0.11467	0.86699		250	0.06876	0.91156		500	0.05504	0.92688
5,001	2	0.25005	0.84527	10,001	2	0.25002	0.75453	50,001	2	0.25000	0.83830
	5	0.19210	0.78896		5	0.19206	0.78891		5	0.19203	0.78891
	10	0.15450	0.84478		10	0.15402	0.84477		10	0.15397	0.84477
	2,500	0.03265	0.95443		5,000	0.02603	0.96320		25,000	0.01533	0.97787

Table 2: Optimal probabilities $P(n,k)$ and the normalized expected time elapsed until stopping $E(n,k)/n$ for selecting the k-th best alternative computed using (34)–(37).

5.1.4 Expected rank type problems

In this section we consider problems (P4) and (P5) discussed in Section 2.

Expected rank minimization. Following (1) we consider the problem of minimization of $\mathrm{E}q(A_{\tau,n})$, where $q(a) = -a$. It is well known that $\mathrm{E}\big[A_{t,n}|R_t = r\big] = (n+1)r/(t+1)$; therefore for $t = 1, \dots, n$

$$U_t(r) = I_{t,n}(r) = \mathrm{E}[q(A_{t,n})|R_t = r] = -\mathrm{E}[A_{t,n}|R_t = r] = -\frac{(n+1)r}{t+1}, \quad r = 1, \dots, t.$$

In this setting

$$\ell_t = t, \ \ \forall t; \quad y_t(j) = U_t(j) = -\frac{n+1}{t+1} j, \ \ j = 1, \dots t; \quad f_t(j) = \frac{1}{t}, \quad \forall j = 1, \dots, t.$$

Substitution to (21) yields $b_1 = -\infty$, $b_2 = -\frac{1}{2}(n+1)$,

$$b_{t+1} = \frac{1}{n-t+1} \sum_{j=1}^{n-t+1} \Big[b_t \vee \Big(-\frac{n+1}{n-t+2} j \Big) \Big], \quad t = 2, \dots, n. \tag{38}$$

Straightforward calculation shows that (38) takes form

$$b_{t+1} = b_t - \frac{1}{n-t+1} \left[\frac{n+1}{n-t+2} \frac{j_t(j_t+1)}{2} + j_t b_t \right], \quad t = 2, \dots, n.$$

where $j_t := \lfloor -b_t \frac{n-t+2}{n+1} \rfloor$. The optimal policy is to stop the first time instance t such that $Y_t > b_{n-t+1}$, i.e.,

$$\tau_* = \min \Big\{ 1 \le t \le n : -\frac{n+1}{t+1} R_t > b_{n-t+1} \Big\} = \min \Big\{ 1 \le t \le n : R_t < j_{n-t+1} \Big\}.$$

Then according to (1) the optimal value of the problem equals to $-b_{n+1}$. We note that the derived recursive procedure coincides with the one of Chow et al. (1964), and the calculation for $n = 10^6$ yields the optimal value $3.86945\dots$

Expected squared rank minimization. This problem was posed in Robbins (1991), and to the best of our knowledge, it was not solved to date. We show that the proposed unified framework can be used in order to compute efficiently the optimal policy and its value.

In this setting $U_t(r) = I_{t,n}(r)$, and the reward is given by $q(a) = -a^2$. It is well known that

$$\mathrm{E}\big[A_{t,n}(A_{t,n}+1)\cdots(A_{t,n}+k-1)\,|\,R_t = r\big] = \frac{(n+1)\cdots(n+k)}{(t+1)\cdots(t+k)} r\cdots(r+k-1);$$

see, e.g., Robbins (1991). Therefore we put

$$U_t(r) = -\mathrm{E}(A_{t,n}^2|R_t = r) = -\frac{(n+1)(n+2)}{(t+1)(t+2)} r\Big(r + \frac{n-t}{n+2}\Big).$$

In this case

$$\ell_t = t,\;\; y_t(j) = U_t(j) = -\frac{(n+1)(n+2)}{(t+1)(t+2)} j\Big(j + \frac{n-t}{n+2}\Big),\;\; f_t(j) = \frac{1}{t},\quad j = 1,\ldots,t.$$

Substituting this to (21) we obtain the following recursive relationship: $b_1 = -\infty$, $b_2 = -\frac{1}{6}(n+1)(2n+1)$,

$$b_{t+1} \;=\; \frac{1}{n-t+1}\sum_{j=1}^{n-t+1}\Big\{b_t \vee \Big[-\frac{(n+1)(n+2)}{(n-t+2)(n-t+3)} j\Big(j + \frac{t-1}{n+2}\Big)\Big]\Big\}.$$

Denote $j_t := \max\{1 \le j \le n-t+1 : b_t \le -j^2 C_{n,t} - jD_{n,t}\}$, where

$$C_{n,t} = \frac{(n+1)(n+2)}{(n-t+2)(n-t+3)},\quad D_{n,t} = \frac{(t-1)(n+1)}{(n-t+2)(n-t+3)}.$$

Then

$$\begin{aligned} j_t = \max\Big\{1 \le j \le n-t+1 : j \le \frac{1}{2C_{n,t}}\Big(-D_{n,t} + \sqrt{D_{n,t}^2 - 4C_{n,t}b_t}\Big)\Big\} \\ = \Big\lfloor \frac{1}{2C_{n,t}}\Big(-D_{n,t} + \sqrt{D_{n,t}^2 - 4C_{n,t}b_t}\Big)\Big\rfloor. \end{aligned}$$

With this notation we have $b_1 = -\infty$, $b_2 = -\frac{1}{6}(n+1)(2n+1)$, and for $t = 2,\ldots,n$

$$b_{t+1} = \frac{1}{n-t+1}\Big[-\frac{1}{6}j_t(j_t+1)(2j_t+1)C_{n,t} - \frac{1}{2}j_t(j_t+1)D_{n,t} + (n-t+1-j_t)b_t\Big]. \quad (39)$$

The optimal policy is to stop the first time instance t such that $Y_t > b_{n-t+1}$ which is equivalent to

$$\tau_* = \min\Big\{1 \le t \le n : R_t < j_{n-t+1}\Big\}.$$

Table 3 presents optimal values $V_*(n) := \mathrm{E}A^2_{\tau_*,n}$ computed with recursive relation (39) for different n.

n	100	250	500	750	1,000	2,500
$V_*(n)$	23.70663	26.49268	27.66697	28.10937	28.34466	28.80553
n	5,000	10,000	20,000	10^5	10^6	10^8
$V_*(n)$	28.97697	29.06969	29.11944	29.16302	29.17431	29.17579

Table 3: Optimal values of $V_*(n) := \mathrm{E}A^2_{\tau_*,n}$ computed using (39).

5.2 Problems with random horizon

This section demonstrates how to apply the proposed framework for solution of selection problems with a random horizon. In these problems we apply Algorithm 1 with ν being the maximal horizon length $N_{\max}$, provided that $N_{\max}$ is finite, or with sufficiently large horizon $\tilde{N}_{\max}$ if $N_{\max}$ is infinite. Moreover, $U_t(r) = J_t(r)$, where $\{J_t(r)\}$ is given by (11).

5.2.1 Classical secretary problem with random horizon

In this setting $q(a) = \mathbf{1}\{a = 1\}$; therefore

$$I_{t,k}(r) = \mathrm{P}(A_{t,N} = 1 \,|\, N = k, R_t = r) = \frac{t}{k}\mathbf{1}\{r = 1\}, \quad k \geq t,$$
$$U_t(r) = J_t(r) = \sum_{k=t}^{N_{\max}} \gamma_k I_{t,k}(r) = t\,\mathbf{1}\{r = 1\} \sum_{k=t}^{N_{\max}} \frac{\gamma_k}{k}\ .$$

Note that if $N_{\max} = \infty$ then condition (13) is trivially fulfilled since

$$t\sum_{k=t}^{\infty} \frac{\gamma_k}{k} \leq \sum_{k=t}^{\infty} \gamma_k \leq 1.$$

The random variables $Y_t = U_t(R_t) = \mathbf{1}(R_t = 1)\, t\, \sum_{k=t}^{\nu} \gamma_k/k$ take two different values $y_t(1) = t\sum_{k=t}^{\nu} \gamma_k/k$ and $y_t(2) = 0$ with corresponding probabilities $f_t(1) = 1/t$ and $f_t(2) = 1-1/t$. Substituting these values in (27) we obtain $b_1 = -\infty$, $b_2 = \gamma_\nu/\nu$, and for $t = 2, \ldots, \nu$

$$b_{t+1} = b_t + \bigg(\sum_{k=\nu-t+1}^{\nu} \frac{\gamma_k}{k} - \frac{b_t}{\nu - t + 1}\bigg)\mathbf{1}\bigg\{b_t < (\nu - t + 1)\sum_{k=\nu-t+1}^{\nu} \frac{\gamma_k}{k}\bigg\}. \tag{40}$$

The optimal policy is to stop at time t if $Y_t > b_{\nu-t+1}$, i.e.,

$$\tau_* = \bigg\{t = 1, \ldots, \nu :\, \mathbf{1}\{R_t = 1\}\, t \sum_{k=t}^{\nu} \frac{\gamma_k}{k} > b_{\nu-t+1}\bigg\}. \tag{41}$$

Presman and Sonin (1972) investigated the structure of optimal stopping rules and showed that, depending on the distribution of N, the stopping region can involve several "islands," i.e., it can be a union of disjoint subsets of $\{1, \ldots, N_{\max}\}$. Note that (41) determines the stopping region automatically. Indeed, it is optimal to stop only at those t's that satisfy $t\sum_{k=t}^{\nu} \gamma_k/k > b_{\nu-t+1}$. We apply the stopping rule (40)–(41) for two examples of distributions of N. In the first example N is assumed to be uniformly distributed on the set $\{1, \ldots, N_{\max}\}$. As it is known, in this case the optimal stopping region has only one

$N_{\max} \mid n$	10	20	40	60	80	10^2	10^3	10^5
$V_*(N_{\max})$	0.35145	0.30760	0.28889	0.28260	0.27949	0.27779	0.27137	0.27067
$E_*(N_{\max})$	0.29290	0.26227	0.280651	0.28605	0.27410	0.27410	0.27995	0.27983
$E_*(n)$	0.61701	0.73421	0.75074	0.73988	0.73436	0.74104	0.73620	0.73576

Table 4: Optimal values $V_*(N_{\max}) := \mathrm{P}\{A_{\tau_*,N} = 1, \tau_* \leq N\}$ for a uniformly distributed horizon length N, normalized expected times until stopping $E_*(N_{\max})$ and $E_*(n)$ for random and fixed horizons.

"island." The second example illustrates a setting in which the stopping region has more than one "island."

1. *Uniform distribution.* In this case $\nu = N_{\max}$, $\gamma_k = 1/N_{\max}$, $k = 1, \ldots, N_{\max}$. It was shown in Presman and Sonin (1972) that the optimal stopping region in this problem has one "island," i.e., the optimal policy selects the first best member appearing in the range $\{k_n, \ldots, n\}$. The recursive relation (40) with $\gamma_k = 1/N_{\max}$, $k = 1, \ldots, N_{\max}$ yields the optimal values $V_*(N_{\max}) := \mathrm{P}\{A_{\tau_*,N} = 1, \tau_* \leq N\}$ given in Table 4. The second line of Table 4 presents the normalized expected time until stopping $E_*(N_{\max}) := \mathrm{E}(\tau_* \wedge N_{\max})/N_{\max}$ computed using (23), (24) and (25). For comparison, we also give the normalized expected time elapsed until stopping $E_*(n) := \mathrm{E}\tau_*/n$ for the optimal stopping rule in the classical secretary problem (see the third line of the table). These numbers are calculated using (22). As expected, $E_*(N_{\max})$ is significantly smaller than $E_*(n)$; the optimal rule is more cautious when the horizon is random.

It was also shown in Presman and Sonin (1972) that $\lim_{N_{\max}\to\infty} V_*(N_{\max}) = 2e^{-2} = 0.27067\ldots$. Note that the numbers in Table 4 are in full agreement with these results. Figure 1(a) displays the sequences $\{b_{N_{\max}-t+1}\}$ and $\left\{t \sum_{k=t}^{N_{\max}} \gamma_k/k\right\}$ for the uniform distribution for $N_{\max} = 100$. Note the stopping region is the set of t's where the blue curve is above the red curve. Thus, there is only one "island" in this case.

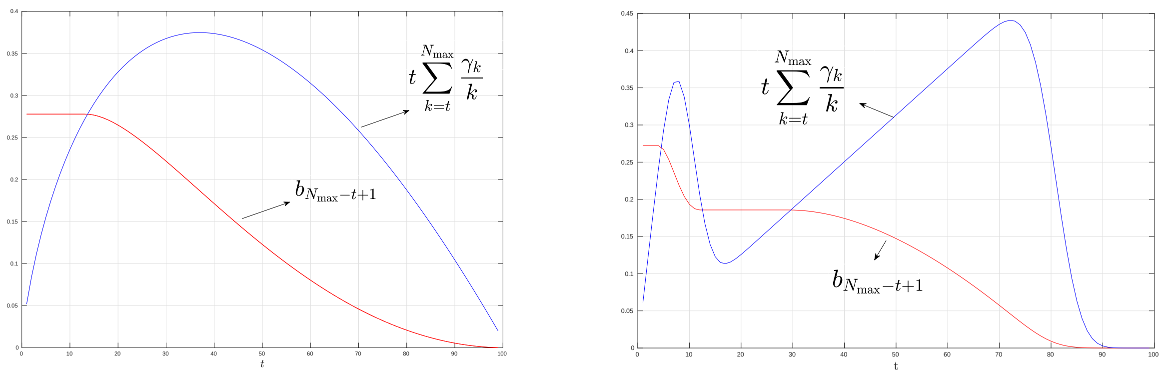
2. *Mixture of two zero–inflated binomial distributions.* Here we assume that the distribution G_N of N is the mixture: $G_N(x) = \frac{1}{2}H_1(x) + \frac{1}{2}H_2(x)$, where $H_i(x) = \mathrm{P}(X_i \leq x | X_i \geq 1)$, $i = 1, 2$, and $X_1 \sim \mathrm{Bin}(50, 0.2)$, $X_2 \sim \mathrm{Bin}(100, 0.8)$. In other words, for $k = 1, \ldots, 100$

$$\gamma_k = \mathrm{P}(N = k) = \frac{1}{2}\binom{50}{k}\left(\frac{1}{4}\right)^k \frac{(0.8)^{50}}{1 - (0.8)^{50}} + \frac{1}{2}\binom{100}{k} 4^k \frac{(0.2)^{100}}{1 - (0.2)^{100}}.$$

The optimal stopping rule is given by (40)–(41) with $\{\gamma_k\}$ indicated above. Figure 1(b) displays the graphs of the sequences $\{b_{N_{\max}-t+1}\}$ and $\left\{t \sum_{k=t}^{N_{\max}} \gamma_k/k\right\}$. It is clearly seen that in this setting the stopping region is a union of two disjoint sets of subsequent integer numbers. These sets correspond to the indices where the graph of $\left\{t \sum_{k=t}^{N_{\max}} \gamma_k/k\right\}$ is above the graph of $\{b_{N_{\max}-t+1}\}$. The stopping region can be easily identified from given formulas.

5.2.2 Expected rank minimization over random horizon

In this setting we would like to minimize the expected absolute rank on the event that the stopping occurs before N; otherwise we receive the absolute rank of the last available



(a) (b)

Figure 1: The graphs of sequences $\{b_{N_{\max}-t+1}\}$ and $\{t\sum_{k=t}^{N_{\max}}\gamma_k/k\}$ for different distributions of N: (a) the uniform distribution; (b) the mixture of two zero–inflated binomial distributions.

observation, $A_{N,N} = R_N$. Formally, the corresponding stopping problem is

$$\begin{aligned} V_*(N_{\max}) &:= \min_{\tau\in\mathscr{T}(\mathscr{R})} \mathrm{E}\big[A_{\tau,N}\mathbf{1}\{N\geq\tau\} + R_N\mathbf{1}\{N<\tau\}\big] \\ &= -\max_{\tau\in\mathscr{T}(\mathscr{R})} \mathrm{E}\big[(R_N - A_{\tau,N})\mathbf{1}\{N\geq\tau\} - R_N\big] \\ &= -\max_{\tau\in\mathscr{T}(\mathscr{R})} \mathrm{E}\big[(R_N - A_{\tau,N})\mathbf{1}\{N\geq\tau\}\big] + \frac{1}{2}(1+\mathrm{E}N). \end{aligned}$$

Thus, letting $q(A_{t,N}) = R_N - A_{t,N}$ we note that

$$I_{t,k}(r) = \mathrm{E}\big[q(A_{t,N})\,|\,N=k, R_1=r_1,\ldots,R_{t-1}=r_{t-1}, R_t=r\big] = \frac{1}{2}(k+1) - \frac{k+1}{t+1}r$$

and therefore

$$U_t(r) = J_t(r) = \sum_{k=t}^{N_{\max}} \gamma_k I_{t,k}(r) = \Big(\frac{1}{2} - \frac{r}{t+1}\Big)\sum_{k=t}^{N_{\max}}(k+1)\gamma_k.$$

If $N_{\max}=\infty$ then we require that $\mathrm{E}N<\infty$; this ensures condition (13).

In this setting $\nu = N_{\max}$ or $\nu = \tilde{N}_{\max}$ depending on support of the distribution of N, and

$$y_t(j) = \Big(\frac{1}{2} - \frac{j}{t+1}\Big)\sum_{k=t}^{\nu}(k+1)\gamma_k, \quad f_t(j) = \frac{1}{t}, \quad j=1,\ldots,t, \quad t=1,\ldots,\nu.$$

The recursion for computation of the optimal value is obtained by substitution of these

$N_{\max}$	100	500	10^3	10^4	10^5	10^6
$\alpha = 1$	4.74437	8.42697	10.70615	23.34298	50.43062	108.71663
$\alpha = 2$	3.83593	4.14133	4.18918	4.23792	4.24381	4.24444
$\alpha = 3$	3.61069	3.80588	3.83549	3.86542	3.86909	3.86947

Table 5: Optimal values $V_*(N_{\max})$ computed using (42).

formulas in (27): $b_1 = -\infty$, $b_2 = 0$, and for $t = 2, \dots, \nu$

$$\begin{aligned} b_{t+1} &= \frac{1}{\nu - t + 1} \sum_{j=1}^{\nu-t+1} \Big[b_t \vee \Big(\frac{1}{2} - \frac{j}{\nu - t + 2}\Big) \sum_{k=\nu-t+1}^{\nu} (k+1)\gamma_k \Big]. \\ &= b_t + \frac{1}{\nu - t + 1} \sum_{j=1}^{\nu-t+1} \Big[\Big(\frac{1}{2} - \frac{j}{\nu - t + 2}\Big) \sum_{k=\nu-t+1}^{\nu} (k+1)\gamma_k - b_t \Big]_+ . \end{aligned} \tag{42}$$

The optimal policy is to stop at time t if $Y_t = U_t(R_t) > b_{\nu-t+1}$, i.e.,

$$\tau_* = \Big\{ t = 1, \dots, \nu : \Big(\frac{1}{2} - \frac{R_t}{t+1}\Big) \sum_{k=\nu-t+1}^{\nu} (k+1)\gamma_k > b_{\nu-t+1} \Big\}.$$

Note that $V_*(N_{\max}) = b_{N_{\max}+1} + \frac{1}{2}(1 + \mathrm{E}N)$.

Gianini-Pettitt (1979) considered distributions of N with finite right endpoint $N_{\max}$ and studied asymptotic behavior of the optimal value $V_*(N_{\max})$ as $N_{\max} \to \infty$. In particular, for distributions satisfying $\mathrm{P}(N = k | N \geq k) = (N_{\max} - k + 1)^{-\alpha}$, $k = 1, \dots, N_{\max}$, $N_{\max} = 1, 2, \dots$ with $\alpha > 0$ one has: (a) if $\alpha < 2$ then $V_*(N_{\max}) \to \infty$ as $N_{\max} \to \infty$; (b) if $\alpha > 2$ then $\lim_{N_{\max}\to\infty} V_*(N_{\max}) = 3.86945\dots$; (c) if $\alpha = 2$ then $\limsup_{N_{\max}\to\infty} V_*(N_{\max})$ is finite and greater than $3.86945\dots$. Thus, if $\alpha > 2$ then the optimal value $V_*(N_{\max})$ coincides asymptotically with the one in the classical problem of minimizing the expected rank studied in Chow et al. (1964); see Problem (P4) in Section 2. On the other hand, if N is uniformly distributed on $\{1, \dots, N_{\max}\}$, i.e. $\alpha = 1$, then $V_*(N_{\max}) \to \infty$ as $N_{\max} \to \infty$.

We illustrate these results in Table 5. The first row of the table, $\alpha = 1$, corresponds to the uniform distribution where $\gamma_k = 1/N_{\max}$, $k = 1, \dots, N_{\max}$, while for general $\alpha > 0$

$$\gamma_k = \frac{1}{(N_{\max} - k + 1)^\alpha} \prod_{j=1}^{k-1} \Big[1 - \frac{1}{(N_{\max} - j + 1)^\alpha} \Big], \quad k = 1, \dots, N_{\max};$$

see Gianini-Pettitt (1979). It is seen from the table that in the case $\alpha = 3$ the optimal value approaches the universal limit of Chow et al. (1964) as $N_{\max}$ goes to infinity. For $\alpha = 2$ the formula (42) yields the optimal value $4.2444\dots$; this complements the result of Gianini-Pettitt (1979) on boundedness of the optimal value.

5.3 Multiple choice problems

The existing literature treats sequential multiple choice problems as problems of multiple stopping. However, if the reward function has an additive structure, and the involved random variables are independent then these problems can be reformulated in terms of the sequential assignment problem of Section 3. Under these circumstances the results of Derman, Lieberman & Ross (1972) are directly applicable and can be used in order to construct optimal selection rules. We illustrate this approach in the next two examples.

5.3.1 Maximizing the probability of selecting the best observation with k choices

This setting was first considered by Gilbert and Mosteller (1966), and it is discussed in Section 2 as Problem (P8). The goal is to maximize the probability for selecting the best observation with k choices, i.e., to maximize

$$\mathrm{P}\big\{ \cup_{j=1}^{k} (A_{\tau_j} = 1)\big\} = \sum_{j=1}^{k} \mathrm{P}(A_{\tau_j} = 1)$$

with respect to the stopping times $\tau^{(k)} = (\tau_1, \ldots, \tau_k)$, $\tau_1 < \cdots < \tau_k$ of the filtration $\mathscr{R}$. This problem is equivalent to the following version of the sequential assignment problem (AP1) [see Section 3].

> Let $0 = p_1 = \cdots = p_{n-k} < p_{n-k+1} = \cdots = p_n = 1$, and let
>
> $$Y_t = \frac{t}{n}\mathbf{1}\{R_t = 1\}, \;\; t = 1, \ldots, n.$$
>
> The goal is to maximize $S(\pi) = \mathrm{E}\sum_{t=1}^{n} p_{\pi_t} Y_t$ with respect to $\pi \in \Pi(\mathscr{Y})$, where $\Pi(\mathscr{Y})$ is the set of all non–anticipating policies of filtration $\mathscr{Y}$, i.e., $\{\pi_t = j\} \in \mathscr{Y}_t$ for all $j = 1, \ldots, n$ and $t = 1, \ldots, n$.

The relationship between sequential assignment and multiple choice problems is evident: if a policy π assigns $p_{\pi_t} = 1$ to the observation Y_t then the corresponding tth observation is selected, i.e., events $\{p_{\pi_t} = 1\}$ and $\cup_{j=1}^{k}\{\tau_j = t\}$ are equivalent.

The optimal policy for the above assignment problem is characterized by Theorem 1. Specifically, for $t = 1, \ldots, n$ let $p_{t_1} \le p_{t_2} \le \cdots \le p_{t_{n-t+1}}$ be the subset of the coefficients $\{p_1, \ldots, p_n\}$ that are left unassigned at time t. Let $s_t = \sum_{i=1}^{n-t+1} p_{t_i}$ denote the number of observations to be selected (unassigned coefficients p's equal to 1). The optimal policy π_* at time t partitions the real line by numbers

$$-\infty = a_{0,n-t+1} \le a_{1,n-t+1} \le \cdots \le a_{n-t,n-t+1} \le a_{n-t+1,n-t+1} = \infty,$$

and prescribes to select the tth observation if $Y_t > a_{n-t+1-s_t,n-t+1}$. In words, the last inequality means that the observation is selected if Y_t is greater than the s_t-th largest number among the numbers $a_{1,n-t+1}, a_{2,n-t+1}, \ldots, a_{n-t,n-t+1}$. These numbers are given by the following formulas: $a_{0,n-t+1} = -\infty$, $a_{n-t+1,n-t+1} = \infty$, and for $j = 1, \ldots, n-t$

$$a_{j,n-t+1} = \int_{a_{j-1,n-t}}^{a_{j,n-t}} z \mathrm{d}F_{t+1}(z) + a_{j-1,n-t}F_{t+1}(a_{j-1,n-t}) + a_{j,n-t}(1 - F_{t+1}(a_{j,n-t})),$$

where F_t is the distribution function of Y_t. The optimal value of the problem is

$$S_*(k) = S(\pi_*; k) = \sum_{j=1}^{k} a_{n-j+1,n+1} \; . \tag{43}$$

k	1	2	3	4	5	6	7	8	25
$S_*(k)$	0.36791	0.59106	0.73217	0.82319	0.88263	0.92175	0.94767	0.96491	0.999997

Table 6: Optimal values $S_*(k)$ in the problem of maximizing the probability of selecting the best option with k choices. The table is computed using (44) and (43) for $n = 10^4$.

In our case $F_t(z) = (1 - \frac{1}{t})\mathbf{1}(z \geq 0) + \frac{1}{t}\mathbf{1}(z \geq \frac{t}{n})$, $t = 1, \ldots, n$ which yields

$$\begin{aligned} a_{j,n-t+1} \;\; = \;\; & \tfrac{1}{n}\mathbf{1}\big(a_{j-1,n-t} < \tfrac{t+1}{n} \leq a_{j,n-t}\big) \\ & + a_{j-1,n-t}\big[\big(1 - \tfrac{1}{t+1}\big)\mathbf{1}(a_{j-1,n-t} \geq 0) + \tfrac{1}{t+1}\mathbf{1}\big(a_{j-1,n-t} \geq \tfrac{t+1}{n}\big)\big] \\ & + a_{j,n-t}\big[\big(1 - \tfrac{1}{t+1}\big)\mathbf{1}(a_{j,n-t} < 0) + \tfrac{1}{t+1}\mathbf{1}\big(a_{j,n-t} < \tfrac{t+1}{n}\big)\big] \end{aligned} \tag{44}$$

for $j = 1, \ldots, n - t$, $a_{0,n-t+1} = -\infty$, $a_{n-t+1,n-t+1} = \infty$, and by convention we set $-\infty \cdot 0 = \infty \cdot 0 = 0$.

Table 6 gives optimal values $S_*(k)$ for $n = 10^4$ and different k. Note that the case $k = 1$ corresponds to the classical secretary problem. It is clearly seen that the optimal probability of selecting the best observideo grows fast with the number of possible choices k. The numbers presented in the table agree with those given in Table 4 of Gilbert and Mosteller (1966).

The structure of the optimal policy allows to compute distribution of the time required for the subset selection. As an illustration, we consider computation of the expected time required for selecting two options ($k = 2$). According to the optimal policy the first choice is made at time $\tau_1 := \min\{t = 1, \ldots, n : Y_t > a_{n-t-1,n-t+1}\}$, while the second choice occurs at time $\tau_2 := \min\{t > \tau_1 : Y_t > a_{n-t,n-t+1}\}$. Then the expected time to the subset selection is

$$\mathrm{E}\tau_2 \;\; = \;\; \mathrm{E}\tau_1 + \mathrm{E}(\tau_2 - \tau_1), \tag{45}$$

where

$$\mathrm{E}\tau_1 = 1 + \sum_{j=1}^{n-1} \prod_{t=1}^{j} F_t(a_{n-t-1,n-1+1}) \tag{46}$$

$$\begin{aligned} \mathrm{E}(\tau_2 - \tau_1) &= 1 + \sum_{i=1}^{n-2} \mathrm{P}(\tau_2 - \tau_1 > i) = 1 + \sum_{j=1}^{n-1} \sum_{i=1}^{n-j-1} \mathrm{P}(\tau_2 - \tau_1 > i \,|\, \tau_1 = j)\mathrm{P}(\tau_1 = j) \\ &= 1 + \sum_{j=1}^{n-1} \sum_{i=1}^{n-j-1} \prod_{t=1}^{j+i} F_t(a_{n-t,n-1+1})\mathrm{P}(\tau_1 = j) \\ &= 1 + \sum_{j=1}^{n-1} \sum_{i=1}^{n-j-1} \prod_{t=j+1}^{j+i} F_t(a_{n-t,n-1+1})\big[1 - F_j(a_{n-j-1,n-j+1})\big] \prod_{t=1}^{j-1} F_t(a_{n-t-1,n-t+1}). \end{aligned} \tag{47}$$

These formulas are clearly computationally amenable and easy to code on a computer.

k	1	2	3	4	5	6	7	8	25
$S_*(k)$	3.86488	4.50590	5.12243	5.72330	6.31262	6.89285	7.46574	8.03255	17.22753

Table 7: The optimal value $S_*(k)$ in the problem of minimization of the expected average rank with k choices for $n = 10^5$.

5.3.2 Minimization of the expected average rank with k choices

In this problem that it is discussed in Section 2 as Problem (P9) we want to minimize the expected average rank of the k selected observations:

$$\min_{\tau^{(k)}} \mathrm{E}\biggl(\frac{1}{k}\sum_{j=1}^{k} A_{\tau_j}\biggr),$$

where $\tau^{(k)} = (\tau_1, \ldots, \tau_k)$, $\tau_1 < \cdots < \tau_k$ are stopping times of filtration $\mathscr{R}$.

This setting is equivalent to the following sequential assignment problem.

Let $0 = p_1 = \cdots = p_{n-k} < p_{n-k+1} = \cdots = p_n = 1$, and let

$$Y_t = -\frac{n+1}{t+1} R_t, \quad t = 1, \ldots, n.$$

The goal is to maximize $S(\pi) = \mathrm{E}\sum_{t=1}^{n} p_{\pi_t} Y_t$ with respect to $\pi \in \Pi(\mathscr{Y})$.

Note that here F_t is a discrete distribution with atoms at $y_t(\ell) = -\frac{n+1}{t+1}\ell$, $\ell = 1, \ldots, t$ and corresponding probabilities $f_t(\ell) := \mathrm{P}\{Y_t = y_t(\ell)\} = \frac{1}{t}$. The structure of the optimal policy is exactly as in the previous section: at time t the real line is partitioned by real numbers $a_{j,n-t+1}$, $j = 0, \ldots, n-t+1$ and tth option if $Y_t > a_{n-t+1-s_t, n-t+1}$, where s_t stands for the number of coefficients p_i equal to 1 at time t. The constants $\{a_{j,n-t+1}\}$ are determined by the following formulas: $a_{0,n-t+1} = -\infty$, $a_{n-t+1,n-t+1} = \infty$, and for $j = 2, \ldots, n-t$

$$\begin{aligned} a_{j,n-t+1} &= \frac{1}{t+1}\sum_{\ell=1}^{t+1} y_{t+1}(\ell)\mathbf{1}\bigl\{y_{t+1}(\ell) \in (a_{j-1,n-t}, a_{j,n-t}]\bigr\} \\ &\quad + \frac{a_{j-1,n-t}}{t+1}\sum_{\ell=1}^{t+1}\mathbf{1}\bigl\{y_{t+1}(\ell) \le a_{j-1,n-t}\bigr\} + \frac{a_{j,n-t}}{t+1}\sum_{\ell=1}^{t+1}\mathbf{1}\bigl\{y_{t+1}(\ell) > a_{j,n-t}\bigr\}. \end{aligned}$$

The optimal value $S_*(k)$ of the problem is again given by (43). Table 7 presents $S_*(k)$ for $n = 10^5$ and different values of k. It worth noting that $k = 1$ corresponds to the standard problem of expected rank minimization [Problem (P4)] with well known asymptotics $S_*(k) \approx 3.8695\ldots$ as n goes to infinity. Using formulas (45), (46) and (47) we also computed expected time required for $k = 2$ selections when $n = 10^3$: $\mathrm{E}\tau_1 \approx 396.25983$ and $\mathrm{E}\tau_2 \approx 610.54822$. Such performance metrics were not established so far and our approach illustrates the simplicity with which this can be done.

5.4 Miscellaneous problems

The next two examples illustrate applicability of the proposed framework to some other problems of optimal stopping.

5.4.1 Moser's problem with random horizon

This is Problem (P10) of Section 2. The stopping problem is

$$V_*(N_{\max}) := \max_{\tau\in\mathscr{T}(\mathscr{X})} \mathrm{E}[(X_\tau - X_N)\mathbf{1}\{\tau \le N\}] + \mu.$$

Define $Y_t = \mathrm{E}\big[(X_t - X_N)\mathbf{1}\{t \le N\} \,|\, \mathscr{X}_t]$; then

$$Y_t = \sum_{k=t}^{N_{\max}} \mathrm{E}\big[(X_t - X_N)\mathbf{1}\{N = k\} \,|\, \mathscr{X}_t\big] = (X_t - \mu)\sum_{k=t+1}^{N_{\max}} \gamma_k,$$

and for any stopping time $\tau \in \mathscr{T}(\mathscr{X})$

$$\mathrm{E}\big[(X_\tau - X_N)\mathbf{1}\{\tau \le N\}\big] = \sum_{t=1}^{\infty} \mathrm{E}\Big[\mathbf{1}\{\tau = t\}\mathrm{E}\big\{(X_t - X_N)\mathbf{1}\{t \le N\} \,|\, \mathscr{X}_t\big\}\Big] = \mathrm{E}Y_\tau.$$

Thus, the original stopping problem is equivalent to the problem of stopping the sequence of independent random variables $Y_t = (X_t - \mu)\sum_{k=t+1}^{N_{\max}} \gamma_k$, $t = 1, \ldots, N_{\max}$, and the optimal value is

$$V_*(N_{\max}) = \mu + \max_{\tau\in\mathscr{T}(\mathscr{Y})} \mathrm{E}Y_\tau.$$

The distribution of Y_t is $F_t(z) = G(\mu + \frac{z}{\sigma_t})$, $t = 1, \ldots, N_{\max}$, where $\sigma_t := \sum_{k=t+1}^{N_{\max}} \gamma_k$. Then applying Corollary 1 we obtain that the optimal stopping rule is given by

$$\begin{aligned}
&b_1 = -\infty, \quad b_2 = \mathrm{E}Y_{N_{\max}},\\
&b_{t+1} = \int_{b_t}^{\infty} z\mathrm{d}F_{N_{\max}-t+1}(z) + b_t F_{N_{\max}-t+1}(b_t), \quad t = 2, \ldots, N_{\max},\\
&\tau_* = \min\{1 \le t \le N_{\max} : Y_t > b_{N_{\max}-t+1}\}.
\end{aligned}$$

In particular, if G is the uniform $[0, 1]$ distribution then straightforward calculation yields: $b_2 = 0$ and

$$b_{t+1} = \frac{1}{2\sigma_{N_{\max}-t+1}}\Big(b_t + \tfrac{1}{2}\sigma_{N_{\max}-t+1}\Big)^2, \quad t = 2, \ldots, N_{\max}.$$

The optimal value of the problem is $V_*(N_{\max}) = b_{N_{\max}+1} + \frac{1}{2}$.

It is worth noting that the case of $\gamma_k = 0$ for all $k = 1, \ldots, N_{\max} - 1$ and $\gamma_{N_{\max}} = 1$ corresponds to the original Moser's problem with fixed horizon $N_{\max}$. In this case $\sigma_t = 1$ for all t, and the above recursive relationship coincides with the one in Moser (1956) which is $E_{t+1} = \frac{1}{2}(1 + E_t^2)$ where $E_t = b_t + \frac{1}{2}$.

5.4.2 Bruss' Odds–Theorem

This is the stopping problem (P11) of Section 2. In this setting we have

$$Y_t := \mathrm{P}\{Z_t = 1, Z_{t+1} = \cdots = Z_n = 0 \,|\, \mathscr{Z}_t\} = \begin{cases} Z_t \prod_{k=t+1}^{n} q_k, & t = 1, \ldots, n-1,\\ Z_t, & t = n, \end{cases} \tag{48}$$

and then

$$V_* := \max_{\tau\in\mathscr{T}(\mathscr{Z})} \mathrm{P}(Z_\tau = 1, Z_{\tau+1} = \cdots = Z_n = 0\} = \max_{\tau\in\mathscr{T}(\mathscr{Y})} \mathrm{E}Y_\tau.$$

Thus, the original stopping problem is equivalent to stopping the sequence $\{Y_t\}$ which is given in (48). Note that Y_t's are independent, and Y_t takes two values $\prod_{k=t+1}^n q_k$ and 0 for $t = 1,\dots,n-1$, and 1 and 0 for $t = n$ with respective probabilities p_t and $q_t = 1 - p_t$. Therefore applying Corollary 1 we obtain that the optimal stopping rule is given by

$$\tau_* = \min\Big\{t = 1,\dots,n : Y_t > b_{n-t+1}\Big\}, \tag{49}$$

where $b_1 = -\infty$, $b_2 = \mathrm{E}Y_n = p_n$, and for $t = 2,3,\dots,n$

$$b_{t+1} = \int_{b_t}^{\infty} z\mathrm{d}F_{n-t+1}(z) + b_t F_{n-t+1}(b_t) = b_t + p_{n-t+1}\Big[\prod_{k=n-t+2}^{n} q_k - b_t\Big]_+, \tag{50}$$

where $[\cdot]_+ = \max\{0,\cdot\}$. The problem optimal value is $V_* = b_{n+1}$.

Now we demonstrate the the stopping rule (49)–(50) is equivalent to the sum–odds–and–stop algorithm of Bruss (2000). According to (49), it is optimal to stop at the first time instance $t \in \{1,\dots,n-1\}$ such that $Z_t = 1$ and $b_{n-t+1}(\prod_{k=t+1}^n q_k)^{-1} < 1$; if such t does not exist then the stopping time is n. Note that

$$\frac{b_{n-t+1}}{\prod_{k=t+1}^n q_k} = \frac{b_{n-t}}{\prod_{k=t+1}^n q_k} + \frac{p_{t+1}}{q_{t+1}}\Big[1 - \frac{b_{n-t}}{\prod_{k=t+2}^n q_k}\Big]_+, \quad t = 0,1,\dots,n-2. \tag{51}$$

Define $u_s := b_s(\prod_{k=n-s+2}^n q_k)^{-1}$, $s = 2,\dots,n+1$. It is evident that $\{u_s\}$ is a monotone increasing sequence, and with this notation (51) takes the form

$$u_{n-t+1} = \frac{1}{q_{t+1}}u_{n-t} + \frac{p_{t+1}}{q_{t+1}}(1 - u_{n-t})_+, \quad t = 0,1,\dots,n-2, \tag{52}$$

$$u_2 = \frac{p_n}{q_n}. \tag{53}$$

In terms of the sequence $\{u_s\}$ the optimal stopping rule (49) is the following: it is optimal to stop at first time $t \in \{1,\dots,n-1\}$ such that $Z_t = 1$ and $u_{n-t+1} < 1$; if such t does not exist then stop at time n. Formally, define $t_* := \min\{t = 1,\dots,n-1 : u_{n-t+1} < 1\}$ if it exists. Then for any $t \in \{t_*, t_*+1.\dots,n-1\}$ we have $u_{n-t+1} < 1$ and iterating (52)-(53) we obtain

$$u_{n-t+1} = u_{n-t} + \frac{p_{t+1}}{q_{t+1}} = \sum_{k=t+1}^{n} \frac{p_k}{q_k}, \quad t = t_*, t_*+1,\dots,n-1. \tag{54}$$

Therefore (49) can be rewritten as

$$\tau_* = \inf\Big\{t = 1,\dots,n-1 : Z_t = 1 \ \text{ and } \sum_{k=t+1}^{n} \frac{p_k}{q_k} < 1\Big\} \wedge n,$$

where by convention $\inf\{\emptyset\} = \infty$. In order to compute the optimal value $V_* = b_{n+1}$ we need to determine u_{n+1}. For this purpose we note that the definition of t_* and (52) imply

$$u_{n-t+1} = \frac{u_{n-t}}{q_{t+1}}, \quad t = t_* - 1, t_* - 2,\dots,1,0, \tag{55}$$

and, in view of (54), $u_{n-t_*+1} = \sum_{k=t_*+1}^{n}(p_k/q_k)$. Therefore iterating (55) we have

$$u_{n+1} = \Big(\prod_{j=1}^{t_*}\frac{1}{q_j}\Big)u_{n-t_*+1} = \Big(\prod_{j=1}^{t_*}\frac{1}{q_j}\Big)\sum_{k=t_*+1}^{n}\frac{p_k}{q_k}.$$

Taking into account that $u_{n+1} = b_{n+1}(\prod_{j=1}^{n} q_j)^{-1}$ we finally obtain the optimal value of the problem:

$$V_* = b_{n+1} = \prod_{j=t_*+1}^{n} q_j \sum_{k=t_*+1}^{n}\frac{p_k}{q_k}.$$

These results coincide with the statement of Theorem 1 in Bruss (2000).

6 Concluding remarks

The proposed framework is applicable to sequential selection problems that can be reduced to settings with independent observations and additive reward function. As we showed, this class is rather broad; in particular, it includes selection problems with no-information, rank-dependent rewards and fixed or random horizon. In addition, the framework covers selection problems with full information when the random variables $\{X_t\}$ are observable, and the reward for stopping at time t is a function of the current observation X_t only. Also, multiple choice problems with random horizon and additive reward can be solved by the proposed framework.

There are selection problems that do not belong to the indicated class. For instance, settings with rank–dependent reward and full information as in Gilbert and Mosteller (1966, Section 3) and Gnedin (2007) cannot be reduced to optimal stopping of a sequence of independent random variables. Another example is the multiple choice problem with zero–one reward; see, e.g., Rose (1982a) and Vanderbei (1980) where the problem of maximizing the probability of selecting k best alternatives was considered. The fact that the results of Derman, Lieberman & Ross (1972) are not applicable to the latter problem was already observed by Rose (1982a) who mentioned this explicitly.

References

Ajtai, M., Megiddo, N. and Waarts, O. (2001). Improved algorithms and analysis for secretary problems and generalizations. *SIAM J. Discrete Math.* **14**, 127.

Albright, S. C., Jr. (1972). Stochastic sequential assignment problems. *Technical report No. 142*, Department of Statistics, Stanford University.

Arlotto, A. and Gurvich, I. (2018) Uniformly bounded regret in the multi-secretary problem. *Stochastic Systems*, forthcoming.

Berezovsky, B.A. and Gnedin, A. B. (1984). *The Problem of Optimal Choice.* Nauka, Moscow (in Russian).

Bruss, F. T. (2000). Sum the odds and stop. *Ann. Probab.* **26**, 1384–1391.

Bruss, F. T. and Louchard, G. (2016). Sequential selection of the best out of n rankable objects. *Discrete Math. Theor. Comput. Sci.* **18**, no. 3, Paper No. 13, 12 pp.

Chow, Y. S., Moriguti, S., Robbins, H. and Samuels, S. M. (1964). Optimal selection based on relative rank (the "Secretary problem"). *Israel J. Math.* **2**, 81–90.

Chow, Y. S., Robbins, H. and Siegmung, D. (1971). *Great Expectations: The Theory of Optimal Stopping.* Houghton Mifflin Company, Boston

Derman, C., Lieberman, G. J. and Ross, S. (1972). A sequential stochastic assignment problem. *Management Science* **18**, 349–355.

Dietz, C., van der Laan, D. and Ridder, A. (2011). Approximate results for a generalized secretary problem. *Probab. Engrg. Inform. Sci.* **25**, 157-169.

Dynkin, E. B. (1963). The optimum choice of the instant for stopping a Markov process. (Russian) *Dokl. Akad. Nauk SSSR* **150**, 238-240. Also in: *Selected papers of E. B. Dynkin with commentary*, 485–488. Edited by A. A. Yushkevich, G. M. Seitz and A. L. Onishchik. American Mathematical Society, Providence, RI; International Press, Cambridge.

Ferguson, T. S. (1989). Who solved the secretary problem? *Statist. Science* **4**, 282–296.

Ferguson, T.S. (2008). *Optimal Stopping and Applications.* https://www.math.ucla.edu/t̃om/Stop–ping/Contents.html

Frank, A. Q. and Samuels, S. M. (1980). On an optimal stopping problem of Gusein–Zade. *Stoch. Proc. Appl.* **10**, 299-311.

Freeman, P. R. (1983). The secretary problem and its extensions: a review. *Int. Statist. Review* **51**, 189-206.

Gianini-Pettitt, J. (1979). Optimal selection based on relative ranks with a random number of individuals. *Adv. Appl. Probab.* **11**, 720–736.

Gilbert, J. and Mosteller, F. (1966). Recognizing the maximum of a sequence. *J. Amer. Statist. Assoc.* **61**, 35-73.

Goldenshluger, A. and Zeevi, A. (2018). Optimal stopping of a random sequence with unknown distribution. *Preprint.*

Gnedin, A. V. (2007). Optimal stopping with rank-dependent loss. *J. Appl. Probab.* **44**, 996–1011.

Gnedin, A. V. and Krengel, U. (1996). Optimal selection problems based on exchangeable trials. *Ann. Appl. Probab.* **6**, 862–882.

Gusein–Zade, S. M. (1966). The problem of choice and the optimal stopping rule for a sequence of independent trials. *Theory Probab. Appl.* **11**, 472–476.

Haggstrom, G. W. (1967). Optimal sequential procedures when more than one stop is required. *Ann. Math. Statist.* **38**, 1618–1626.

Krieger, A. M. and Samuel-Cahn, E. (2009). The secretary problem of minimizing the expected rank: a simple suboptimal approach with generalizations. *Adv. in Appl. Probab.* **41**, 1041–1058

Krieger, A. M., Pollak, M. and Samuel-Cahn, E. (2007). Select sets: rank and file. *Ann. Appl. Probab.* **17**, 360–385.

Krieger, A. M., Pollak, M. and Samuel-Cahn, E. (2008). Beat the mean: sequential selection by better than average rules. *J. Appl. Probab.* **45**, 244–259.

Lin, Y.-S., Hsiau, S.-R. and Yao, Y.-C. (2016). Optimal selection of the k-th best candidate. `arXiv:1701.00052v1`.

Lindley, D. V. (1961). Dynamic programming and decision theory. *Appl. Statist.* **10**, 39–51.

Moser, L. (1956). On a problem of Cayley. *Scripta Math.* **22**, 289–292.

Mucci, A. G. (1973). Differential equations and optimal choice problems. *Ann. Statist.* **1**, 104–113.

Nikolaev, A. G. and Jacobson, S. H. (2010). Stochastic sequential decision–making with a random number of jobs. *Oper. Res.* **58**, 1023–1027.

Nikolaev, M. L. and Sofronov, G. (2007). A multiple optimal stopping rule for sums of independent random variables. *Discrete Math. Appl.* **17**, 463–473.

Presman, E. L. and Sonin, I. M. (1972). The best choice problem for a random number of objects. *Teor. Veroyatnost. i Primenen.* **17**, 695-706.

Quine, M. P. and Law, J. S. (1996). Exact results for a secretary problem. *J. Appl. Probab.* **33**, 630–639.

Robbins, H. (1970). Optimal stopping. *Amer. Math. Monthly* **77**, 333-343.

Robbins, H. (1991). Remarks on the secretary problem. *Amer. J. Math. and Management Sci.* **11**, 25–37.

Rose, J.S (1982a). A problem of optimal choice and assignment. *Oper. Res.* **30**, 172-181.

Rose, J.S. (1982b). Selection of nonextremal candidates from a sequence. *J. Optimization Theory Applic.* **38**, 207-219.

Samuels, S. (1991). Secretary problems. In: *Handbook of Sequential Analysis* edited by B. K. Ghosh and P.K.Sen. Marcel Dekker Inc., New York.

Vanderbei, R.J. (1980). The optimal choice of a subset of a population. *Math. Oper. Res.* **5**, 481-486.

Vanderbei, R.J. (2012). The postdoc variant of the secretary problem. *Tech. Report*.

Yeo, G. F. (1997). Duration of a secretary problem. *J. Appl. Probab.* **34**, 556-558.