MAXIMIZING THE PROBABILITY OF STOPPING ON ANY OF THE LAST *m* SUCCESSES IN INDEPENDENT BERNOULLI TRIALS WITH RANDOM HORIZON

MITSUSHI TAMAKI,* *Aichi University*

Abstract

We consider the problem of maximizing the probability of stopping on any of the last *m* successes in independent Bernoulli trials with random horizon of length *N*, where *m* is a predetermined integer. A prior is given for *N*. It is known that, when *N* is degenerate, i.e. P{*N* = *n*} = 1 for a given *n* > *m*, the sum-the-multiplicative-odds theorem gives the solution and shows that the optimal rule is a threshold rule, i.e. it stops on the first success appearing after a given stage. However, when *N* is nondegenerate, the optimal rule is not necessarily a threshold rule. So our main concern in Section 2 is to give a sufficient condition for the optimal rule to be a threshold rule when *N* is a bounded random variable such that P{*N* ≤ *n*} = 1. Application will be made to the usual (discrete arrival time) secretary problem with a random number *N* of applicants in Section 3. When *N* is uniform or curtailed geometric, the optimal rules are shown to be threshold rules and their asymptotic results are obtained. We also examine, as a nonhomogeneous Poisson process model, an intermediate prior that allows *N* to be uniform or degenerate. In Section 4 we consider a continuous arrival time version of the secretary problem with a random number *M* of applicants. It is shown that, whatever the distribution of *M*, we can win with probability greater than or equal to u_m^*, where u_m^* is, as given in (1.4), the asymptotic win probability of the usual secretary problem when *N* degenerates to *n* and $n \to \infty$.

Keywords: Optimal stopping; secretary problem; best choice; sum-the-odds theorem; sum-the-multiplicative-odds theorem; threshold rule; e^{-1}-law; continuous arrival time model

2010 Mathematics Subject Classification: Primary 60G40; 62L15

1. Introduction and summary

Before considering our problems, we briefly review the sum-the-multiplicative-odds theorem and the secretary problem. Let *n* be a given positive integer, and suppose that *n* independent Bernoulli trials are performed one at a time, each of which results in a *success* or a *failure*. That is, if we let X_j equal 1 if the *j*th trial is a success and 0 if it is a failure, then $X_1, X_2, \ldots, X_n$ are independent Bernoulli random variables that are observed sequentially. When we seek an optimal stopping rule of this sequential observation problem with the objective of maximizing the probability of stopping on any of the last *m* successes for a predetermined *m* (we assume that *n* > *m* throughout this paper unless otherwise specified, because, for *n* ≤ *m*, the optimal rule evidently stops on the first success), the following theorem gives a solution if we let

$a_j = \mathrm{P}\{X_j = 1\}$, $b_j = 1 - a_j$, and $r_j = a_j/b_j$, and define, for $j \geq 1$ and $k \geq 0$,

$$R_{k,i,j} = \sum_{k < t_1 < t_2 < \cdots < t_j \leq i} r_{t_1} r_{t_2} \cdots r_{t_j} \tag{1.1}$$

for $k + j \leq i \leq n$ and $R_{k,i,j} = 0$ for $k + j > i$.

Theorem 1.1. (Sum-the-multiplicative-odds theorem (STMOT).) *For the problem of maximizing the probability of stopping on any of the last m successes in n independent Bernoulli trials, the optimal rule stops on the first success $X_k = 1$ with $k \geq s_m(n)$, if any, where*

$$s_m(n) = \min\{k \geq 1 : R_{k,n,m} \leq 1\}. \tag{1.2}$$

Moreover, the maximal probability of win (i.e. achieving the objective) is

$$u_m(n) = \left(\prod_{j=s_m(n)}^{n} b_j \right) \left(\sum_{j=1}^{m} R_{s_m(n)-1,n,j} \right).$$

See Tamaki (2010) for further details of the STMOT. When $m = 1$, the STMOT is referred to as the sum-the-odds theorem, which was obtained in Bruss (2000) and later extended into several directions in Ferguson (2008). See also Hill and Krengel (1992), Hsiau and Yang (2002), Bruss (2003), Bruss and Paindaveine (2000), Bruss and Louchard (2009), and Ano *et al.* (2010) for related works.

An interesting application of the STMOT appears in the *secretary problem* described as follows. A known number n of rankable applicants (1 being the best and n the worst) appear one at a time in random order with all $n!$ permutations equally likely. That is, each of the successive ranks of n applicants constitutes a random permutation. Suppose that all that can be observed are the relative ranks of the applicants as they appear. If Y_j denotes the relative rank of the jth applicant among the first j applicants, the sequentially observed random variables are $Y_1, Y_2, \ldots, Y_n$. It is well known that

(a) $Y_1, Y_2, \ldots, Y_n$ are independent random variables;

(b) $\mathrm{P}\{Y_j = i\} = 1/j$, $1 \leq i \leq j$, $1 \leq j \leq n$.

The jth applicant is called a *candidate* if he/she is relatively best, i.e. $Y_j = 1$. If the objective is to stop on any of the last m successes, that is, any of the last m candidates (stopping is identified with selection of an applicant in the secretary problem), independent Bernoulli random variables $X_1, X_2, \ldots, X_n$ are specified by $X_j = I(Y_j = 1)$, where $I(E)$ is the indicator function of an event E, so the STMOT gives the solutions $s_m(n)$ and $u_m(n)$ corresponding to $a_j = 1/j$. In particular, as n tends to ∞, we have

$$s_m^* = \lim_{n \to \infty} \frac{s_m(n)}{n} = \exp\{-(m!)^{1/m}\} \tag{1.3}$$

and

$$u_m^* = \lim_{n \to \infty} u_m(n) = \exp\{-(m!)^{1/m}\} \sum_{j=1}^{m} \frac{(m!)^{j/m}}{j!}. \tag{1.4}$$

See Lemma 3.2 and Table 1 of Tamaki (2010) for (1.3), (1.4), and their numerical values. The secretary problem with $m = 1$, referred to as the *best-choice problem* (because the last candidate is best overall), gives the well-known result $s_1^* = u_1^* = e^{-1}$. The reader is referred to Ferguson (1989) and Samuels (1991) for reviews of the secretary problem.

In Section 2, the STMOT is extended to allow for a random horizon of length N. That is, N represents the random number of Bernoulli trials to be performed, and is assumed to be a bounded random variable that is also independent of Bernoulli trials. A prior distribution will be given for N. A stopping rule is said to be a *threshold rule* if it stops only on the first success appearing after a given stage. In particular, the optimal rule, as described in the STMOT, is said to be a threshold rule with value $s_m(n)$. It is known that, for a random N, the optimal rule is not necessarily a threshold rule (see, e.g. Presman and Sonin (1972)). Hence, our main concern in Section 2 is to give a simple sufficient condition for the optimal rule to be a threshold rule.

An application of this condition again appears in the secretary problem (i.e. $a_j = 1/j$) with a random number N of applicants. In particular, for the problem with N uniform on $\{1, 2, \ldots, n\}$, the optimal rule will be shown to be a threshold rule with value $t_m(n)$ having the limiting property

$$t_m^* = \lim_{n \to \infty} \frac{t_m(n)}{n} = \exp\{-[(m+1)!]^{1/m}\}. \tag{1.5}$$

The corresponding probability of win $v_m(n)$ also has the limit

$$v_m^* = \lim_{n \to \infty} v_m(n) = \exp\{-[(m+1)!]^{1/m}\} \sum_{j=1}^{m} \frac{[(m+1)!]^{(j+1)/m}}{(j+1)!}. \tag{1.6}$$

See Lemma 3.1 and Table 1 in Section 3 for (1.5), (1.6), and their numerical values. We see that $t_1^* = e^{-2}$ and $v_1^* = 2e^{-2}$, which coincides with the result derived in Presman and Sonin (1972), who were the first to study the best-choice problem with a random number of applicants. See also Irle (1980) and Petruccelli (1983). It may be interesting to compare (t_m^*, v_m^*) and (s_m^*, u_m^*) as two extremes. A generalized uniform prior that can be a bridge between (s_m^*, u_m^*) and (t_m^*, v_m^*) will be discussed as a nonhomogeneous Poisson process model in Section 3.1. In addition to the uniform prior, a curtailed geometric prior is also examined in detail. See also the best-choice problem with random freeze in Samuel-Cahn (1995) as a related work.

We consider a variation of the secretary problem in Section 4 as a contrast to the problem studied in Sections 2 and 3. Instead of having the applicants appear in discrete time, we have them appear in continuous time. As such, we consider the *continuous arrival time model* posed in Bruss (1984) and give a generalization of the so-called e^{-1}-law for $m = 1$.

2. Bernoulli trials with random horizon

Let $X_1, X_2, \ldots$ be a sequence of independent Bernoulli random variables with $P\{X_i = 1\} = a_i$ for $1 \leq i$, and let N be a bounded random variable having a prior distribution $p_k = P\{N = k\}$ such that $\sum_{k=1}^{n} p_k = 1$ and $p_n > 0$ for a given n. We also assume that N is independent of $X_1, X_2, \ldots$. The objective of the problem is to find a stopping rule that maximizes the probability of stopping on any of the last m successes in N trials. Remember that the problem with $\{a_i = 1/i, \ i \geq 1\}$ is referred to as a secretary problem.

For ease of description, let, for a given prior $\{p_k, \ 1 \leq k \leq n\}$,

$$\pi_k = P\{N \geq k\} = p_k + p_{k+1} + \cdots + p_n, \qquad 1 \leq k \leq n,$$

with $\pi_1 = 1$ and $\pi_n > 0$ (π_0 is interpreted as 1 if it appears). Write $b_i = 1 - a_i$ and $r_i = a_i/b_i$ as before, and let

$$B_{k,i} = b_{k+1}b_{k+2}\ldots b_i, \qquad 0 \le k < i \le n,$$

with $B_{k,k} = 1$ for convenience. Define $R_{k,i,j}$ as in (1.1). Then we have the following preliminary lemma concerning the number of successes.

Lemma 2.1. *For a given $k \ge 0$, assume that $0 < a_l < 1$ for $l \ge k + 1$. Then, for $j \ge 0$,*

  (i) $P\{X_{k+1} + \cdots + X_i = j\} = B_{k,i} R_{k,i,j}, \ k + j \le i \le n;$

  (ii) $P\{X_{k+1} + \cdots + X_N = j \mid N \ge k\} = \sum_{i=k+j}^{n}(B_{k,i} R_{k,i,j})p_i/\pi_k, \ k + j \le n,$

where we use the additional convention that $R_{k,i,0} = 1$ for $i \ge k$ to make (i) and (ii) valid for $j = 0$ as well.

  Proof. (i) See the proof of Theorem 1 of Bruss (2000). Since the X_js are independent of N, (ii) is immediate from (i) by conditioning on the value of $N \ge k$.

Remark 2.1. In the secretary problem, $a_1 = 1$, so Lemma 2.1 is not applicable to the secretary problem for $k = 0$. However, since $X_1 = 1$ and $N \ge 1$ with probability 1, it is easy to relate $k = 0$ to $k = 1$ as follows:

  (i) $P\{X_1 + \cdots + X_i = j\} = P\{X_2 + \cdots + X_i = j - 1\};$

  (ii) $P\{X_1 + \cdots + X_N = j \mid N \ge 0\} = P\{X_2 + \cdots + X_N = j - 1 \mid N \ge 1\}.$

To avoid an unnecessary complication, we assume that $0 < a_i < 1$ for $i \ge 1$ throughout this paper except for the secretary problem.

  Denote by k the *state* in which we have observed the kth trial to be successful, $1 \le k \le n$. Let $S(k)$ and $C(k)$ represent the win probabilities when stopping on the present success in state k and when continuing observations (after leaving state k) in an optimal manner, respectively. Then $V(k) = \max\{S(k), C(k)\}$ represents the optimal value attained by starting from state k. Let $Q_j(k)$ be the probability that the number of successes resulting from the future trials, after leaving state k, is j. This is just the probability given in Lemma 2.1(ii), i.e.

$$Q_j(k) = \sum_{i=k+j}^{n} (B_{k,i} R_{k,i,j})\frac{p_i}{\pi_k}, \qquad k + j \le n, \tag{2.1}$$

with $Q_j(k) = 0$, $k + j > n$, and we can write

$$S(k) = \sum_{j=0}^{\min(m-1,n-k)} Q_j(k) = \begin{cases} \displaystyle\sum_{j=0}^{m-1} Q_j(k) & \text{if } k \le n - m, \\ 1 & \text{if } k > n - m, \end{cases} \tag{2.2}$$

because, by stopping in state k, we win if the number of future successes is less than m.

  On the other hand, if we decide not to stop in state k and proceed to the next stage, we can observe the $(k + 1)$th trial with probability π_{k+1}/π_k, but observe no trial with the remaining probability $1 - \pi_{k+1}/\pi_k$, in which case we come to know that the kth trial was the last trial and

we have lost. The $(k + 1)$th trial, if performed, results in a success or a failure with respective probabilities of a_{k+1} and b_{k+1}. Hence, we have

$$C(k) = \frac{\pi_{k+1}}{\pi_k}[a_{k+1}V(k + 1) + b_{k+1}C(k + 1)], \qquad 1 \le k < n, \tag{2.3}$$

with the obvious boundary condition $C(n) = 0$. Note that the same reasoning can be used to yield the recursive form of $Q_j(k)$, in addition to (2.1), as follows:

$$Q_j(k) = \frac{\pi_{k+1}}{\pi_k}[a_{k+1}Q_{j-1}(k + 1) + b_{k+1}Q_j(k + 1)], \qquad j \ge 1, \tag{2.4}$$

$$Q_0(k) = \frac{p_k}{\pi_k} + \frac{\pi_{k+1}}{\pi_k}b_{k+1}Q_0(k + 1). \tag{2.5}$$

Define $\tilde{C}(k)$ to be the win probability when using a rule of stopping on the first success after leaving state k, if any. Then, from (2.3), $\tilde{C}(k)$ satisfies the relation

$$\tilde{C}(k) = \frac{\pi_{k+1}}{\pi_k}[a_{k+1}S(k + 1) + b_{k+1}\tilde{C}(k + 1)], \qquad 1 \le k < n. \tag{2.6}$$

Under this rule, we win if and only if the number of future successes is greater than 0 but less than $m + 1$, so $\tilde{C}(k)$ must be expressed as

$$\tilde{C}(k) = \sum_{j=1}^{\min(m, n-k)} Q_j(k) = \begin{cases} \sum_{j=1}^{m} Q_j(k) & \text{if } k \le n - m, \\ 1 - Q_0(k) & \text{if } k > n - m. \end{cases} \tag{2.7}$$

Remark 2.2. We can formally check that (2.7) holds by induction as follows. It clearly holds for $k > n - m$. Assume that $\tilde{C}(k + 1) = \sum_{j=1}^{m} Q_j(k + 1)$ for $k + 1 \le n - m$. Then we have, from (2.6), (2.2), and the induction hypothesis,

$$\tilde{C}(k) = \frac{\pi_{k+1}}{\pi_k}\left[a_{k+1}\sum_{j=0}^{m-1} Q_j(k + 1) + b_{k+1}\sum_{j=1}^{m} Q_j(k + 1)\right]$$

$$= \frac{\pi_{k+1}}{\pi_k}\left[a_{k+1}\sum_{j=1}^{m} Q_{j-1}(k + 1) + b_{k+1}\sum_{j=1}^{m} Q_j(k + 1)\right]$$

$$= \sum_{j=1}^{m} Q_j(k)$$

as desired, where the last equality follows from (2.4).

Let

$$t_m(n) = \min\{j : S(k) \ge \tilde{C}(k) \text{ for all } j \le k \le n\}. \tag{2.8}$$

Then the backward induction shows that it is optimal to stop in state $k \ge t_m(n)$. We now have the following result.

Lemma 2.2. *Whatever the distribution of $N(\le n)$, the first time the optimal rule stops on a success occurs no later than the $s_m(n)$th trial, where $s_m(n)$ is as defined in (1.2). Moreover, the optimal rule stops on the first success among trials $s_m(n), s_m(n) + 1, \ldots, n$ if stopping has not occurred previously.*

Proof. It is sufficient to show that $t_m(n) \leq s_m(n)$. Since $t_m(n) \leq n - m + 1$ from (2.2) and (2.7), and $S(k) \geq \tilde{C}(k)$ is equivalent to $Q_0(k) \geq Q_m(k)$ for $k \leq n - m$, we can write

$$t_m(n) = \min\{j : Q_0(k) - Q_m(k) \geq 0 \text{ for all } j \leq k \leq n - m\} \tag{2.9}$$

with $\min\{\phi\} = n - m + 1$, and obtain, from (2.1), the expression

$$Q_0(k) - Q_m(k) = (\pi_k^{-1})\left[\sum_{i=k}^{k+m-1} B_{k,i} p_i + \sum_{i=k+m}^{n} B_{k,i} p_i (1 - R_{k,i,m})\right]. \tag{2.10}$$

By definition, $R_{k,i,m}$ is nonincreasing in k and nondecreasing in i, so we easily see that $t_m(n) \leq s_m(n)$ from (1.2), (2.9), and (2.10).

Theorem 2.1. *A necessary and sufficient condition for the optimal rule to be a threshold rule with value $t_m(n)$ is that, for all $1 \leq k < t_m(n) - 1$ (if such a k exists),*

$$\pi_k \sum_{j=0}^{m-1} Q_j(k) < \pi_{t_m(n)-1} \sum_{j=1}^{m} Q_j(t_m(n) - 1). \tag{2.11}$$

Proof. The optimal rule obviously stops in state k if $k \geq t_m(n)$. If $t_m(n) = 1$, the proof is complete. Suppose now that $t_m(n) > 1$. Then, from the definition of $t_m(n)$ given in (2.8) we have

$$S(t_m(n) - 1) < \tilde{C}(t_m(n) - 1),$$

which is, from (2.2) and (2.7), equivalent to

$$\sum_{j=0}^{m-1} Q_j(t_m(n) - 1) < \sum_{j=1}^{m} Q_j(t_m(n) - 1),$$

implying that (2.11) holds for $k = t_m(n) - 1$ and the optimal rule does not stop in state $t_m(n) - 1$. Suppose that there exists $j < t_m(n) - 1$ such that the optimal rule stops in state j. This prevents the optimal rule from being a threshold rule. Let k be the largest among such js. Considering that, if we stop in state k, we win with probability $S(k)$, whereas if we follow, after leaving state k, a threshold rule with value $t_m(n)$, we will win with probability $(\pi_{t_m(n)-1}/\pi_k)\tilde{C}(t_m(n) - 1)$ (because we win with probability $\tilde{C}(t_m(n) - 1)$ by pretending we are leaving state $t_m(n) - 1$ if the Bernoulli trials continue up to time $t_m(n) - 1$ inclusive, otherwise we lose), we must have

$$S(k) \geq \frac{\pi_{t_m(n)-1}}{\pi_k}\tilde{C}(t_m(n) - 1),$$

or, equivalently,

$$\pi_k \sum_{j=0}^{m-1} Q_j(k) \geq \pi_{t_m(n)-1} \sum_{j=1}^{m} Q_j(t_m(n) - 1),$$

which violates (2.11). Hence, (2.11) ensures that the optimal rule never stops before $t_m(n)$.

Let

$$A(k) = \frac{Q_m(k)}{Q_0(k)}, \qquad 1 \leq k \leq n - m.$$

Then we can give a simpler condition.

Lemma 2.3. *A sufficient condition for (2.11) to hold, i.e. for the optimal rule to be a threshold rule with value $t_m(n)$, is that the sequence $\{A(k)\}$ is nonincreasing in k.*

Proof. Let $\Phi(k)$ denote the probability of win attained by the threshold rule with value k, $1 \le k \le n - m + 1$. Then it can be computed from

$$\Phi(k) = \pi_{k-1} \tilde{C}(k-1). \tag{2.12}$$

We thus have, from (2.6),

$$
\begin{aligned}
\Phi(k) - \Phi(k+1) &= \pi_{k-1} \tilde{C}(k-1) - \pi_k \tilde{C}(k) \\
&= \pi_k [a_k S(k) + b_k \tilde{C}(k)] - \pi_k \tilde{C}(k) \\
&= a_k \pi_k [S(k) - \tilde{C}(k)] \\
&= a_k \pi_k [Q_0(k) - Q_m(k)].
\end{aligned} \tag{2.13}
$$

Considering that

$$A(1) \ge \cdots \ge A(t_m(n) - 1) > 1 \ge A(t_m(n)), \tag{2.14}$$

from the assumption of the lemma and the definition of $t_m(n)$ given in (2.9), we find from (2.13) that $\Phi(k)$ is unimodal with peak $\Phi(t_m(n))$, implying that, for $k \le t_m(n) - 1$,

$$\Phi(k-1) < \Phi(t_m(n)). \tag{2.15}$$

Again applying (2.14), i.e. $Q_0(k) < Q_m(k)$ for $k \le t_m(n) - 1$, to $\Phi(k-1)$ yields

$$\Phi(k-1) = \pi_k \sum_{j=1}^{m} Q_j(k) > \pi_k \sum_{j=0}^{m-1} Q_j(k),$$

which, combined with (2.15), ensures condition (2.11).

Lemma 2.3 is considered a special case of Theorem 2.1 of Tamaki (2010), giving a condition for the *one-stage look-ahead rule* (1-sla rule) to be optimal in a monotone problem. See also Ferguson (2006) for the 1-sla rule. The STMOT is obtained by letting N be degenerate, i.e. $P\{N = n\} = 1$, because $A(k)$ then proves to be $R_{k,n,m}$, which is nonincreasing in k. However, when N is nondegenerate, this condition does not work very well. For example, we have, for the secretary problem,

$$B_{k,i} = \frac{k}{i}, \qquad R_{k,i,j} = \sum_{k < i_1 < i_2 < \cdots < i_j \le i} \prod_{t=1}^{j} \frac{1}{i_t - 1}. \tag{2.16}$$

Hence,

$$A(k) = \frac{\sum_{i=k+m}^{n} R_{k,i,m} p_i / i}{\sum_{i=k}^{n} p_i / i}.$$

The numerator and the denominator of $A(k)$ are both nonincreasing in k, but it is not easy to check the monotonicity of $A(k)$ as a whole even for such a simple prior as uniform (see Example 2.3(a) given later).

To find a more convenient sufficient condition, we make the following transformations:

$$s(k) = B_{1,k} \pi_k S(k),$$
$$c(k) = B_{1,k} \pi_k C(k),$$
$$v(k) = B_{1,k} \pi_k V(k) = \max\{s(k), c(k)\}.$$

Let $\mu(k) = s(k) - c(k)$. Then the following recursive relation holds.

Lemma 2.4. *For* $1 \le k \le n - m$,

$$\mu(k) = \mu(k+1) + r_{k+1} \min\{0, \mu(k+1)\} + B_{1,k} G(k) \tag{2.17}$$

with $\mu(n-m+1) = \sum_{i=n-m+1}^{n} B_{1,i} p_i$, *where*

$$G(k) = p_k - r_{k+1} \sum_{i=k+m}^{n} B_{k,i} R_{k+1,i,m-1} p_i. \tag{2.18}$$

Proof. Let $q_j(k) = B_{1,k} \pi_k Q_j(k)$. Then, for $k \le n - m$, we have

$$s(k) = \sum_{j=0}^{m-1} q_j(k) \tag{2.19}$$

from (2.2) and

$$c(k) = c(k+1) + r_{k+1} v(k+1) \tag{2.20}$$

from (2.3). Equations (2.4) and (2.5) can be written as

$$q_j(k) = q_j(k+1) + r_{k+1} q_{j-1}(k+1), \qquad j \ge 1, \tag{2.21}$$
$$q_0(k) = q_0(k+1) + B_{1,k} p_k, \tag{2.22}$$

respectively. Summing both sides of (2.21) over $1 \le j \le m - 1$ and then adding it to (2.22) yields, via (2.19),

$$s(k) = s(k+1) + B_{1,k} p_k + r_{k+1}\{s(k+1) - q_{m-1}(k+1)\}. \tag{2.23}$$

Hence, subtracting (2.20) from (2.23) yields

$$\mu(k) = \mu(k+1) + r_{k+1}\{s(k+1) - v(k+1)\} + B_{1,k}\left\{ p_k - \frac{r_{k+1} q_{m-1}(k+1)}{B_{1,k}} \right\}$$
$$= \mu(k+1) + r_{k+1} \min\{0, \mu(k+1)\} + B_{1,k} G(k),$$

as desired. For $k > n - m$, $S(k) - C(k) = Q_0(k)$ from (2.2) and (2.7). Hence, $\mu(k) = B_{1,k} \pi_k Q_0(k) = \sum_{i=k}^{n} B_{1,i} p_i$.

Observe that (2.17) implies that if $\mu(k+1) \ge 0$ and $G(k) \ge 0$ then $\mu(k) \ge 0$, and that, since $\mu(j) \ge 0$ for $j \ge s_m(n)$ from Lemma 2.2, we only require that, for $k < s_m(n)$,

$$G(j) \ge 0 \quad \text{for all } k \le j < s_m(n) \tag{2.24}$$

to guarantee that $\mu(j) \geq 0$ for $j \geq k$. The following theorem shows that, under an additional condition on $G(k)$ given below in (2.25), $\mu(k)$ for $k \leq s_m(n)$ is unimodal with a peak attained at $\min(\xi, s_m(n))$, where

$$\xi = \min\{1 \leq k \leq n - m : G(k) \geq 0\}$$

with $\min\{\phi\} = n - m + 1$.

Theorem 2.2. *A sufficient condition for the optimal rule to be a threshold rule is that $G(k)$ changes its sign from negative to positive at most once before $s_m(n)$, that is,*

$$\text{once } G(k) \geq 0 \text{ for some } k, \text{ then } G(j) \geq 0 \text{ for all } k \leq j < s_m(n). \quad (2.25)$$

Proof. Assume first that $\xi < s_m(n)$, which implies that $G(k) \geq 0$ for $\xi \leq k < s_m(n)$ and $G(k) < 0$ for $k \leq \xi - 1$ from (2.25). Then, for $\xi \leq k < s_m(n)$, we have, from (2.17) and (2.24),

$$\mu(k) - \mu(k + 1) = B_{1,k} G(k) \geq 0,$$

whereas, for $k \leq \xi - 1$, we have, from (2.17),

$$\mu(k) - \mu(k + 1) \leq B_{1,k} G(k) < 0.$$

When $\xi \geq s_m(n)$, $G(k) < 0$ for all $k < s_m(n)$ from (2.25), and so $\mu(k) - \mu(k + 1) \leq B_{1,k} G(k) < 0$ from (2.17). Thus, the unimodality of $\mu(k)$ is shown. This property, combined with $\mu(s_m(n)) \geq 0$, ensures that the optimal rule is a threshold rule.

The foregoing argument to derive Lemmas 2.2 and 2.4, and Theorems 2.1 and 2.2 follows the same lines as in Petruccelli (1983), who considered the best-choice problem.

If the optimal rule turns out to be a threshold rule, the threshold value $t_m(n)$ and the corresponding probability of win $v_m(n) = \Phi(t_m(n))$ are respectively given as follows from the preceding argument (see (2.1), (2.7), (2.8), and (2.12)):

$$t_m(n) = \min\left\{k : \sum_{i=k}^{n} B_{k,i} p_i \geq \sum_{i=k+m}^{n} B_{k,i} R_{k,i,m} p_i\right\}, \quad (2.26)$$

$$v_m(n) = \sum_{j=1}^{m}\left(\sum_{i=t_m(n)-1+j}^{n} B_{t_m(n)-1,i} R_{t_m(n)-1,i,j} p_i\right). \quad (2.27)$$

Example 2.1. Let N take on only the value greater than $s_m(n) - 1$, i.e. $p_1 = p_2 = \cdots = p_{s_m(n)-1} = 0$. In such a case, $G(k) < 0$ for all $k < s_m(n)$, so the optimal rule is a threshold rule.

Example 2.2. (*Secretary problem with fixed population size.*) Let N be degenerate, i.e. $P\{N = n\} = 1$. This is the problem considered in Tamaki (2010, Section 3). As a simple consequence of Example 2.1, the optimal rule is a threshold rule.

For the secretary problem, $B_{k,i}$ and $R_{k,i,j}$ are given by (2.16), so we have, from (2.18),

$$G(k) = p_k - \sum_{j=m}^{n-k} \frac{1}{k + j} R_{k+1,k+j,m-1} p_{k+j}. \quad (2.28)$$

For the purposes of most applications to the secretary problem, the following corollary is useful.

Corollary 2.1. *For the secretary problem with $p_k > 0$ for all $1 \leq k \leq n$, a sufficient condition for the optimal rule to be a threshold rule is that*

$$\frac{p_{k+j}}{p_k} \text{ is nonincreasing in } k (< s_m(n)) \tag{2.29}$$

for each possible value of j.

Proof. From the assumption that $p_k > 0$ and (2.28),

$$G(k) = p_k(1 - H(k)),$$

where

$$H(k) = \sum_{j=m}^{n-k} \frac{1}{k+j} R_{k+1,k+j,m-1} \frac{p_{k+j}}{p_k}. \tag{2.30}$$

To verify condition (2.25), it suffices to show that $H(k)$ is nonincreasing in k. This immediately follows from two observations:

(i) In the secretary problem, $r_i = 1/(i-1)$ is decreasing in i, so $R_{k+1,k+j,m-1}$ is decreasing in k by definition for each j. Moreover, from condition (2.29), each term on the right-hand side of (2.30) is nonincreasing in k.

(ii) Each term on the right-hand side of (2.30) is nonnegative, so removing it does not increase the sum.

Example 2.3. It is easy to see that the following priors satisfy condition (2.29) with no restriction on k.

(a) Uniform: $p_k = 1/n$, $1 \leq k \leq n$.

(b) Curtailed geometric: for a given parameter $0 < q < 1$,

$$p_k = \frac{(1-q)q^{k-1}}{1-q^n}, \qquad 1 \leq k \leq n.$$

(c) Curtailed Poisson: for a given parameter $0 < \lambda$,

$$p_k = e^{-\lambda} \frac{\lambda^k}{k!} \Big/ \sum_{j=1}^{n} e^{-\lambda} \frac{\lambda^j}{j!}, \qquad 1 \leq k \leq n.$$

(d) Curtailed binomial: for a given parameter $0 < p < 1$,

$$p_k = \frac{\binom{n}{k} p^k (1-p)^{n-k}}{1-(1-p)^n}, \qquad 1 \leq k \leq n.$$

The following example serves as a unification of the two secretary problems.

Example 2.4. (*Generalized uniform.*) Let N be a uniform random variable on $\{T, T+1, \ldots, n\}$ for a given parameter $T = 1, 2, \ldots, n$, i.e.

$$p_k = \begin{cases} 0 & \text{if } 1 \leq k \leq T - 1, \\ \dfrac{1}{n-T+1} & \text{if } T \leq k \leq n. \end{cases}$$

The cases $T = 1$ and $T = n$ correspond to Example 2.3(a) and Example 2.2, respectively. The optimality of the threshold rule for $1 < T < n$ can be ascertained from (2.25) by noting that $G(k) < 0$ for $k \le T - 1$, whereas $G(k)$ is nondecreasing in k for $k \ge T$.

The asymptotic values for some secretary problems will be given in Section 3. Before concluding this section, we briefly consider two problems other than the secretary problem. A problem of interest may include the case of constant success probability $a_k \equiv a$ (hence, $b_k \equiv b$ and $r_k \equiv r = a/b$). We then have

$$B_{k,i} = b^{i-k}, \qquad R_{k,i,j} = \binom{i-k}{j} r^j.$$

Therefore, for $p_k > 0$ for all $1 \le k \le n$, $G(k)$ of (2.18) can be written as

$$G(k) = p_k \left[1 - r^m \sum_{j=m}^{n-k} b^j \binom{j-1}{m-1} \frac{p_{k+j}}{p_k} \right],$$

which implies that condition (2.29) also gives a sufficient condition for this case to have a threshold rule.

Another problem of interest may be to examine the condition on the sequence $\{a_k, \ 1 \le k \le n\}$ for the optimal rule to be a threshold rule for a given prior. For example, when N is uniform, the optimal rule is a threshold rule if a_k satisfies the condition

$$a_{k+1} \le \frac{a_k}{1 + a_k}, \quad \text{or, equivalently,} \quad a_k \ge \frac{a_{k+1}}{1 - a_{k+1}}.$$

This can be shown as follows. Let $A_{k,i} = B_{k,i} r_{k+1} R_{k+1,i,m-1}$. Then,

$$G(k) = \frac{1}{n} \left(1 - \sum_{i=k+m}^{n} A_{k,i} \right).$$

From Theorem 2.2, it suffices to show that $A_{k,i}$ is nonincreasing in k for each i, i.e. $A_{k-1,i} \ge A_{k,i}$. However, we easily see that

$$\begin{aligned}
A_{k-1,i} - A_{k,i} &= B_{k,i}(r_k b_k R_{k,i,m-1} - r_{k+1} R_{k+1,i,m-1}) \\
&\ge B_{k,i} R_{k+1,i,m-1}(r_k b_k - r_{k+1}) \\
&= B_{k,i} R_{k+1,i,m-1} \left(a_k - \frac{a_{k+1}}{1 - a_{k+1}} \right) \\
&\ge 0,
\end{aligned}$$

where the first inequality follows since $R_{k,i,m-1}$ is nonincreasing in k and the second inequality follows from the above condition on $\{a_k\}$.

3. Asymptotic results for the secretary problem

When the optimal rule of the secretary problem is a threshold rule, we have, from (2.16), (2.26), and (2.27),

$$t_m(n) = \min \left\{ k : \sum_{i=k}^{k+m-1} \frac{k}{i} p_i + \sum_{i=k+m}^{n} \frac{k}{i} p_i (1 - R_{k,i,m}) \ge 0 \right\}, \tag{3.1}$$

$$v_m(n) = (t_m(n) - 1) \sum_{j=1}^{m} \sum_{i=t_m(n)-1+j}^{n} \frac{p_i}{i} R_{t_m(n)-1,i,j}, \tag{3.2}$$

where $R_{k,i,j}$ is given in (2.16). Equation (3.2) is valid for $t_m(n) \geq 2$. For $t_m(n) = 1$, we obviously have, from (2.2),

$$v_m(n) = S(1) = \sum_{j=0}^{m-1} \sum_{i=j+1}^{n} \frac{p_i}{i} R_{1,i,j}.$$

For later use, define, as a function of t, $v_m(n, t)$ to be $v_m(n)$ by replacing $t_m(n)$ by t. Hence, $v_m(n, t_m(n)) = v_m(n)$, and $v_m(n, t)$ is interpreted as the probability of win attained by the threshold rule with value t.

When N is degenerate (i.e. Example 2.2), we use the notation $s_m^*(n)$ and $u_m^*(n)$ for $t_m(n)$ and $v_m(n)$, respectively, conforming to the notation of the STMOT. Their asymptotics when $n \to \infty$, i.e. $s_m^* = \lim_{n \to \infty} s_m^*(n)/n$ and $u_m^* = \lim_{n \to \infty} u_m^*(n)$, are already given in (1.3) and (1.4). When N is nondegenerate, we are mainly concerned with examining the asymptotic behaviors for a uniform prior, curtailed geometric prior, and generalized uniform prior.

Lemma 3.1. (Uniform prior.) *Let n tend to ∞ for the uniform prior given in Example 2.3(a). Then, (1.5) and (1.6) hold asymptotically.*

Proof. For a uniform prior, we have, from (3.2),

$$v_m(n, t) = \frac{t-1}{n} \sum_{j=1}^{m} \sum_{i=t-1+j}^{n} \frac{1}{i} \sum_{t-1<k_1<k_2<\cdots<k_j\leq i} \frac{1}{k_1-1} \frac{1}{k_2-1} \cdots \frac{1}{k_j-1}.$$

If we let n tend to ∞ and write z as the limit of t/n, then, using y for i/n and x_s for k_s/n, $1 \leq s \leq m$, and dy and dx_s for $1/n$, $v_m(n, t)$ becomes a Riemann approximation to a multiple integral, i.e. $v_m(n, t) \to \Psi_m(z)$, where

$$\Psi_m(z) = z \sum_{j=1}^{m} \int_{z<y<1} \frac{dy}{y} \int_{z<x_1<x_2<\cdots<x_j<y} \frac{dx_1}{x_1} \frac{dx_2}{x_2} \cdots \frac{dx_j}{x_j}$$

$$= z \sum_{j=1}^{m} \int_{z}^{1} \frac{dy}{y} \left(\int_{z}^{y} \frac{dx_1}{x_1} \int_{x_1}^{y} \frac{dx_2}{x_2} \cdots \int_{x_{j-1}}^{y} \frac{dx_j}{x_j} \right)$$

$$= z \sum_{j=1}^{m} \int_{z}^{1} \frac{dy}{y} \frac{[\log(y/z)]^j}{j!}$$

$$= z \sum_{j=1}^{m} \frac{(-\log z)^{j+1}}{(j+1)!}. \tag{3.3}$$

The value z that maximizes $\Psi_m(z)$ is easily found by setting the derivative with respect to z equal to 0 and then solving for z. When this is done, we obtain (1.5) because

$$\Psi_m'(z) = \frac{(-\log z)^{m+1}}{(m+1)!} + \log z.$$

TABLE 1: Values of t_m^* and v_m^* for several m.

			m			
	1	2	3	4	5	10
t_m^*	0.1353	0.0863	0.0559	0.0365	0.0240	0.0032
v_m^*	0.2707	0.4705	0.6172	0.7243	0.8020	0.9635

and the optimal $z \in (0, 1)$ can be considered as $t_m^* = \lim_{n\to\infty} t_m(n)/n$ from the above approximation. Substituting the optimal value back into (3.3) yields the desired probability $\Psi_m(t_m^*) = v_m^* (= \lim_{n\to\infty} v_m(n))$ as given in (1.6). This completes the proof.

Remark 3.1. The left-hand side of the inequality in (3.1),

$$\sum_{i=k}^{k+m-1} \frac{k}{i}\frac{1}{n} + \sum_{i=k+m}^{n} \frac{k}{i}\frac{1}{n}(1 - R_{k,i,m}),$$

can be approximated, as $n \to \infty$, by

$$\int_{z<y<1} \frac{z\,\mathrm{d}y}{y}\left(1 - \int_{z<x_1<x_2<\cdots<x_m<y} \frac{\mathrm{d}x_1}{x_1}\frac{\mathrm{d}x_2}{x_2}\cdots\frac{\mathrm{d}x_m}{x_m}\right),$$

because the first term vanishes. This integral is simplified to

$$(-z\log z)\left[1 - \frac{(-\log z)^m}{(m+1)!}\right],$$

so equating this to 0 and then solving for $z \in (0, 1)$ yields

$$\exp\{-[(m+1)!]^{1/m}\}.$$

This coincides with (1.5).

In Table 1 we present some numerical values of t_m^* and v_m^*.

To obtain interesting asymptotic results for the curtailed geometric prior given in Example 2.3(b), we must allow the parameter q to depend on n.

Lemma 3.2. (Curtailed geometric prior.) *Let q depend on n through $q = 1 - c/n$ for a fixed positive value $c(< n)$, and define $t_{m,c} = \lim_{n\to\infty} t_m(n)/n$. Then $t_{m,c}$ is a unique root z of the equation*

$$\int_1^{1/z} \frac{\mathrm{e}^{-czx}}{x}\left[1 - \frac{(\log x)^m}{m!}\right]\mathrm{d}x = 0. \tag{3.4}$$

Moreover, as $n \to \infty$, the optimal probability tends to

$$v_{m,c} = \frac{ct}{1 - \mathrm{e}^{-c}} \int_1^{1/t} \frac{\mathrm{e}^{-ctx}}{x}\left[\sum_{j=1}^{m} \frac{(\log x)^j}{j!}\right]\mathrm{d}x, \tag{3.5}$$

where $t = t_{m,c}$.

TABLE 2: Values of $t_{m,c}$ (upper) and $v_{m,c}$ (lower) for several pairs of (m, c).

c	m				
	1	2	3	4	5
0	0.1353	0.0863	0.0559	0.0365	0.0240
	0.2707	0.4705	0.6172	0.7243	0.8020
0.1	0.1317	0.0840	0.0543	0.0355	0.0234
	0.2689	0.4681	0.6149	0.7222	0.8004
1	0.1008	0.0643	0.0416	0.0272	0.0179
	0.2546	0.4494	0.5964	0.7059	0.7867
5	0.0346	0.0225	0.0148	0.0098	0.0065
	0.2337	0.4209	0.5671	0.6792	0.7641
10	0.0174	0.0113	0.0075	0.0049	0.0033
	0.2329	0.4196	0.5656	0.6778	0.7628
50	0.0035	0.0023	0.0015	0.0010	0.0007
	0.2329	0.4196	0.5656	0.6778	0.7628

We omit the proof of Lemma 3.2 because it is similar to that of Lemma 3.1. See Remark 3.3 below for a more transparent proof via a nonhomogeneous Poisson process model.

Remark 3.2. It is easy to see that, from (3.4) and (3.5),

$$\lim_{c \to 0} t_{m,c} = t_m^*, \qquad \lim_{c \to 0} v_{m,c} = v_m^*.$$

That is, as $c \to 0$, a curtailed geometric prior behaves like a uniform prior.

In Table 2 we present some numerical values of $t_{m,c}$ and $v_{m,c}$ for several m and c. It seems that, as $c \to \infty$, $v_{m,c}$ converges to some positive value. For $m = 1$, this value is approximately 0.2329, which coincides with the value given in Presman and Sonin (1972) (see Example 3.4 of Samuel-Cahn (1995) for this value).

For the generalized uniform prior given in Example 2.4, we let T depend on n in such a way that $T/n \to \alpha$ as $n \to \infty$ for a fixed $0 < \alpha < 1$. For ease of treatment, we consider this problem as a nonhomogeneous Poisson process model in Section 3.1.

3.1. Nonhomogeneous Poisson process model

Presman and Sonin (1972, p. 772) found the nonhomogeneous Poisson process (NPP) to be an appropriate setting in which to define the infinite version of the secretary problem as the limit of the corresponding finite problems. See also Section 3 of Tamaki (2010). Suppose that n applicants appear at fractional times $1/n, 2/n, \ldots, n/n$ instead of at times $1, 2, \ldots, n$. As $k/n \to x$ with $n \to \infty$, the occurrence of candidates in the limiting problem constitutes an NPP with intensity function $\lambda(x) = 1/x, 0 < x \leq 1$, due to properties (a) and (b) of the secretary problem described in Section 1. That is, if we denote by $M(s, t)$ the number of candidates appearing in the time interval $(s, t), 0 < s < t \leq 1$, then $M(s, t)$ is distributed as a Poisson random variable with parameter $\int_s^t \lambda(x)\mathrm{d}x = \log(t/s)$, namely,

$$\mathrm{P}\{M(s, t) = k\} = \frac{s}{t} \frac{\{\log(t/s)\}^k}{k!}, \qquad k \geq 0. \tag{3.6}$$

The secretary problem with a random number N of applicants is represented as an NPP with an appropriate random truncation time $V(\le 1)$ adjusted to N. The objective is to stop on any of the last m candidates before V. Let $\Psi(z)$ be the probability of win when we use a (limiting) threshold rule with value z, which stops on the first candidate that appears after time z, if any. Then, conditioning on the value of V, we have, from (3.6),

$$\Psi(z) = \sum_{j=1}^{m} P\{M(z, V) = j\} = \sum_{j=1}^{m} \int_{z}^{1} \frac{z}{v} \frac{\{\log(v/z)\}^j}{j!} f_V(v)\, dv, \tag{3.7}$$

where $f_V(v)$ is the density of V. When N is a generalized uniform random variable with parameter $T = \alpha n$, it is obvious that the corresponding density is given by

$$f_V(v) = \begin{cases} 0 & \text{if } 0 < v \le \alpha, \\ \dfrac{1}{1 - \alpha} & \text{if } \alpha < v < 1. \end{cases}$$

Write $\Psi_{m,\alpha}(z)$ for $\Psi(z)$ corresponding to this density. Then, the straightforward calculation from (3.7) yields

$$\Psi_{m,\alpha}(z) = \frac{z}{1 - \alpha} \sum_{j=1}^{m} \int_{\max(\alpha, z)}^{1} \frac{\{\log(v/z)\}^j}{j!} \frac{dv}{v}$$

$$= \begin{cases} \dfrac{z}{1 - \alpha} \displaystyle\sum_{j=1}^{m} \dfrac{\{\log(1/z)\}^{j+1} - \{\log(\alpha/z)\}^{j+1}}{(j+1)!} & \text{if } 0 < z \le \alpha, \\[2ex] \dfrac{z}{1 - \alpha} \displaystyle\sum_{j=1}^{m} \dfrac{\{\log(1/z)\}^{j+1}}{(j+1)!} & \text{if } \alpha < z \le 1. \end{cases} \tag{3.8}$$

Lemma 3.3. (Generalized uniform prior.) *Let $t_{m,\alpha}$ denote the optimal threshold value, and let $v_{m,\alpha}$ be the optimal probability ($t_{m,0}$ and $v_{m,0}$ are already given as t_m^* in (1.5) and v_m^* in (1.6), respectively). Two cases are distinguished according to whether $\alpha \le t_m^*$ or $\alpha > t_m^*$.*

Case (i): $0 \le \alpha \le t_m^*$. In this case

$$t_{m,\alpha} = t_m^*, \tag{3.9}$$

$$v_{m,\alpha} = \frac{v_m^*}{1 - \alpha}. \tag{3.10}$$

Case (ii): $t_m^* < \alpha < 1$. The optimal threshold value $t_{m,\alpha}$ is a unique root $z \in (0, \alpha)$ of the equation

$$\left\{\log\left(\frac{1}{z}\right)\right\}^{m+1} - \left\{\log\left(\frac{\alpha}{z}\right)\right\}^{m+1} = (m+1)!\, \log\left(\frac{1}{\alpha}\right), \tag{3.11}$$

or, equivalently,

$$\sum_{j=0}^{m} \frac{(\log \alpha)^j (-\log z)^{m-j}}{(j+1)!\,(m-j)!} = 1. \tag{3.12}$$

Moreover,

$$v_{m,\alpha} = \frac{t_{m,\alpha}}{1-\alpha} \sum_{j=1}^{m} \frac{\{\log(1/t_{m,\alpha})\}^j - \{\log(\alpha/t_{m,\alpha})\}^j}{j!}. \tag{3.13}$$

In particular,

$$\lim_{\alpha\to 1} t_{m,\alpha} = s_m^*,$$
$$\lim_{\alpha\to 1} v_{m,\alpha} = u_m^*, \tag{3.14}$$

where s_m^* and u_m^* are as given in (1.3) and (1.4), respectively.

Proof. Differentiating $\Psi_{m,\alpha}(z)$ in (3.8) with respect to z yields

$$\Psi'_{m,\alpha}(z) = \begin{cases} \dfrac{a_\alpha(z)}{(1-\alpha)(m+1)!} & \text{if } 0 < z \le \alpha, \\[2mm] \dfrac{b(z)\log(1/z)}{(1-\alpha)(m+1)!} & \text{if } \alpha < z \le 1, \end{cases} \tag{3.15}$$

where

$$a_\alpha(z) = \left\{\log\left(\frac{1}{z}\right)\right\}^{m+1} - \left\{\log\left(\frac{\alpha}{z}\right)\right\}^{m+1} - (m+1)!\log\left(\frac{1}{\alpha}\right),$$
$$b(z) = \left\{\log\left(\frac{1}{z}\right)\right\}^m - (m+1)!.$$

To prove (3.9) and (3.11), it suffices to show that

(a) $\Psi'_{m,\alpha}(z) = 0$ has a unique root $z \in (0, 1)$ for each α;

(b) $\Psi'_{m,\alpha}(z) = 0$ is reduced to $b(z) = 0$ for $\alpha \le t_m^*$, whereas $\Psi'_{m,\alpha}(z) = 0$ is reduced to $a_\alpha(z) = 0$ for $\alpha > t_m^*$.

Differentiating $\Psi'_{m,\alpha}(z)$ in (3.15) again yields

$$\Psi''_{m,\alpha}(z) = \begin{cases} \dfrac{\{\log(\alpha/z)\}^m - \{\log(1/z)\}^m}{(1-\alpha)m!\,z} & \text{if } 0 < z \le \alpha, \\[3mm] \dfrac{m! - \{\log(1/z)\}^m}{(1-\alpha)m!\,z} & \text{if } \alpha < z \le 1, \end{cases}$$

showing that $\Psi''_{m,\alpha}(z)$ changes its sign at $\max(\alpha, s_m^*)$ from negative to positive. This property, together with the boundary conditions $\lim_{z\to 0}\Psi'_{m,\alpha}(z) > 0$ and $\Psi'_{m,\alpha}(1) = 0$ from (3.15), ensures part (a). For part (b), it is important to see that

$$\Psi'_{m,\alpha}(\alpha) = \frac{-\log\alpha}{(1-\alpha)(m+1)!}[(-\log\alpha)^m - (m+1)!],$$

which implies that, from the definition of t_m^*, $\Psi'_{m,\alpha}(\alpha) < 0$ if $\alpha > t_m^*$, whilst $\Psi'_{m,\alpha}(\alpha) \ge 0$ if $\alpha \le t_m^*$. This just proves part (b) because, if $\alpha > t_m^*$, $t_{m,\alpha}$ must be less than α, whereas, if $\alpha \le t_m^*$, $t_{m,\alpha}$ must be greater than or equal to α.

The optimal probability $v_{m,\alpha}$ is obtained from (3.8) through $v_{m,\alpha} = \Psi_{m,\alpha}(t_{m,\alpha})$. Note that (3.10) is also due to (1.6) and that, to obtain (3.13), (3.11) with z replaced by $t_{m,\alpha}$ must be taken into consideration. Letting $\alpha \to 1$ in (3.12) yields

$$\frac{(-\log z)^m}{m!} = 1,$$

which proves (3.14). Moreover, from (3.13), (3.14), (1.3), and (1.4),

$$\lim_{\alpha \to 1} v_{m,\alpha} = s_m^* \sum_{j=1}^{m} \frac{1}{j!} \lim_{\alpha \to 1} \frac{(-\log s_m^*)^j - (-\log s_m^* + \log \alpha)^j}{1 - \alpha}$$

$$= s_m^* \sum_{j=1}^{m} \frac{1}{j!} j(-\log s_m^*)^{j-1}$$

$$= s_m^* \sum_{j=1}^{m} \frac{(m!)^{(j-1)/m}}{(j-1)!}$$

$$= u_m^*.$$

This completes the proof.

For $m = 1$ and $m = 2$, we can give explicit expressions for $t_{m,\alpha}$ and $v_{m,\alpha}$.

Example 3.1. Suppose that $m = 1$ and $t_1^* = e^{-2}$.

Case (i): $0 \leq \alpha \leq t_1^*$. In this case

$$t_{1,\alpha} = e^{-2}, \qquad v_{1,\alpha} = \frac{2}{1 - \alpha} e^{-2}.$$

Case (ii): $t_1^* < \alpha < 1$. In this case

$$t_{1,\alpha} = \sqrt{\alpha} e^{-1}, \qquad v_{1,\alpha} = \frac{-\sqrt{\alpha} \log \alpha}{1 - \alpha} e^{-1}.$$

Example 3.2. Suppose that $m = 2$ and $t_2^* = e^{-\sqrt{6}}$.

Case (i): $0 \leq \alpha \leq t_2^*$. In this case

$$t_{2,\alpha} = e^{-\sqrt{6}}, \qquad v_{2,\alpha} = \frac{3 + \sqrt{6}}{1 - \alpha} e^{-\sqrt{6}}.$$

Case (ii): $t_2^* < \alpha < 1$. In this case

$$t_{2,\alpha} = \sqrt{\alpha} \exp\{-\nu(\alpha)\}, \qquad v_{2,\alpha} = \frac{-\sqrt{\alpha} \log \alpha}{1 - \alpha} (1 + \nu(\alpha)) \exp\{-\nu(\alpha)\},$$

where

$$\nu(\alpha) = \frac{\sqrt{72 - 3(\log \alpha)^2}}{6}.$$

Remark 3.3. (i) When N is a curtailed geometric random variable with parameter $q = 1 - c/n$, the corresponding V is, as is easily seen, a curtailed exponential random variable with parameter c, namely,

$$f_V(v) = \frac{ce^{-cv}}{1 - e^{-c}}, \qquad 0 < v < 1.$$

Therefore, from (3.7),

$$\Psi(z) = \frac{cz}{1 - e^{-c}} \int_1^{1/z} \frac{e^{-czv}}{v} \left[\sum_{j=1}^m \frac{(\log v)^j}{j!} \right] dv,$$

which is consistent with (3.5).

(ii) Let, for $j = 0, 1, \ldots,$

$$F_j(z) = \int_z^1 \frac{\{\log(v/z)\}^j}{j!} \frac{f_V(v)}{v} \, dv.$$

Then, (3.7) can be written as

$$\Psi(z) = \sum_{j=1}^m z F_j(z).$$

We easily find that

$$F_j'(z) = -\frac{1}{z} F_{j-1}(z), \qquad j \geq 1,$$

and so

$$\Psi'(z) = F_m(z) - F_0(z),$$

implying that, if the optimal rule is a threshold rule for a truncation time V, the threshold value is a unique root z of the equation $F_m(z) = F_0(z)$.

4. Continuous arrival time model

Let $Z_1, Z_2, \ldots$ be continuous independent and identically distributed random variables with a common distribution function F, and let M be an integer-valued random variable independent of the Z_ks. The random variable M represents the total number of applicants and Z_k, $1 \leq k \leq M$, denotes the arrival time of the kth best applicant. Unlike N, M is not assumed to be bounded. When we consider the problem of stopping on any of the last m candidates in this framework, the optimal rule becomes very complicated in general, because it depends not only on the arrival time of the candidate but also on the number of arrivals observed up to that time. See, for example, Tamaki and Wang (2010) for the exact form of the optimal rule of the best-choice problem with M uniform on $\{1, 2, \ldots, n\}$.

Here we consider, as a class of suboptimal stopping rules, the t-strategy ($0 < t < 1$), which passes up all the applicants that appear before time t and then stops on the first candidate, if any. The function F can be assumed to be uniform on $[0, 1]$ without loss of generality. Let K_n represent the number of candidates to appear when the total number of applicants is n (i.e. this

corresponds to the case $P\{M = n\} = 1$). Then the probability mass function of K_n is given as

$$p_n(k) = P\{K_n = k\} = \frac{1}{n!}\begin{bmatrix} n \\ k \end{bmatrix}, \qquad 1 \le k \le n,$$

where the notation $\begin{bmatrix} n \\ k \end{bmatrix}$, $1 \le k \le n, 1 \le n$, is a real number called the *Stirling number of the first kind*, and $p_n(k)$ satisfies the recursion

$$p_n(k) = \frac{1}{n}p_{n-1}(k-1) + \left(1 - \frac{1}{n}\right)p_{n-1}(k), \quad 1 \le k \le n, \ 2 \le n, \tag{4.1}$$

with $p_1(1) = 1$ and $p_n(k) = 0$ for $k = 0$ or $k > n$ (see, e.g. Equation (2.5.9) of Arnold *et al.* (1998)). An important identity of the Stirling number of the first kind is, for any positive integer k,

$$\sum_{n=k}^{\infty} \begin{bmatrix} n \\ k \end{bmatrix} \frac{z^n}{n!} = \frac{1}{k!}\left(\log \frac{1}{1-z}\right)^k, \qquad 0 < z < 1, \tag{4.2}$$

as listed in Graham *et al.* (1989, Equation (7.50), p. 337). See also Tamaki (2009) for a probabilistic proof of this identity.

Let $R_n(t)$ represent the (true) rank of the (relatively) best applicant among those that have arrived by time t, if any. If no applicant appears before t, $R_n(t)$ is assumed to be $n + 1$ for convenience. Then conditioning on $R_n(t)$ yields

$$P_n(t) = P\{\text{win by the } t\text{-strategy}\}$$
$$= \sum_{j=0}^{n} P\{\text{win by the } t\text{-strategy} \mid R_n(t) = j + 1\}P\{R_n(t) = j + 1\}. \tag{4.3}$$

Since $R_n(t) = j + 1$ occurs if and only if the $(j + 1)$th best applicant appears before t, while the top j applicants (i.e. best, second best, …, jth best) all appear after t for $0 \le j < n$, we have

$$P\{R_n(t) = j + 1\} = \begin{cases} t(1-t)^j & \text{if } 0 \le j \le n - 1, \\ (1-t)^n & \text{if } j = n. \end{cases}$$

Moreover, given $R_n(t) = j + 1$, we win by the t-strategy if the number of future candidates is greater than 0 but less than $m + 1$. This occurs with probability d_j if we let $d_j = P\{K_j \le m\}$, $j \ge 1$ ($d_0 = 0$ for convention), because, given $R_n(t) = j + 1$, future applicants under consideration are restricted to the top j and their arrival orders are equally likely. Thus, we have, from (4.3),

$$P_n(t) = \sum_{j=1}^{n-1} d_j t(1-t)^j + d_n(1-t)^n, \qquad n \ge 1. \tag{4.4}$$

Note that $P_n(t) = 1 - t$ for $1 \le n \le m$, because $d_j = 1$ for $1 \le j \le m$.

Lemma 4.1. *Some properties of $P_n(t)$ are as follows.*

(i) $P_n(t) \ge P_{n+1}(t)$, $n \ge 1$. *More precisely,*

$$P_n(t) - P_{n+1}(t) = \frac{(1-t)^{n+1}}{n+1}p_n(m). \tag{4.5}$$

(ii) *For* $n \geq 1$,

$$P_n'(t) = \sum_{k=m}^{n-1} p_k(m)(1-t)^k - 1, \qquad P_n''(t) = -\sum_{k=m}^{n-1} k p_k(m)(1-t)^{k-1},$$

where the vacuous sum is assumed to be 0.

(iii) *Let n tend to* ∞. *Then*

$$P_n(t) \to P(t) = t \sum_{i=1}^{m} \frac{(-\log t)^i}{i!}.$$

Moreover, $P(t)$, $0 \leq t \leq 1$, *is maximized at* $t = s_m^*$ *and* $P(s_m^*) = u_m^*$, *where* s_m^* *and* u_m^* *are given in (1.3) and (1.4), respectively.*

Proof. (i) From (4.4),

$$P_n(t) - P_{n+1}(t) = (d_n - d_{n+1})(1-t)^{n+1}.$$

On the other hand, from (4.1),

$$
\begin{aligned}
d_n - d_{n+1} &= \sum_{j=1}^{m} \{p_n(j) - p_{n+1}(j)\} \\
&= \sum_{j=1}^{m} \left\{ p_n(j) - \frac{1}{n+1} p_n(j-1) - \frac{n}{n+1} p_n(j) \right\} \\
&= \frac{1}{n+1} p_n(m).
\end{aligned}
$$

Hence, (4.5) is obtained.

(ii) The proof follows by a straightforward calculation from (4.4).

(iii) As n tends to ∞, $P_n(t)$ in (4.4) converges uniformly to

$$
\begin{aligned}
P(t) &= t \sum_{j=1}^{\infty} d_j (1-t)^j \\
&= t \sum_{j=1}^{\infty} \left[\sum_{i=1}^{m} p_j(i) \right] (1-t)^j \\
&= t \sum_{i=1}^{m} \left[\sum_{j=i}^{\infty} p_j(i)(1-t)^j \right] \\
&= t \sum_{i=1}^{m} \frac{(-\log t)^i}{i!},
\end{aligned}
$$

where the last equality follows from (4.2). The equation $P'(t) = 0$ is solved to yield $t = s_m^*$ and $P(s_m^*) = u_m^*$ is easily checked. This completes the proof.

From Lemma 4.1 we see that $P_n(t)$ has a unique maximum at $t = t_n$ and that, as $n \to \infty$,

$$P_n(t) \searrow P(t) \quad \text{and} \quad t_n \nearrow s_m^*.$$

Also, note that $t_n = 0$ for $n < n^*$ and $t_n > 0$ for $n \geq n^*$, where

$$n^* = \min\{n \geq m + 1 : P_n'(0) > 0\} = \min\left\{n \geq m + 1 : \sum_{k=m}^{n-1} p_k(m) > 1\right\}.$$

Let $H(t)$ be the probability of win by the t-strategy. Then

$$H(t) = \sum_{n \geq 1} P_n(t) \, \mathrm{P}\{M = n\} \geq P(t),$$

implying that, in total ignorance of the distribution of M, the s_m^*-strategy attains the win with probability greater than or equal to u_m^*. This result can be seen as a generalization of the e^{-1}-law for $m = 1$.

Acknowledgements

References

ANO, K, KAKINUMA, H. AND MIYOSHI, N. (2010). Odds theorem with multiple selection chances. *J. Appl. Prob.* **47,** 1093–1104.

ARNOLD, B. C., BARLAKRISHNAN, N. AND NAGARAJA, H. N. (1998). *Records.* John Wiley, New York.

BRUSS, F. T. (1984). A unified approach to a class of best choice problems with an unknown number of options. *Ann. Prob.* **12,** 882–889.

BRUSS, F. T. (2000). Sum the odds to one and stop. *Ann. Prob.* **28,** 1384–1391.

BRUSS, F. T. (2003). A note on bounds for the odds theorem of optimal stopping. *Ann. Prob.* **31,** 1859–1861.

BRUSS, F. T. AND LOUCHARD, G. (2009). The odds algorithm based on sequential updating and its performance. *Adv. Appl. Prob.* **41,** 131–153.

BRUSS, F. T. AND PAINDAVEINE, D. (2000). Selecting a sequence of last successes in independent trials. *J. Appl. Prob.* **37,** 389–399.

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

FERGUSON, T. S. (2006). *Optimal Stopping and Applications.* Available at http://www.math.ucla.edu/~tom/Stopping/Contents.html.

FERGUSON, T. S. (2008). The sum-the-odds theorem with application to a stopping game of Sakaguchi. Preprint.

GRAHAM, R. L., KNUTH, D. E. AND PATASHNIK, O. (1989). *Concrete Mathematics.* Addison-Wesley, Reading, MA.

HILL, T. P. AND KRENGEL, U. (1992). A prophet inequality related to the secretary problem. In *Strategies for Sequential Search and Selection in Real Time* (Amherst, MA, 1990; Contemp. Math. **125**), American Mathematical Society, Providence, RI, pp. 209–215.

HSIAU, S.-R. AND YANG, J.-R. (2002). Selecting the last success in Markov-dependent trials. *J. Appl. Prob.* **39,** 271–281.

IRLE, A. (1980). On the best choice problem with random population size. *Z. Operat. Res.* **24,** 177–190.

PETRUCCELLI, J. D. (1983). On the best-choice problem when the number of observations is random. *J. Appl. Prob.* **20,** 165–171.

PRESMAN, E. L. AND SONIN, I. M. (1972). The best choice problem for a random number of objects. *Theoret. Prob. Appl.* **17,** 657–668.

SAMUEL-CAHN, E. (1995). The best-choice secretary problem with random freeze on jobs. *Stoch. Process. Appl.* **55,** 315–327.

SAMUELS, S. M. (1991). Secretary problem. In *Handbook of Sequential Analysis,* eds B. K. Ghosh and P. K. Sen, Dekker, New York, pp. 381–405.

TAMAKI, M. (2009). A probabilistic proof of an identity related to the Stirling number of the first kind. In *Recent Advances in Stochastic Operations Research II*, eds T. Dohi, S. Osaki and K. Sawaki, World Scientific, Hackensack, NJ, pp. 3–9.

TAMAKI, M. (2010). Sum the multiplicative odds to one and stop. *J. Appl. Prob.* **47,** 761–777.

TAMAKI, M. AND WANG, Q. (2010). A random arrival time best-choice problem with uniform prior on the number of arrivals. In *Optimization and Optimal Control* eds, A. Chinchuluun, *et al.*, Springer, New York, pp. 499–510.