Adv. Appl. Prob. **48,** 726–743 (2016)
*doi:*10.1017/apr.2016.25

URN SAMPLING DISTRIBUTIONS GIVING ALTERNATE CORRESPONDENCES BETWEEN TWO OPTIMAL STOPPING PROBLEMS

MITSUSHI TAMAKI,* *Aichi University*

Abstract

The best-choice problem and the duration problem, known as versions of the secretary problem, are concerned with choosing an object from those that appear sequentially. Let $(B, \boldsymbol{p})$ denote the best-choice problem and $(D, \boldsymbol{p})$ the duration problem when the total number N of objects is a bounded random variable with prior $\boldsymbol{p} = (p_1, p_2, \dots, p_n)$ for a known upper bound n. Gnedin (2005) discovered the correspondence relation between these two quite different optimal stopping problems. That is, for any given prior $\boldsymbol{p}$, there exists another prior $\boldsymbol{q}$ such that $(D, \boldsymbol{p})$ is equivalent to $(B, \boldsymbol{q})$. In this paper, motivated by his discovery, we attempt to find the alternate correspondence $\{\boldsymbol{p}^{(m)}, m \geq 0\}$, i.e. an infinite sequence of priors such that $(D, \boldsymbol{p}^{(m-1)})$ is equivalent to $(B, \boldsymbol{p}^{(m)})$ for all $m \geq 1$, starting with $\boldsymbol{p}^{(0)} = (0, \dots, 0, 1)$. To be more precise, the duration problem is distinguished into $(D_1, \boldsymbol{p})$ or $(D_2, \boldsymbol{p})$, referred to as model 1 or model 2, depending on whether the planning horizon is N or n. The aforementioned problem is model 1. For model 2 as well, we can find the similar alternate correspondence $\{\boldsymbol{p}^{[m]}, m \geq 0\}$. We treat both the no-information model and the full-information model and examine the limiting behaviors of their optimal rules and optimal values related to the alternate correspondences as $n \to \infty$. A generalization of the no-information model is given. It is worth mentioning that the alternate correspondences for model 1 and model 2 are respectively related to the urn sampling models without replacement and with replacement.

Keywords: Secretary problem; best-choice problem; alternate correspondence; duration problem; Bruss extension

2010 Mathematics Subject Classification: Primary 60G40
Secondary 62L15

1. Introduction

In the *best-choice problem*, a version of the secretary problem (see, e.g. Samuels (1991) for a survey), a fixed known number n of rankable objects appear one at a time in random order with all $n!$ permutations equally likely (1 being the best and n the worst). Each time an object appears, we must decide either to select it and stop observing or reject it and continue observing, based on the relative rank of the current object with respect to its predecessors. The objective is to find a stopping rule that maximizes the probability of selecting the best of all n objects. Evidently we can confine our selection to a relatively best object. For ease of description, we often call an object a *candidate* if it is relatively best upon arrival.

As a different version of the secretary problem, Ferguson *et al.* (1992) considered the optimal stopping problem, referred to as the *duration problem*, in the same framework described above. We only select a candidate. Define T_k as the time of the first candidate after k if there is one,

Received 18 August 2015; revision received 30 October 2015.
* Postal address: Department of Business Administration, Aichi University, Nagoya Campus, Hiraike 4-60-6, Nakamura, Nagoya, Aichi, 453-8777, Japan. Email address: tamaki@vega.aichi-u.ac.jp

and as $n+1$ if there is none. Then the duration of holding a candidate selected at time k is $(T_k - k)/n$ (division by n is for normalization) and the objective of this problem is to find a stopping rule that maximizes the expected duration of holding a candidate.

These two classical problems with fixed horizon n were generalized to the problems with random horizon by introducing uncertainty about the number N of the available objects. The selection must be made by time N. See, e.g. Presman and Sonin (1972), Irle (1980), Petruccelli (1983), and Tamaki (2011) for the best-choice problem, and Gnedin (2004), (2005) and Tamaki (2013) for the duration problem. Throughout this paper, we assume that the random variable N, independent of the arrival order of the objects, is bounded by n and has a prior distribution $\boldsymbol{p} = (p_1, p_2, \dots, p_n)$, where $p_k = \mathbb{P}\{N = k\}$ are such that $\sum_{k=1}^{n} p_k = 1$ and $p_n > 0$. It is also assumed that $n \geq 2$, unless otherwise specified. Define, for later use, $\boldsymbol{\pi} = (\pi_1, \pi_2, \dots, \pi_n)$ and $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \dots, \sigma_n)$ as functions of $\boldsymbol{p}$ where, for $1 \leq k \leq n$,

$$\pi_k = p_k + p_{k+1} + \dots + p_n, \qquad \sigma_k = \pi_k + (n-k)p_k.$$

When N has a prior $\boldsymbol{p}$, we simply denote the best-choice problem by $(B, \boldsymbol{p})$ and the duration problem by $(D, \boldsymbol{p})$. Though the objective of $(B, \boldsymbol{p})$ is to select the best of all N objects, $(D, \boldsymbol{p})$ can be distinguished into two problems denoted by $(D_1, \boldsymbol{p})$ or $(D_2, \boldsymbol{p})$ depending on whether the final stage of the planning horizon is N or n. That is, the duration of a candidate selected at time k is defined as $(T_k - k)/n$, as before, but if no further candidate appears by time N, T_k is interpreted as $N+1$ for $(D_1, \boldsymbol{p})$ and as $n+1$ for $(D_2, \boldsymbol{p})$. The problem $(D_k, \boldsymbol{p})$, $k = 1, 2$, is referred to as model k of the duration problem. We denote the optimal values of the problems $(B, \boldsymbol{p})$ and $(D_k, \boldsymbol{p})$ by $v_n^B(\boldsymbol{p})$ and $v_n^{D_k}(\boldsymbol{p})$, respectively, to make explicit the dependence on n and $\boldsymbol{p}$. Note that the classical problems occur if N degenerates to n (i.e. $\boldsymbol{p} = (0, \dots, 0, 1)$), in which case there exists no difference between $(D_1, \boldsymbol{p})$ and $(D_2, \boldsymbol{p})$.

A stopping rule is said to be *simple* if, for a given positive integer $s_n (\leq n)$, it passes over the first $s_n - 1$ objects and stops with the first, if any, candidate. The value s_n is referred to as the *critical* number of the simple rule. It is well known that the optimal rules of the classical problems are simple. However, the form of the optimal rule depends on $\boldsymbol{p}$, implying that it is not necessarily simple. An example of $\boldsymbol{p}$ for which the optimal rule of $(B, \boldsymbol{p})$ when $n = 8$ is not simple is when $\boldsymbol{p} = (p_1, p_2, \dots, p_8)$, where $p_1 = 0$, $p_2 = 0.895$, $p_3 = \dots = p_7 = 0.001$, and $p_8 = 0.1$, see Irle (1980) .

Define

$$\boldsymbol{p}^{(0)} = (0, \dots, 0, 1), \qquad \boldsymbol{p}^{(1)} = \left(\frac{1}{n}, \dots, \frac{1}{n}, \frac{1}{n}\right)$$

as two special priors. Then $\boldsymbol{p}^{(0)}$ corresponds to the fixed horizon and $\boldsymbol{p}^{(1)}$ to the random horizon with N uniform on $\{1, 2, \dots, n\}$. Ferguson *et al.* (1992) recognized the equivalence between $(D_1, \boldsymbol{p}^{(0)})$ and $(B, \boldsymbol{p}^{(1)})$. Extending this equivalence, Gnedin (2005) discovered the further correspondences between model 1 of the duration problem and the best-choice problem (see also Samuels (2004) and Porosiński (2002) for related works). According to Gnedin (2005, Proposition 4.1 and Corollary 4.1), this discovery can be stated as follows for our framework.

Proposition 1.1. (Equivalence between $(D_1, \boldsymbol{p})$ and $(B, \boldsymbol{q})$.) *For any given prior*

$$\boldsymbol{p} = (p_1, p_2, \dots, p_n) \quad \textit{on } N,$$

there exists another prior $\boldsymbol{q} = (q_1, q_2, \dots, q_n)$ *defined from* $\boldsymbol{p}$ *as*

$$\boldsymbol{q} = \frac{\boldsymbol{\pi}}{\mathbb{E}[N]} \tag{1.1}$$

https://doi.org/10.1017/apr.2016.25 Published online by Cambridge University Press

such that $(D_1, \boldsymbol{p})$ *is equivalent to* $(B, \boldsymbol{q})$ *in the sense that these two problems have the same optimal rules. Moreover, their optimal values only differ by the factor* $\mathbb{E}[N]/n$*; namely,*

$$v_n^{D_1}(\boldsymbol{p}) = \frac{\mathbb{E}[N]}{n} v_n^B(\boldsymbol{q}). \tag{1.2}$$

Remark 1.1. Note that (1.2) can be written as $n v_n^{D_1}(\boldsymbol{p}) = \mathbb{E}[N] v_n^B(\boldsymbol{q})$, and that Gnedin considered the left-hand side, i.e. $n v_n^{D_1}(\boldsymbol{p})$, as the optimal value of the duration problem, because the duration is, in his setting, not normalized (see the symbolic expression below his Corollary 4.1). It is also noted that Propositions 1.1 and 1.2, given below, hold not only for the no-information model but also for a wide variety of stochastic processes including the generalized no-information model in Section 2.2 and the full-information model in Section 3.

In order to show that the prior $\boldsymbol{p}^{(2)} = (p_1^{(2)}, p_2^{(2)}, \dots, p_n^{(2)})$, for which $(D_1, \boldsymbol{p}^{(1)})$ is equivalent to $(B, \boldsymbol{p}^{(2)})$, is given by

$$p_k^{(2)} = \frac{2(n-k+1)}{n(n+1)}, \qquad 1 \le k \le n,$$

from (1.1), Gnedin (2005) suggested a problem of finding iteratively an infinite sequence of priors

$$\boldsymbol{p}^{(m)} = (p_1^{(m)}, p_2^{(m)}, \dots, p_n^{(m)}), \qquad m \ge 1,$$

with $\boldsymbol{p}^{(0)} = (0, \dots, 0, 1)$, such that $(D_1, \boldsymbol{p}^{(m-1)})$ is equivalent to $(B, \boldsymbol{p}^{(m)})$. In this paper we are motivated by this suggestion. The set $\{\boldsymbol{p}^{(m)}, m \ge 0\}$ is then referred to as the *alternate correspondence of type-1*.

We have a similar correspondence between model 2 of the duration problem and the best-choice problem, which can be stated as follows.

Proposition 1.2. (Equivalence between $(D_2, \boldsymbol{p})$ and $(B, \boldsymbol{q})$.) *For any given prior*

$$\boldsymbol{p} = (p_1, p_2, \dots, p_n) \quad \textit{on } N,$$

there exists another prior $\boldsymbol{q} = (q_1, q_2, \dots, q_n)$ *defined from* $\boldsymbol{p}$ *as*

$$\boldsymbol{q} = \frac{\boldsymbol{\sigma}}{n} \tag{1.3}$$

such that $(D_2, \boldsymbol{p})$ *is equivalent to* $(B, \boldsymbol{q})$ *in the sense that these two problems have the same optimal rules and the same optimal values. Thus,*

$$v_n^{D_2}(\boldsymbol{p}) = v_n^B(\boldsymbol{q}).$$

Proof. We omit the proof because it is similar to the proof of Gnedin (2005, Proposition 4.1). □

Let $\boldsymbol{p}^{[0]} = (0, \dots, 0, 1)$. Then the set $\{\boldsymbol{p}^{[m]} = (p_1^{[m]}, p_2^{[m]}, \dots, p_n^{[m]}), m \ge 0\}$ is referred to as the *alternate correspondence of type-2* if $(D_2, \boldsymbol{p}^{[m-1]})$ is equivalent to $(B, \boldsymbol{p}^{[m]})$ for all $m \ge 1$. From (1.3), it is easy to see that

$$p_k^{[1]} = \frac{1}{n}, \qquad p_k^{[2]} = \frac{2(n-k)+1}{n^2}, \qquad 1 \le k \le n.$$

Our main concerns are to find the explicit expressions of the two alternate correspondences (type-k corresponds to model k of the duration problem, where $k = 1, 2$). We also examine the optimal rules and the optimal values related to these alternate correspondences. It is of further interest to derive the limiting values of $v_n^B(\boldsymbol{p}^{(m)})$ and $v_n^B(\boldsymbol{p}^{[m]})$ as $n \to \infty$. These are discussed in Section 2.1. In Section 2.2 we generalize the above problems by allowing the objects to appear in accordance with Bernoulli trials. It is worth mentioning that the alternate correspondences of type-1 and type-2 are respectively related to the urn sampling models without replacement and with replacement.

In contrast to the above no-information model, in which the observations are the relative ranks of the objects, the *full-information* model is the problem in which the observations are the true values of N objects $X_1, X_2, \dots, X_N$, assumed to be independent and identically distributed (i.i.d.) random variables from a known continuous distribution, taken without loss of generality to be the uniform distribution on the interval [0, 1]. We also assume that N is independent of $X_1, X_2, \dots$. Let $L_k = \max\{X_1, X_2, \dots, X_k\}$ and call the kth object (or X_k) a candidate if it is a relative maximum, i.e. $X_k = L_k$. Consider a class of stopping rules of the form

$$\tau_N(\boldsymbol{a}) = \min\{k\colon X_k = L_k \geq a_k\} \wedge N,$$

where $\boldsymbol{a} = (a_1, a_2, \dots, a_n)$ is a given sequence of thresholds satisfying the monotone condition $1 \geq a_1 \geq a_2 \geq \cdots \geq a_n \geq 0$. This rule is said to be a *monotone rule* (with thresholds $\boldsymbol{a}$). It is well known that the optimal rules of the classical problems, i.e. problems with fixed horizon, are monotone (cf. Gilbert and Mosteller (1966) and Ferguson *et al.* (1992)). This full-information model is considered in Section 3.

2. Alternate correspondences

Propositions 1.1 and 1.2 show how to construct the distribution which makes the duration problem equivalent to the best-choice problem, but they are of little help for specifying the optimal rules and deriving the optimal values. Fortunately, we can give the explicit expressions of these quantities for the alternate correspondences from the existing literature.

2.1. No-information model

If the optimal rule of $(B, \boldsymbol{p})((D_k, \boldsymbol{p}),\ k = 1, 2)$ is simple, we denote its (optimal) critical number by $s_n^B(\boldsymbol{p})(s_n^{D_k}(\boldsymbol{p}))$. The following results give a sufficient condition for the optimal rule to be simple for each of these problems and also give, when this condition is met, the explicit expressions for the critical numbers and the optimal values.

Lemma 2.1. *Define, for given n and positive vector* $\boldsymbol{t} = (t_1, t_2, \dots, t_n)$ *satisfying the condition that* t_{k+j}/t_k *is nonincreasing in k for each possible value of j,*

$$s_n(\boldsymbol{t}) = \min\left\{k \geq 1\colon \sum_{i=k}^{n} \frac{t_i}{i} \geq \sum_{i=k+1}^{n} \left(\sum_{j=k+1}^{i} \frac{1}{j-1}\right) \frac{t_i}{i}\right\},$$

and then

$$v_n(\boldsymbol{t}) = \frac{s_n(\boldsymbol{t}) - 1}{n} \sum_{i=s_n(\boldsymbol{t})}^{n} \left(\sum_{j=s_n(\boldsymbol{t})}^{i} \frac{1}{j-1}\right) \frac{n t_i}{i} \quad \text{for } s_n(\boldsymbol{t}) \geq 2,$$

and $v_n(\boldsymbol{t}) = \sum_{i=1}^{n} t_i/i$ *for* $s_n(\boldsymbol{t}) = 1$.

The $(B, \boldsymbol{p})$*-problem.* (i) *A sufficient condition for the optimal rule to be simple is that* p_{k+j}/p_k *is nonincreasing in k for each possible value of j.*

(ii) *Suppose that the optimal rule is simple. Then the corresponding critical number and the optimal value are respectively*

$$s_n^B(\boldsymbol{p}) = s_n(\boldsymbol{p}), \tag{2.1}$$

$$v_n^B(\boldsymbol{p}) = v_n(\boldsymbol{p}). \tag{2.2}$$

The $(D_1, \boldsymbol{p})$*-problem.* (iii) *A sufficient condition for the optimal rule to be simple is that* π_{k+j}/π_k *is nonincreasing in k for each possible value of j.*

(iv) *Suppose that the optimal rule is simple. Then the corresponding critical number and the optimal value are given as*

$$s_n^{D_1}(\boldsymbol{p}) = s_n(\boldsymbol{\pi}), \tag{2.3}$$

$$v_n^{D_1}(\boldsymbol{p}) = \frac{1}{n} v_n(\boldsymbol{\pi}).$$

The $(D_2, \boldsymbol{p})$*-problem.* (v) *A sufficient condition for the optimal rule to be simple is that* σ_{k+j}/σ_k *is nonincreasing in k for each possible value of j.*

(vi) *Suppose that the optimal rule is simple. Then the corresponding critical number and the optimal value are given as*

$$s_n^{D_2}(\boldsymbol{p}) = s_n(\boldsymbol{\sigma}), \qquad v_n^{D_2}(\boldsymbol{p}) = \frac{1}{n} v_n(\boldsymbol{\sigma}).$$

Proof. See the proof of Theorem 2.3 because the $(D_2, \boldsymbol{p})$-problem is a special case of the Bruss extension (corresponding to $\boldsymbol{a} = (1, \frac{1}{2}, \frac{1}{3}, \dots, 1/n) \in A^*$). For more detail, see Tamaki (2011) and (2013). □

The following theorem gives the main results concerning the alternate correspondences.

Theorem 2.1. (i) Alternate correspondence of type-1. *Let* $\boldsymbol{p}^{(0)} = (0, \dots, 0, 1)$ *and*

$$p_k^{(m)} = \binom{n+m-1-k}{m-1} \bigg/ \binom{n+m-1}{m}, \qquad 1 \le k \le n, \tag{2.4}$$

for $m \ge 1$*. Then* $\{\boldsymbol{p}^{(m)}, m \ge 0\}$ *is the alternate correspondence of type-1. The optimal rule is simple for both* $(D_1, \boldsymbol{p}^{(m)})$ *and* $(B, \boldsymbol{p}^{(m)})$*, and we have the following relations for* $m \ge 0$*:*

$$s_n^{D_1}(\boldsymbol{p}^{(m)}) = s_n^B(\boldsymbol{p}^{(m+1)}), \tag{2.5}$$

$$v_n^{D_1}(\boldsymbol{p}^{(m)}) = \frac{m+n}{(m+1)n} v_n^B(\boldsymbol{p}^{(m+1)}). \tag{2.6}$$

(ii) Alternate correspondence of type-2. *Let* $\boldsymbol{p}^{[0]} = (0, \dots, 0, 1)$ *and*

$$p_k^{[m]} = \left(\frac{n-k+1}{n}\right)^m - \left(\frac{n-k}{n}\right)^m, \qquad 1 \le k \le n, \tag{2.7}$$

for $m \ge 1$*. Then* $\{\boldsymbol{p}^{[m]}, m \ge 0\}$ *is the alternate correspondence of type-2. The optimal rule is simple for both* $(D_2, \boldsymbol{p}^{[m]})$ *and* $(B, \boldsymbol{p}^{[m]})$*, and we have the following relations for* $m \ge 0$*:*

$$s_n^{D_2}(\boldsymbol{p}^{[m]}) = s_n^B(\boldsymbol{p}^{[m+1]}), \tag{2.8}$$

$$v_n^{D_2}(\boldsymbol{p}^{[m]}) = v_n^B(\boldsymbol{p}^{[m+1]}). \tag{2.9}$$

https://doi.org/10.1017/apr.2016.25 Published online by Cambridge University Press

Proof. For convenience, when N has a prior $\boldsymbol{p}^{(m)}$ ($\boldsymbol{p}^{[m]}$), denote the corresponding N, $\boldsymbol{\pi}$, and $\boldsymbol{\sigma}$ by $N^{(m)}$, $\boldsymbol{\pi}^{(m)}$, and $\boldsymbol{\sigma}^{(m)}$ ($N^{[m]}$, $\boldsymbol{\pi}^{[m]}$, and $\boldsymbol{\sigma}^{[m]}$), respectively.

(i) We first show that, for each m,

$$\boldsymbol{p}^{(m+1)} = \frac{m+1}{m+n}\boldsymbol{\pi}^{(m)} = \frac{1}{\mathbb{E}[N^{(m)}]}\boldsymbol{\pi}^{(m)}. \tag{2.10}$$

Dividing both sides of the well-known identity

$$\binom{n+m-k}{m} = \sum_{j=k}^{n}\binom{n+m-j-1}{m-1}$$

by $\binom{n+m-1}{m}$ immediately yields

$$\frac{m+n}{m+1}p_k^{(m+1)} = \pi_k^{(m)}. \tag{2.11}$$

Summing up with respect to k on both sides yields

$$\frac{m+n}{m+1} = \mathbb{E}[N^{(m)}], \tag{2.12}$$

owing to $\mathbb{E}[N^{(m)}] = \sum_{k=1}^{n}\pi_k^{(m)}$. Then (2.10) is obtained from (2.11) and (2.12). In order to show that the optimal rule is simple, it suffices to show from Lemmas 2.1(i) and 2.1(iii) that $p_{k+j}^{(m)}/p_k^{(m)}$ and $\pi_{k+j}^{(m)}/\pi_k^{(m)}$, for $m \geq 1$, are both nonincreasing in k for each j. For this, see Tamaki (2016, Appendix B, Example (e)) together with the property $p_{k+j}^{(m)}/p_k^{(m)} = \pi_{k+j}^{(m-1)}/\pi_k^{(m-1)}$ obtained from (2.10). Equations (2.5) and (2.6) are immediate consequences from Proposition 1.1 because (2.10) is just (1.1) by taking $\boldsymbol{p}^{(m)}$ as $\boldsymbol{p}$ and $\boldsymbol{p}^{(m+1)}$ as $\boldsymbol{q}$, respectively ((2.5) holds because the same optimal rules have the same critical numbers).

(ii) We have, for each m,

$$\boldsymbol{p}^{[m+1]} = \frac{1}{n}\boldsymbol{\sigma}^{[m]} \tag{2.13}$$

because

$$\pi_k^{[m]} = \sum_{j=k}^{n}p_j^{[m]} = \left(\frac{n-k+1}{n}\right)^m \tag{2.14}$$

and, hence,

$$\begin{aligned}
\sigma_k^{[m]} &= \pi_k^{[m]} + (n-k)p_k^{[m]} \\
&= \left(\frac{n-k+1}{n}\right)^m + (n-k)\left[\left(\frac{n-k+1}{n}\right)^m - \left(\frac{n-k}{n}\right)^m\right] \\
&= n\left[\left(\frac{n-k+1}{n}\right)^{m+1} - \left(\frac{n-k}{n}\right)^{m+1}\right] \\
&= np_k^{[m+1]}.
\end{aligned}$$

In order to show that the optimal rule is simple, it suffices to show from Lemmas 2.1(i) and 2.1(v) that $p_{k+j}^{[m]}/p_k^{[m]}$ and $\sigma_{k+j}^{[m]}/\sigma_k^{[m]}$, for $m \geq 1$, are both nonincreasing in k for

https://doi.org/10.1017/apr.2016.25 Published online by Cambridge University Press

each j. For this, see Tamaki (2016, Appendix B, Example (d)) together with the property $p_{k+j}^{[m]}/p_k^{[m]} = \sigma_{k+j}^{[m-1]}/\sigma_k^{[m-1]}$ obtained from (2.13). Equations (2.8) and (2.9) are then immediate consequences from Proposition 1.2 because (2.13) is just (1.3) by taking $\boldsymbol{p}^{[m]}$ as $\boldsymbol{p}$ and $\boldsymbol{p}^{[m+1]}$ as $\boldsymbol{q}$, respectively.

This completes the proof. □

Remark 2.1. The critical numbers and the optimal values in (2.5), (2.6), (2.8), and (2.9) are computed from Lemma 2.1. It is noted that the validity of (2.5), (2.6), (2.8), and (2.9) can be, in turn, ascertained directly from Lemma 2.1. For example, we obtain (2.5) through

$$\begin{aligned} s_n^{D_1}(\boldsymbol{p}^{(m)}) &= s_n(\boldsymbol{\pi}^{(m)}) \quad \text{from (2.3)} \\ &= s_n^B(\boldsymbol{\pi}^{(m)}) \quad \text{from (2.1)} \\ &= s_n^B\left(\frac{m+n}{m+1}\boldsymbol{p}^{(m+1)}\right) \quad \text{from (2.10)} \\ &= s_n^B(\boldsymbol{p}^{(m+1)}), \end{aligned}$$

where the last equality follows from $s_n(c\boldsymbol{t}) = s_n(\boldsymbol{t})$ for any positive constant c. We obtain (2.6) in a similar manner, taking account of the property $v_n(c\boldsymbol{t}) = cv_n(\boldsymbol{t})$. Equations (2.8) and (2.9) are also obtained similarly. The study of the uniqueness of the optimal rule, as touched on by Szajowski (1992), (1993), might be interesting.

Remark 2.2. The two random variables $N^{(m)}$ and $N^{[m]}$ can be related to sampling balls from an urn without replacement and with replacement, respectively. Suppose that there exists an urn containing $n+m-1$ balls numbered $1, 2, \dots, n+m-1$. We draw m balls randomly from the urn without replacement. Then $N^{(m)}$ denotes the smallest of the m numbers drawn. Suppose that there exists an urn containing n balls numbered $1, 2, \dots, n$. We draw m balls one at a time randomly from the urn with replacement. Then $N^{[m]}$ denotes the smallest of the m numbers drawn.

To obtain the further properties of $N^{(m)}$ and $N^{[m]}$, we briefly review the concept of 'stochastically larger' and the distribution of the smallest value from the uniform sample. We say that $N^{(m)}$ is *stochastically larger* than $N^{[m]}$, written as $N^{(m)} \geq_{\text{st}} N^{[m]}$, if $\mathbb{P}\{N^{(m)} \geq k\} \geq \mathbb{P}\{N^{[m]} \geq k\}$, or, equivalently,

$$\pi_k^{(m)} \geq \pi_k^{[m]}, \qquad 1 \leq k \leq n. \tag{2.15}$$

Let $Y_1, Y_2, \dots, Y_m$ be independent random variables each uniformly distributed on $(0, 1)$. Then it is known that $V_m = \min\{Y_1, Y_2, \dots, Y_m\}$ representing the smallest of these m random variables has a density $f_{V_m}(v) = m(1-v)^{m-1}, 0 \leq v \leq 1$, with $\mathbb{E}[V_m] = 1/(m+1)$.

Corollary 2.1. (i) *We have, for* $m \geq 1$,

$$N^{(m)} \geq_{\text{st}} N^{[m]}.$$

(ii) *If, for a given* m, *we let* $n, k \to \infty$ *in such a way that* $k/n \to v$, *then, for* $0 \leq v \leq 1$,

$$np_k^{(m)} \to f_{V_m}(v), \tag{2.16}$$

$$np_k^{[m]} \to f_{V_m}(v). \tag{2.17}$$

https://doi.org/10.1017/apr.2016.25 Published online by Cambridge University Press

Proof. (i) We show (2.15) by induction on m. This is obvious for $m = 1$. Assume that (2.15) holds for some $m \geq 1$. Then, using the recursive relations

$$\pi_k^{(m+1)} = \frac{n-k+1+m}{n+m}\pi_k^{(m)}, \qquad \pi_k^{[m+1]} = \frac{n-k+1}{n}\pi_k^{[m]},$$

which are straightforward from (2.4), (2.10), and (2.14), and the fact that

$$\frac{n-k+1+m}{n+m} \geq \frac{n-k+1}{n},$$

we have

$$\pi_k^{(m+1)} - \pi_k^{[m+1]} \geq \frac{n-k+1}{n}[\pi_k^{(m)} - \pi_k^{[m]}],$$

thus completing the induction.

(ii) From (2.4) and (2.7), we can write

$$np_k^{(m)} = m\left(1 - \frac{k}{n+1}\right)\left(1 - \frac{k}{n+2}\right)\cdots\left(1 - \frac{k}{n+m-1}\right),$$

$$np_k^{[m]} = m\left(1 - \frac{k}{n}\right)^{m-1} + \sum_{j=0}^{m-2}\frac{1}{n^{m-1-j}}\binom{m}{j}\left(1 - \frac{k}{n}\right)^j, \tag{2.18}$$

verifying (2.16) and (2.17) by letting $k/n \to v$, because the second term on the right-hand side of (2.18) vanishes as $n \to \infty$. □

Remark 2.3. As a result of $N^{(m)} \geq_{\text{st}} N^{[m]}$, we have $\mathbb{E}[N^{(m)}] \geq \mathbb{E}[N^{[m]}]$, or, equivalently, $(m+n)/(m+1) \geq \sum_{k=1}^n (k/n)^m$ from (2.12) and (2.14). From Corollary 2.1(ii) we see that, as $n \to \infty$, the difference fades away between the two urn sampling schemes without replacement and with replacement, implying that both $N^{(m)}/n$ and $N^{[m]}/n$ converge in distribution to V_m.

Now we examine the limiting properties of the optimal values and the optimal rules as $n \to \infty$. The duration problems are solved through the corresponding best-choice problems, so we concentrate on the study of the best-choice problems. First we consider the alternate correspondence of type-1. Let, for $m \geq 0$,

$$s^{(B,m)} = \lim_{n\to\infty}\frac{s_n^B(\boldsymbol{p}^{(m)})}{n}, \qquad v^{(B,m)} = \lim_{n\to\infty} v_n^B(\boldsymbol{p}^{(m)}).$$

It is well known that $s^{(B,0)} = v^{(B,0)} = \mathrm{e}^{-1}$ (see, e.g. Gilbert and Mosteller (1966)) and $s^{(B,1)} = \mathrm{e}^{-2}$ and $v^{(B,1)} = 2\mathrm{e}^{-2}$ (see, e.g. Presman and Sonin (1972)). We have the following results for $m \geq 2$.

Theorem 2.2. *For $m \geq 2$, $s^{(B,m)}$ and $v^{(B,m)}$ are calculated as follows.*

(i) *The value of $s^{(B,m)}$ is given as a solution $x \in (0, 1)$ to*

$$\frac{1}{2}\log^2 x + (1 + h_{m-1})\log x + \sum_{j=1}^{m-1}\frac{1 + h_{m-1} - h_{j-1}}{j}(1-x)^j = 0, \tag{2.19}$$

where $h_k = \sum_{j=1}^k 1/j$, $k \geq 1$ and $h_0 = 0$.

(ii) *The value of* $v^{(B,m)}$ *is given by*

$$v^{(B,m)} = ms \sum_{j=m}^{\infty} \frac{(1-s)^j}{j} = -ms\biggl(\log s + \sum_{j=1}^{m-1} \frac{(1-s)^j}{j}\biggr), \qquad (2.20)$$

where, for ease of notation, $s^{(B,m)}$ *is abbreviated to* s.

Proof. It is noted that, as is easily suggested from (2.2), the probability of selecting the best for $(B, \boldsymbol{p}^{(m)})$ by using a simple rule with critical number i is given by

$$v^B_{n,i}(\boldsymbol{p}^{(m)}) = \frac{i-1}{n} \sum_{k=i}^{n} \biggl(\sum_{j=i}^{k} \frac{1}{j-1}\biggr) \frac{np^{(m)}_k}{k}. \qquad (2.21)$$

If we let $n \to \infty$ and write x, y, and v as the limits of i/n, j/n, and k/n, respectively, $v^B_{n,i}(\boldsymbol{p}^{(m)})$, combined with (2.16), becomes a Riemann approximation to an integral

$$v^{(m)}(x) = x \int_x^1 \biggl(\int_x^v \frac{1}{y}\,\mathrm{d}y\biggr) \frac{m(1-v)^{m-1}}{v}\,\mathrm{d}v, \qquad x \in (0,1). \qquad (2.22)$$

In Appendix A we show that $v^{(m)}(x)$ can be simplified to

$$v^{(m)}(x) = mxg^{(m)}(x), \qquad (2.23)$$

where

$$g^{(m)}(x) = \frac{1}{2}\log^2 x + h_{m-1}\log x + \sum_{j=1}^{m-1} \frac{h_{m-1} - h_{j-1}}{j}(1-x)^j. \qquad (2.24)$$

(i) The value of x that maximizes $v^{(m)}(x)$ is easily found by setting the derivative with respect to x equal to 0 and then solving for x. The value obtained in this manner is obviously $s^{(B,m)}$. Thus, $\mathrm{d}v^{(m)}(x)/\mathrm{d}x = 0$, or, equivalently,

$$g^{(m)}(x) + x\frac{\mathrm{d}g^{(m)}(x)}{\mathrm{d}x} = 0 \qquad (2.25)$$

from (2.23) yields (2.19), because a straightforward calculation from (2.24) gives

$$\frac{\mathrm{d}g^{(m)}(x)}{\mathrm{d}x} = \frac{1}{x}\biggl(\log x + \sum_{j=1}^{m-1} \frac{(1-x)^j}{j}\biggr). \qquad (2.26)$$

(ii) Considering that $v^{(B,m)}$ in (2.20) is obtained as $v^{(m)}(s^{(B,m)})$ and that (2.25) holds for $x = s^{(B,m)}$, we immediately have (2.20) through (2.23) and (2.26).

This completes the proof. □

Note that (2.19) and (2.20) in Theorem 2.2 are still valid for $m = 1$ since the vacuous sum is 0. In Table 1 we present some numerical values of $s^{(B,m)}$ and $v^{(B,m)}$.

The following corollary gives the additional limiting relations if we let, for $m \ge 0$,

$$s^{[B,m]} = \lim_{n\to\infty} \frac{s^B_n(\boldsymbol{p}^{[m]})}{n}, \qquad v^{[B,m]} = \lim_{n\to\infty} v^B_n(\boldsymbol{p}^{[m]}), \qquad s^{(D_1,m)} = \lim_{n\to\infty} \frac{s^{D_1}_n(\boldsymbol{p}^{(m)})}{n},$$

$$v^{(D_1,m)} = \lim_{n\to\infty} v^{D_1}_n(\boldsymbol{p}^{(m)}), \qquad s^{[D_2,m]} = \lim_{n\to\infty} \frac{s^{D_2}_n(\boldsymbol{p}^{[m]})}{n}, \qquad v^{[D_2,m]} = \lim_{n\to\infty} v^{D_2}_n(\boldsymbol{p}^{[m]}).$$

https://doi.org/10.1017/apr.2016.25 Published online by Cambridge University Press

Table 1: Values of $s^{(B,m)}$ and $v^{(B,m)}$ for several m.

	m						
	0	1	2	3	4	5	10
$s^{(B,m)}$	0.3679	0.1353	0.0775	0.0539	0.0412	0.0334	0.0171
$v^{(B,m)}$	0.3679	0.2707	0.2535	0.2469	0.2435	0.2414	0.2372

Corollary 2.2. *We have the following relations for* $m \geq 0$*:*

(i) $s^{[B,m]} = s^{(B,m)}$, $v^{[B,m]} = v^{(B,m)}$;

(ii) $s^{(D_1,m)} = s^{(B,m+1)}$, $v^{(D_1,m)} = (m+1)^{-1} v^{(B,m+1)}$;

(iii) $s^{[D_2,m]} = s^{[B,m+1]}$, $v^{[D_2,m]} = v^{[B,m+1]}$;

(iv) $s^{[D_2,m]} = s^{(D_1,m)}$, $v^{[D_2,m]} = (m+1) v^{(D_1,m)}$.

Proof. Corollary 2.2(i) is obvious because the same argument as in the proof of Theorem 2.2 applies if (2.17) is used instead of (2.16) in (2.21). Of course, this coincidence is intuitively clear from Remark 2.3. Corollaries 2.2(ii) and 2.2(iii) follow from (2.5), (2.6), (2.8), and (2.9). Corollary 2.2(iv) follows from Corollaries 2.2(i)–2.2(iii). □

Remark 2.4. From Corollary 2.2(iv) we see remarkable features between the two models of the duration problem with alternate correspondences. From the first expression in Corollary 2.2(iv) we see that the critical numbers of the two models are asymptotically the same. This does not hold in general (compare, e.g. Lemma 3.2 with Lemma 4.2 for the generalized uniform prior in Tamaki (2013)). The second expression says that the optimal value of model 2 is just $m+1$ times as large as that of model 1 asymptotically. This may be viewed in a sense as the equivalence between two models, because it can be written as $v^{[D_2,m]} = v^{(D_1,m)}/\mathbb{E}[V_m]$ and the right-hand side is interpreted as the normalized value of model 1.

2.2. Bruss extension

Here we attempt to generalize the no-information model by allowing the objects to appear in accordance with Bernoulli trials. A total number N of objects appear one at a time and each object is judged either to be a candidate or not upon arrival. Let I_k, $1 \leq k \leq n$, be the indicator of the event that the kth object is a candidate and suppose that $I_1, I_2, \dots, I_n$ is a sequence of independent Bernoulli random variables with $\mathbb{P}\{I_k = 1\} = a_k$, $1 \leq k \leq n$, where $0 < a_k \leq 1$ for simplicity. Here the best-choice problem must be considered as a problem of choosing the last candidate prior to N. It is easy to see that the no-information model occurs as a special case if $a_k = 1/k$, $1 \leq k \leq n$ (this case satisfies $a_{k+1} = a_k/(1+a_k)$, $1 \leq k \leq n$, starting with $a_1 = 1$, implying that Theorem 2.3 below is applicable).

This generalization was introduced by Bruss (2000) to describe the celebrated odds theorem. Though the optimal rule of the best-choice problem with fixed horizon is simple, the form of the optimal rule depends both on $\boldsymbol{p}$ and $\boldsymbol{a} = (a_1, a_2, \dots, a_n)$. Note that, as in the no-information model, the stopping rule is said to be simple with critical number s_n if it passes over the first $s_n - 1$ objects and stops with the first candidate, if any. Let $b_k = 1 - a_k$ and $r_k = a_k/b_k$, $1 \leq k \leq n$, and define

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

with $B_{k,k} = 1$ for convenience. For given $\boldsymbol{p}$ and $\boldsymbol{a}$, denote simply the best-choice problem by $(B, \boldsymbol{p}, \boldsymbol{a})$ and model k of the duration problem by $(D_k, \boldsymbol{p}, \boldsymbol{a})$, $k = 1, 2$. The optimal values of these problems are denoted by $v_n^B(\boldsymbol{p}, \boldsymbol{a})$ and $v_n^{D_k}(\boldsymbol{p}, \boldsymbol{a})$, and the critical numbers of the optimal rules are denoted by $s_n^B(\boldsymbol{p}, \boldsymbol{a})$ and $s_n^{D_k}(\boldsymbol{p}, \boldsymbol{a})$ if they are simple. The following theorem gives a sufficient condition (on $\boldsymbol{a}$) for the optimal rule to be simple for the alternate correspondences of type-1 and type-2.

Theorem 2.3. *Let*

$$A^* = \left\{\boldsymbol{a} \colon a_{k+1} \le \frac{a_k}{1 + a_k},\ 1 \le k < n \text{ with } 0 < a_1 \le 1\right\}.$$

Then the optimal rules of the problems $(B, \boldsymbol{p}^{(m)}, \boldsymbol{a})$, $(B, \boldsymbol{p}^{[m]}, \boldsymbol{a})$, $(D_1, \boldsymbol{p}^{(m)}, \boldsymbol{a})$, *and* $(D_2, \boldsymbol{p}^{[m]}, \boldsymbol{a})$ *are simple if* $\boldsymbol{a} \in A^*$. *Moreover, for* $\boldsymbol{a} \in A^*$, *we have the following.*

(i) *Alternate correspondence of type-1. We have*

$$s_n^{D_1}(\boldsymbol{p}^{(m)}, \boldsymbol{a}) = s_n^B(\boldsymbol{p}^{(m+1)}, \boldsymbol{a}), \qquad v_n^{D_1}(\boldsymbol{p}^{(m)}, \boldsymbol{a}) = \frac{m+n}{(m+1)n} v_n^B(\boldsymbol{p}^{(m+1)}, \boldsymbol{a}).$$

(ii) *Alternate correspondence of type-2. We have*

$$s_n^{D_2}(\boldsymbol{p}^{[m]}, \boldsymbol{a}) = s_n^B(\boldsymbol{p}^{[m+1]}, \boldsymbol{a}), \qquad v_n^{D_2}(\boldsymbol{p}^{[m]}, \boldsymbol{a}) = v_n^B(\boldsymbol{p}^{[m+1]}, \boldsymbol{a}).$$

These values are computed via the following formulae for the problem $(B, \boldsymbol{p}, \boldsymbol{a})$ *with simple optimal rule:*

$$s_n^B(\boldsymbol{p}, \boldsymbol{a}) = \min\left\{k \ge 1 \colon \sum_{i=k}^{n} p_i B_{k,i} \ge \sum_{i=k+1}^{n} \left(\sum_{j=k+1}^{i} r_j\right) p_i B_{k,i}\right\} \tag{2.27}$$

and

$$v_n^B(\boldsymbol{p}, \boldsymbol{a}) = \sum_{i=s_n}^{n} \left(\sum_{j=s_n}^{i} r_j\right) p_i B_{s_n-1,i}, \tag{2.28}$$

where s_n *is understood to be* $s_n^B(\boldsymbol{p}, \boldsymbol{a})$.

Proof. For given $\boldsymbol{p}$ and $\boldsymbol{a}$, define, for $j = 0, 1, 2,$

$$G^{(j)}(k) = \alpha_k^{(j)} - r_{k+1} \sum_{i=k+1}^{n} B_{k,i} \alpha_i^{(j)}, \qquad 1 \le k < n, \tag{2.29}$$

where $\alpha_i^{(0)} = p_i$, $\alpha_i^{(1)} = \pi_i$, and $\alpha_i^{(2)} = \sigma_i$, for each i. Then Tamaki (2013, Remark 4.1) states that these functions give a unified approach to the three problems to determine whether their optimal rule is simple or not ('$j = 0$' corresponds to the best-choice problem and '$j \neq 0$' to model j of the duration problem). That is, a sufficient condition for the optimal rule to be simple for problem 'j' is that $G^{(j)}(k)$, as a function of k, changes its sign from negative to positive at most once. This condition is clearly met if

$$U^{(j)}(k) = \sum_{i=1}^{n-k} Q_{k,i} \left(\frac{\alpha_{k+i}^{(j)}}{\alpha_k^{(j)}}\right) \tag{2.30}$$

https://doi.org/10.1017/apr.2016.25 Published online by Cambridge University Press

is nonincreasing in k, where $Q_{k,i} = a_{k+1}B_{k+1,k+i} = a_{k+1}b_{k+2}\cdots b_{k+i}$, because (2.29) can be written as

$$G^{(j)}(k) = \alpha_k^{(j)}[1 - U^{(j)}(k)].$$

To prove that $U^{(j)}(k)$ is nonincreasing in k, it suffices to show that the following two properties are satisfied:

(i) $Q_{k,i}$ is nonincreasing in k for each i;

(ii) $\alpha_{k+i}^{(j)}/\alpha_k^{(j)}$ is nonincreasing in k for each i

because each term on the right-hand side of (2.30) is nonnegative, so removing it does not increase the sum. For the alternate correspondences, (ii) has already been shown to hold in Theorem 2.1 (exactly speaking, (ii) is shown for $j = 0, 1$ for type-1 and for $j = 0, 2$ for type-2). Property (i) is equivalently written as $Q_{k-1,i} \geq Q_{k,i}$ or

$$\frac{a_k b_{k+1}}{a_{k+1}} \geq b_{k+i} \quad \text{for each possible } i. \tag{2.31}$$

Hence, it holds that $\boldsymbol{a}$ satisfies (2.31) if $\boldsymbol{a} \in A^*$ because $a_{k+1} \leq a_k/(1 + a_k)$ is equivalent to $a_k b_{k+1}/a_{k+1} \geq 1$ and because $b_{k+i} \leq 1$. Considering that, for $\boldsymbol{p}$ and $\boldsymbol{a}$ satisfying (i) and (ii), (2.27) and (2.28) are given by Tamaki (2011, Equations (2.26) and (2.27)), with $m = 1$, and

$$s_n^{D_1}(\boldsymbol{p}, \boldsymbol{a}) = s_n^B(\boldsymbol{\pi}, \boldsymbol{a}), \qquad v_n^{D_1}(\boldsymbol{p}, \boldsymbol{a}) = \frac{1}{n} v_n^B(\boldsymbol{\pi}, \boldsymbol{a}),$$

$$s_n^{D_2}(\boldsymbol{p}, \boldsymbol{a}) = s_n^B(\boldsymbol{\sigma}, \boldsymbol{a}), \qquad v_n^{D_2}(\boldsymbol{p}, \boldsymbol{a}) = \frac{1}{n} v_n^B(\boldsymbol{\sigma}, \boldsymbol{a})$$

hold from Tamaki (2013, Section 3, Equations (3.11) and (3.12)) and Tamaki (2013, Section 4), respectively, we immediately obtain Theorems 2.3(i) and 2.3(ii) from Propositions 1.1 and 1.2 in a similar manner as in Theorem 2.1. □

3. Full-information model

For the full-information model, the following results give a sufficient condition for the optimal rule to be monotone for each of the three problems and also give, when this condition is met, the explicit expressions for the optimal monotone thresholds and the optimal values.

Lemma 3.1. *For given n and positive vector* $\boldsymbol{t} = (t_1, t_2, \ldots, t_n)$ *satisfying the condition that* t_{k+j}/t_k *is nonincreasing in k for each possible value of j, let, depending on* $\boldsymbol{t}$, $r(\boldsymbol{t}) = \min\{i : t_i \geq \sum_{j=1}^{n-i} t_{i+j}/j\}$ *and define*

$$\boldsymbol{a}(\boldsymbol{t}) = (a_1(\boldsymbol{t}), a_2(\boldsymbol{t}), \ldots, a_n(\boldsymbol{t}))$$

as a vector of monotone thresholds such that $a_k(\boldsymbol{t})$ *is a unique root* $x \in (0, 1)$ *to*

$$\sum_{i=k}^{n} t_i x^i = \sum_{i=k}^{n-1} t_i x^i \sum_{j=1}^{n-i} \frac{t_{i+j}}{t_i} \frac{1 - x^j}{j}, \qquad 1 \leq k < r(\boldsymbol{t}), \tag{3.1}$$

and $a_k(\boldsymbol{t}) = 0$ *for* $r(\boldsymbol{t}) \leq k \leq n$. *Define also*

$$v_n(\boldsymbol{t}) = \sum_{i=1}^{n} \frac{t_i}{i}\left[1 + \sum_{k=1}^{i-1}\sum_{j=k}^{i-1}\left(\frac{1}{j} + \frac{1}{i-j}\right)a_k^j - \sum_{k=1}^{i}(1 + h_{i-k})a_k^i\right], \tag{3.2}$$

where $a_k(\boldsymbol{t})$ *is abbreviated to* a_k.

https://doi.org/10.1017/apr.2016.25 Published online by Cambridge University Press

The $(B, \boldsymbol{p})$*-problem.* (i) *A sufficient condition for the optimal rule to be monotone is that* p_{k+j}/p_k *is nonincreasing in k for each possible value of j.*

(ii) *Suppose that the optimal rule is monotone. Then the corresponding thresholds* $\boldsymbol{a}^B(\boldsymbol{p})$ *and the optimal value* $v_n^B(\boldsymbol{p})$ *are given as*

$$\boldsymbol{a}^B(\boldsymbol{p}) = \boldsymbol{a}(\boldsymbol{p}), \tag{3.3}$$

$$v_n^B(\boldsymbol{p}) = v_n(\boldsymbol{p}). \tag{3.4}$$

The $(D_1, \boldsymbol{p})$*-problem.* (iii) *A sufficient condition for the optimal rule to be monotone is that* π_{k+j}/π_k *is nonincreasing in k for each possible value of j.*

(iv) *Suppose that the optimal rule is monotone. Then the corresponding thresholds* $\boldsymbol{a}^{D_1}(\boldsymbol{p})$ *and the optimal value* $v_n^{D_1}(\boldsymbol{p})$ *are given as*

$$\boldsymbol{a}^{D_1}(\boldsymbol{p}) = \boldsymbol{a}(\boldsymbol{\pi}), \tag{3.5}$$

$$v_n^{D_1}(\boldsymbol{p}) = \frac{1}{n} v_n(\boldsymbol{\pi}). \tag{3.6}$$

The $(D_2, \boldsymbol{p})$*-problem.* (v) *A sufficient condition for the optimal rule to be monotone is that* σ_{k+j}/σ_k *is nonincreasing in k for each possible value of j.*

(vi) *Suppose that the optimal rule is monotone. Then the corresponding thresholds* $\boldsymbol{a}^{D_2}(\boldsymbol{p})$ *and the optimal value* $v_n^{D_2}(\boldsymbol{p})$ *are given as*

$$\boldsymbol{a}^{D_2}(\boldsymbol{p}) = \boldsymbol{a}(\boldsymbol{\sigma}), \tag{3.7}$$

$$v_n^{D_2}(\boldsymbol{p}) = \frac{1}{n} v_n(\boldsymbol{\sigma}). \tag{3.8}$$

Proof. Define, depending on $\boldsymbol{t}$,

$$f_i(\boldsymbol{t}) = h_i + \sum_{k=1}^{i} \sum_{j=k}^{i} \frac{1}{j}(h_{i-j} - h_{j-k} - 1)a_k^j,$$

$$g_i(\boldsymbol{t}) = \frac{1 - a_1^i}{i} + \sum_{j=1}^{i-1} \left[\sum_{k=1}^{j} \frac{a_k^j}{j(i-j)} - \sum_{k=1}^{j} \frac{a_k^i}{i(i-j)} - \frac{a_{j+1}^i}{i} \right] \quad \text{for } 1 \le i \le n,$$

where $a_k(\boldsymbol{t})$ is abbreviated to a_k as before. Then it is a simple matter to show that

$$g_{i+1}(\boldsymbol{t}) = f_{i+1}(\boldsymbol{t}) - f_i(\boldsymbol{t}), \qquad 0 \le i < n, \tag{3.9}$$

with $f_0(\boldsymbol{t}) = 0$. Observe also that (3.2) is written as

$$v_n(\boldsymbol{t}) = \sum_{i=1}^{n} t_i g_i(\boldsymbol{t}).$$

First we dispose of the $(D_1, \boldsymbol{p})$-problem. Lemma 3.1(iii) and (3.5) are just Tamaki (2016, Corollary 3.1 and Theorem 3.1). We have the expression of

$$v_n^{D_1}(\boldsymbol{p}) = \frac{1}{n} \sum_{i=1}^{n} p_i f_i(\boldsymbol{\pi})$$

from Tamaki (2016, Lemma 3.2 and Theorem 3.1) (notations are not consistent). To show (3.6), it suffices to show that

$$\sum_{i=1}^{n} p_i f_i(\boldsymbol{\pi}) = \sum_{i=1}^{n} \pi_i g_i(\boldsymbol{\pi})$$

because of $v_n(\boldsymbol{\pi}) = \sum_{i=1}^{n} \pi_i g_i(\boldsymbol{\pi})$. Putting $p_i = \pi_i - \pi_{i+1}$ and then using (3.9), we have

$$\sum_{i=1}^{n} p_i f_i(\boldsymbol{\pi}) = \sum_{i=1}^{n} \pi_i f_i(\boldsymbol{\pi}) - \sum_{i=1}^{n-1} \pi_{i+1}(f_{i+1}(\boldsymbol{\pi}) - g_{i+1}(\boldsymbol{\pi})) = \sum_{i=1}^{n} \pi_i g_i(\boldsymbol{\pi}),$$

where the last equality follows from $f_1(\boldsymbol{\pi}) = g_1(\boldsymbol{\pi})$. This is the desired result. Note that, in a similar manner, we also have

$$\sum_{i=1}^{n} p_i f_i(\boldsymbol{\sigma}) = \sum_{i=1}^{n} \pi_i g_i(\boldsymbol{\sigma}). \tag{3.10}$$

We now turn to the $(D_2, \boldsymbol{p})$-problem. Lemma 3.1(v) and (3.7) are just Tamaki (2016, Corollary 4.1 and Theorem 4.1). To show (3.8), observe that

$$v_n^{D_2}(\boldsymbol{p}) = \frac{1}{n} \sum_{i=1}^{n} p_i [f_i(\boldsymbol{\sigma}) + (n-i) g_i(\boldsymbol{\sigma})] = \frac{1}{n} \sum_{i=1}^{n} \sigma_i g_i(\boldsymbol{\sigma}),$$

where the first equality follows from Tamaki (2016, Lemma 4.2 and Theorem 4.1) and the second from (3.10). Then the desired result is immediate because $v_n(\boldsymbol{\sigma}) = \sum_{i=1}^{n} \sigma_i g_i(\boldsymbol{\sigma})$.

Finally we dispose of the $(B, \boldsymbol{p})$-problem. Equations (3.3) and (3.4) follow from Porosiński (1987, Theorem 2). More specifically, (3.4) is just Porosiński (1987, Equation (19)) and (3.3) follows because $a_k^B(\boldsymbol{p})$ is the solution x of the Porosiński's equation $c(k, x) = 0$, or, equivalently,

$$\sum_{i=k}^{n} p_i x^{i-k} - \sum_{i=k+1}^{n} x^{i-k-1} \int_x^1 \left(\sum_{j=i}^{n} p_j y^{j-i} \right) \mathrm{d}y = 0,$$

which reduces to (3.1) with t_i replaced by p_i. The sufficient condition Lemma 3.1(i) is derived in a similar manner as for the duration problem (see, e.g. Tamaki (2016, Lemma 3.1 and Corollary 3.1)). □

Since Propositions 1.1 and 1.2 hold for the full-information model as well, we can give the following results analogous to Theorem 2.1.

Theorem 3.1. (i) *Alternate correspondence of type-1. The optimal rule is monotone for both* $(D_1, \boldsymbol{p}^{(m)})$ *and* $(B, \boldsymbol{p}^{(m)})$*, and we have the following relations for* $m \geq 0$*:*

$$\boldsymbol{a}^{D_1}(\boldsymbol{p}^{(m)}) = \boldsymbol{a}^{B}(\boldsymbol{p}^{(m+1)}), \tag{3.11}$$

$$v_n^{D_1}(\boldsymbol{p}^{(m)}) = \frac{m+n}{(m+1)n} v_n^{B}(\boldsymbol{p}^{(m+1)}). \tag{3.12}$$

(ii) *Alternate correspondence of type-2. The optimal rule is monotone for both* $(D_2, \boldsymbol{p}^{[m]})$ *and* $(B, \boldsymbol{p}^{[m]})$*, and we have the following relations for* $m \geq 0$*:*

$$\boldsymbol{a}^{D_2}(\boldsymbol{p}^{[m]}) = \boldsymbol{a}^{B}(\boldsymbol{p}^{[m+1]}), \qquad v_n^{D_2}(\boldsymbol{p}^{[m]}) = v_n^{B}(\boldsymbol{p}^{[m+1]}). \tag{3.13}$$

Note that, for consistency with the propositions, we use the same notation $v_n^B(\boldsymbol{p})$ and $v_n^{D_k}(\boldsymbol{p})$ to denote the optimal values for the full-information model if no confusion occurs.

Remark 3.1. We can give the full-information analogue of Remark 2.1. The optimal thresholds and the optimal values in (3.11)–(3.13) are computed from Lemma 3.1 and the validity of (3.11)–(3.13) can in turn be ascertained from Lemma 3.1. For example, we obtain (3.11) through

$$\begin{aligned}\boldsymbol{a}^{D_1}(\boldsymbol{p}^{(m)}) &= \boldsymbol{a}(\boldsymbol{\pi}^{(m)}) \quad \text{from (3.5)}\\ &= \boldsymbol{a}^{B}(\boldsymbol{\pi}^{(m)}) \quad \text{from (3.3)}\\ &= \boldsymbol{a}^{B}\left(\frac{m+n}{m+1}\boldsymbol{p}^{(m+1)}\right) \quad \text{from (2.10)}\\ &= \boldsymbol{a}^{B}(\boldsymbol{p}^{(m+1)}),\end{aligned}$$

where the last equality follows from $\boldsymbol{a}(c\boldsymbol{t}) = \boldsymbol{a}(\boldsymbol{t})$ for any positive constant c. We obtain (3.12) in a similar manner, taking account of the property $v_n(c\boldsymbol{t}) = cv_n(\boldsymbol{t})$. The equations in (3.13) are also obtained similarly.

We now consider the limiting optimal values as $n \to \infty$. Let, for $m \geq 0$,

$$v^{(B,m)} = \lim_{n\to\infty} v_n^{B}(\boldsymbol{p}^{(m)}), \qquad v^{[B,m]} = \lim_{n\to\infty} v_n^{B}(\boldsymbol{p}^{[m]}),$$
$$v^{(D_1,m)} = \lim_{n\to\infty} v_n^{D_1}(\boldsymbol{p}^{(m)}), \qquad v^{[D_2,m]} = \lim_{n\to\infty} v_n^{D_2}(\boldsymbol{p}^{[m]}).$$

Then, from Theorem 3.1, we obviously have the following results analogous to Corollary 2.2.

Corollary 3.1. *We have the following limiting relations for $m \geq 0$:*

(i) $v^{[B,m]} = v^{(B,m)}$;

(ii) $v^{(D_1,m)} = (m+1)^{-1}v^{(B,m+1)}$;

(iii) $v^{[D_2,m]} = v^{[B,m+1]}$;

(iv) $v^{[D_2,m]} = (m+1)v^{(D_1,m)}$.

The explicit expressions of $v^{(D_1,m)}$ and $v^{[D_2,m]}$ were obtained in Tamaki (2016), however, $v^{(D_1,m)}$ appeared as $v_m^{(1)}$ in Theorem 3.2 and $v^{[D_2,m]}$ as $v_m^{(2)}$ in Theorem 4.2 (the derivation is based on a planar Poisson process approach developed by Gnedin (1996), (2004) or Samuels (2004)). Since $v^{(B,m)} = mv^{(D_1,m-1)}$, for $m \geq 1$, from Corollary 3.1(ii), if we let

$$I(c) = \int_c^{\infty} \frac{\mathrm{e}^{-x}}{x}\,\mathrm{d}x, \qquad J(c) = \int_0^{c} \frac{\mathrm{e}^{x}-1}{x}\,\mathrm{d}x,$$

and introduce the additional functions

$$I_m(c) = \int_c^{\infty} \frac{m!\,\mathrm{e}^{-x}}{x^{m+1}}\,\mathrm{d}x, \qquad K_m(c) = \int_0^{c} \frac{x^m \mathrm{e}^{x}}{m!}\,\mathrm{d}x, \qquad L_m(c) = \int_0^{c} \frac{m!\,\mathrm{e}^{-x}}{x^{m+1}} K_m(x)\,\mathrm{d}x,$$

for $m \geq 0$, we can state $v^{(B,m)}$ as follows.

Theorem 3.2. *For $m \geq 1$, let c_m be a unique root c of*

$$\sum_{k=0}^{m-1} \frac{(-c)^k}{k!}(1 - L_k(c)) = \mathrm{e}^{-c}(1 - J(c)).$$

Then

$$v^{(B,m)} = m\left(\frac{(m-1)!\,K_{m-1}(c)}{c^{m-1}} - \frac{c\mathrm{e}^{c}L_{m-1}(c)}{m+c}\right)\left(\frac{c^{m-1}I_{m-1}(c)}{(m-1)!} - \frac{c^{m}I_{m}(c)}{m!}\right) + \frac{m}{m+c}L_{m-1}(c),$$

where c_m *is abbreviated to* c*. For* $m = 0$*,*

$$v^{(B,0)} = \mathrm{e}^{-c_0} + (\mathrm{e}^{c_0} - c_0 - 1)I(c_0) \approx 0.580\,16,$$

where $c_0 \approx 0.804\,35$ *is a unique root* c *of* $J(c) = 1$*.*

Proof. See Tamaki (2016, Theorem 3.2) for $m \geq 1$ (note that his c_{m-1} is equal to our c_m), and see Samuels (1991) or Berezovskiy and Gnedin (1984) for $m = 0$. ☐

In Table 2 we present some numerical values of c_m and $v^{(B,m)}$.

Appendix A.

It is easy to see that

$$\frac{1}{2}\log^2 x = \sum_{k=1}^{\infty}\frac{h_k}{k+1}(1-x)^{k+1} \tag{A.1}$$

from the power series expansion for $\log x = -\sum_{k=1}^{\infty}(1-x)^k/k$. We first show that, as a generalization of this identity,

$$\sum_{k=1}^{\infty}\frac{h_k}{k+m}(1-x)^{k+m} = \frac{1}{2}\log^2 x - \left(h_{m-1} - \sum_{k=1}^{m-1}\frac{1}{k}(1-x)^k\right)\log x - \sum_{k=1}^{m-1}\frac{h_{m-1}-h_{k-1}}{k}(1-x)^k \quad \text{for } m \geq 1, \tag{A.2}$$

which can be proved by induction on m. For $m = 1$, (A.2) reduces to (A.1). Suppose that (A.2) holds for some m. Considering that $h_k = h_{k+1} - 1/(k+1)$, we have

$$\sum_{k=1}^{\infty}\frac{h_k}{k+m+1}(1-x)^{k+m+1} = \sum_{k=1}^{\infty}\frac{h_k}{k+m}(1-x)^{k+m} - \sum_{k=1}^{\infty}\frac{1}{k(k+m)}(1-x)^{k+m}. \tag{A.3}$$

TABLE 2: Values of c_m and $v^{(B,m)}$ for several m.

	m						
	0	1	2	3	4	5	10
c_m	0.8044	2.1198	3.6925	5.3520	7.0411	8.7423	17.3014
$v^{(B,m)}$	0.5802	0.4352	0.4045	0.3926	0.3865	0.3827	0.3753

https://doi.org/10.1017/apr.2016.25 Published online by Cambridge University Press

The second term is written as

$$\begin{aligned}
&\frac{1}{m}\sum_{k=1}^{\infty}\left(\frac{1}{k}-\frac{1}{k+m}\right)(1-x)^{k+m}\\
&\quad=\frac{1}{m}\left[(1-x)^m\sum_{k=1}^{\infty}\frac{(1-x)^k}{k}-\sum_{k=m+1}^{\infty}\frac{(1-x)^k}{k}\right]\\
&\quad=\frac{1}{m}\left[-(1-x)^m\log x-\left(-\log x-\sum_{k=1}^{m}\frac{(1-x)^k}{k}\right)\right]. \qquad \text{(A.4)}
\end{aligned}$$

Substituting (A.2) (from the induction hypothesis) and (A.4) into the right-hand side of (A.3) yields, after some rearrangements, the right-hand side of (A.2) with m replaced by $m+1$. Thus, the induction is complete.

We are ready to prove (2.23). From (2.22),

$$\begin{aligned}
\frac{v^{(m)}(x)}{mx}&=\int_x^1\left(\int_x^v\frac{1}{y}\,\mathrm{d}y\right)\frac{(1-v)^{m-1}}{v}\,\mathrm{d}v\\
&=\int_x^1\log v\frac{(1-v)^{m-1}}{v}\,\mathrm{d}v-\log x\int_x^1\frac{(1-v)^{m-1}}{v}\,\mathrm{d}v\\
&=\int_0^{1-x}\frac{u^{m-1}}{1-u}\log(1-u)\,\mathrm{d}u-\log x\int_0^{1-x}\frac{u^{m-1}}{1-u}\,\mathrm{d}u. \qquad \text{(A.5)}
\end{aligned}$$

The second integral can be expressed as

$$\int_0^{1-x}\frac{u^{m-1}}{1-u}\,\mathrm{d}u=\int_0^{1-x}u^{m-1}\left(\sum_{j=0}^{\infty}u^j\right)\mathrm{d}u=\sum_{k=m}^{\infty}\frac{(1-x)^k}{k}=-\log x-\sum_{k=1}^{m-1}\frac{(1-x)^k}{k}. \qquad \text{(A.6)}$$

The first integral becomes

$$\begin{aligned}
\int_0^{1-x}\frac{u^{m-1}}{1-u}\log(1-u)\,\mathrm{d}u&=-\int_0^{1-x}u^{m-1}\left(\sum_{j=0}^{\infty}u^j\right)\left(\sum_{i=1}^{\infty}\frac{u^i}{i}\right)\mathrm{d}u\\
&=-\int_0^{1-x}u^{m-1}\left(\sum_{k=1}^{\infty}h_k u^k\right)\mathrm{d}u\\
&=-\sum_{k=1}^{\infty}\frac{h_k}{k+m}(1-x)^{k+m}. \qquad \text{(A.7)}
\end{aligned}$$

Substituting (A.6) and (A.7), combined with (A.2), into (A.5) immediately shows that

$$\frac{v^{(m)}(x)}{mx}=g^{(m)}(x)$$

as desired, where $g^{(m)}(x)$ is defined in (2.24). □

https://doi.org/10.1017/apr.2016.25 Published online by Cambridge University Press

Acknowledgements

The author is grateful to Dr Q. Wang for her help with the numerical evaluations and to an anonymous referee for his/her careful reading and many helpful suggestions.

References

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

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

Ferguson, T. S., Hardwick, J. P. and Tamaki, M. (1992). Maximizing the duration of owning a relatively best object. *Contemp. Math.* **125,** 37–57.

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

Gnedin, A. V. (1996). On the full information best-choice problem. *J. Appl. Prob.* **33,** 678–687.

Gnedin, A. V. (2004). Best choice from the planar Poisson process. *Stoch. Process. Appl.* **111,** 317–354.

Gnedin, A. V. (2005). Objectives in the best-choice problems. *Sequential Anal.* **24,** 177–188.

Irle, A. (1980). On the best choice problem with random population size. *Z. Operat. Res. Ser. A-B* **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.

Porosiński, Z. (1987). The full-information best choice problem with a random number of observations. *Stoch. Process. Appl.* **24,** 293–307.

Porosiński, Z. (2002). On best choice problems having similar solutions. *Statist. Prob. Lett.* **56,** 321–327.

Presman, E. L. and Sonin, I. M. (1972). The best choice problem for a random number of objects. *Theory Prob. Appl.* **17,** 657–668.

Samuels, S. M. (1991). Secretary problems. In *Handbook of Sequential Analysis*, Dekker, New York, pp. 381–405.

Samuels, S. M. (2004). Why do these quite different best-choice problems have the same solutions? *Adv. Appl. Prob.* **36,** 398–416.

Szajowski, K. (1992). On nonzero sum game with priority in the secretary problem. *Math. Japon.* **37,** 415–426.

Szajowski, K. (1993). Double stopping by two decision-makers. *Adv. Appl. Prob.* **25,** 438–452.

Tamaki, M. (2011). Maximizing the probability of stopping on any of the last *m* successes in independent Bernoulli trials with random horizon. *Adv. Appl. Prob.* **43,** 760–781.

Tamaki, M. (2013). Optimal stopping rule for the no-information duration problem with random horizon. *Adv. Appl. Prob.* **45,** 1028–1048.

Tamaki, M. (2016). Optimal stopping rule for the full-information duration problem with random horizon. *Adv. Appl. Prob.* **48,** 52–68.