ORIGINAL ARTICLE

Full-information best choice game with hint

Marek Skarupski[1] ⓘ

Abstract
In the classical full-information best choice problem a decision maker aims to select the best opportunity. His decision is based only on the exact values of the observed sequence. In this paper we consider two modifications of the above problem. We add a second player who can either propose additional information or block the observed object and demand an extortion. Our goal is to establish an optimal reward for the second player and the best moment to interrupt the decision process. The situation when the number of observations tends to reach infinity has been studied.

Keywords Optimal stopping · Best choice problem · Matrix game · Markov chain · Threshold strategy

Mathematics Subject Classification 90C40 · 60G40

1 Introduction and literature review

The best choice problems are the most inspiring problems in modern mathematics. Its origin is the so-called secretary problem. A comprehensive work on this issue can be found in Ferguson (1989). Gilbert and Mosteller (1966) considered different variants of the best choice problems for the very first time and solved by the heuristic arguments. They categorize the rank-based problem as a no-information case (which includes the classical secretary problem). On the other hand, we have the so-called full-information case, where we may base our choice of the stopping time on the true values of the object. This is a much more complex problem. In other words, we can say that the no-information problem is a simplified, full-information problem. It is always possible to

✉ Marek Skarupski
marek.skarupski@pwr.edu.pl

1 Faculty of Pure and Applied Mathematics, Wrocław University of Science and Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland

calculate the rank of the currently observed object by counting how many predecessors were better. Unfortunately, the opposite operation is not possible. In this work we focus on the full-information case. A Markovian approach, which is widely used in this article was presented by Bojdecki (1978). Exact solutions for initial problems were presented by Samuels (1982). Many modifications of the problem have been made. Porosiński (1987) presented a model in which the random horizon was introduced. The link between the infinite problem and planar Poisson process was presented in Gnedin (1996), and following that, the exact results for the initial problem were derived by Gnedin and Miretskiy (2007). The full-information best choice problem where two choices are allowed was presented in Porosinski (1992). Petruccelli (1982) allowed the solicitation in the choice. A modification in which the decision maker can go back only in some fixed number was considered by Tamaki (1986). Recently, Kuchta (2017) derived an optimal strategy for the iterated full-information best choice problem.

The game version of the best choice problem was treated by many authors, e.g. (Porosiński and Szajowski 1996; Sakaguchi and Szajowski 1997; Sakaguchi 1984). The game with a hint was presented in Dotsenko and Marynych (2014). The authors considered the problem for the no-information case. In their version, the decision maker can observe only the ranks of objects.

2 Preliminaries: full-information best choice problem

Consider a probability space $(\Omega, \mathcal{F}, P)$. By $E[\cdot]$ we denote the usual expectation with respect to the probability measure P. Fix $n \in \mathbb{N}$ and consider an i.i.d. sequence $X_1, \ldots, X_n$ from continuous distribution $F(x)$. Without loss of generality we can assume that it is a uniform distribution $\mathcal{U}(0, 1)$, i.e. $F(x) = x, x \in [0, 1]$. Define a filtration

$$\mathcal{F}_k = \sigma(X_1, \ldots, X_k), \quad k = 1, \ldots, n.$$

By $\mathcal{T}$ denote a set of all stopping moments with respect to the family $(\mathcal{F}_k)_{k=1,\ldots,n}$. The aim is to find the stopping moment $\tau^* \in \mathcal{T}$ such that

$$E\left[\mathbb{I}_{X_{\tau^*}=\max\{X_1,\ldots,X_n\}}\right] = \sup_{\tau \in \mathcal{T}} E\left[\mathbb{I}_{X_\tau=\max\{X_1,\ldots,X_n\}}\right], \tag{1}$$

where $\mathbb{I}_A(\omega)$ denotes an indicator of the set A

$$\mathbb{I}_A(\omega) = \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A. \end{cases} \tag{2}$$

The moments of consecutive local maximum (cf. Bojdecki 1978) are given by

$$\tau_1 \equiv 1, \quad \tau_{j+1} = \inf\{k : \tau_j \leq k \leq n, X_k = \max\{X_1, \ldots, X_k\}\}. \tag{3}$$

By $\mathcal{T}_0$ let us denote the set of the moments defined in (3). Note that $\mathcal{T}_0 \subset \mathcal{T}$. Consider the following sequence

$$\xi_j = (\tau_j, X_{\tau_j}) \quad \{\tau_j < \infty\} \tag{4}$$

$\tau_j \in \mathcal{T}_0$ for all j. In the case when $\{\tau_j = \infty\}$ we introduce a special absorbing state δ. The sequence in (4) is a homogeneous Markov chain on the state space $\mathfrak{E} = (1, \ldots, n) \times (0, 1) \cup \delta$ with sigma algebra $\mathcal{E}$. One step transition probabilities (i.e. $P(\xi_{j+1} = \mathfrak{b} | \xi_j = \mathfrak{a})$) for the above chain are defined as

$$p(\mathfrak{a}, \mathfrak{b}) = \begin{cases} x^{m-k-1} \int_B dy & \mathfrak{a} = (k, x), \mathfrak{b} = (m, B) \in (k+1, \ldots, n) \times (x, 1], \\ x^{n-k}, & \mathfrak{a} = (k, x), \mathfrak{b} = \delta, k \leq n \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

It is sufficient to find the optimal stopping time in the set $\mathcal{T}_0$. Knowing that kth object is relatively the best and its value is $X_k = x$ in case of selecting it we obtain a gain function given by

$$g(\mathfrak{a}) = \begin{cases} x^{n-k}, & \mathfrak{a} = (k, x) \quad k = 1, \ldots, n; \quad x \in (0, 1) \\ 0, & \mathfrak{a} = \delta. \end{cases} \tag{6}$$

(6) is provided by property of a Markov chain (4). Let T be an operator of a conditional expectation and $V(\cdot)$ be the value function of the problem (cf. Shiryayev 1978). From general theory we know that V satisfies a Bellman equation

$$V(\mathfrak{a}) = \max \{g(\mathfrak{a}), TV(\mathfrak{a})\}.$$

Let $E_{\mathfrak{a}}[g(\xi_1)]$ be the expected payoff for one step starting from the state $\mathfrak{a}$.

$$E_{\mathfrak{a}}[g(\xi_1)] = Tg(\mathfrak{a}) = \sum_{j=1}^{n-k} x^{j-1} \int_x^1 y^{n-(k+j)} dy = x^{n-k} \sum_{m=1}^{n-k} \frac{x^{-m} - 1}{m}. \tag{7}$$

Consider the set of states where the inequality $g(\mathfrak{a}) \geq Tg(\mathfrak{a})$ holds. Let

$$D_k = \{x : \mathfrak{a} = (k, x), g(\mathfrak{a}) \geq Tg(\mathfrak{a})\} = \left\{ x : 1 \geq \sum_{m=1}^{n-k} \frac{x^{-m} - 1}{m} \right\}. \tag{8}$$

Since the problem is monotone the One-Step-Look-Ahead rule is optimal (see Bojdecki 1978) and the optimal region is

$$D = \{\mathfrak{a} = (k, x) : x \geq d_{n-k}, 1 \leq k \leq n\} \tag{9}$$

where $d_{n-k} \in (0, 1)$ is the solution of equality

$$\sum_{m=1}^{n-k} \frac{x^{-m} - 1}{m} = 1. \tag{10}$$

The optimal stopping rule is given by

$$\tau^* = \inf \{k : 1 \leq k \leq n, X_k = \max\{X_1, \ldots, X_k\}, X_k \geq d_{n-k}\}. \tag{11}$$

The value of the problem (cf. Sakaguchi 1973) is

$$V(0, 0) = \frac{1}{n} \left[1 + \sum_{k=1}^{n-1} \sum_{l=1}^{k} \frac{d_k^{n-l}}{n-l} \right]. \tag{12}$$

Let $i = n - k$. Then, d_i is an increasing sequence: $d_0 \leq d_1 \leq \cdots \leq d_{n-1}$. We show some elementary properties of these thresholds. Recall Bernoulli's inequality.

Theorem 1 *For $a \geq -1$ and $b \geq 1$*

$$(1 + a)^b \geq 1 + ab.$$

The proof can be found in Bullen (2013) and here it is omitted.

Lemma 1 *For any i*

$$d_i \geq 1 - \frac{1}{i+1}$$

Proof Since the sum in (10) is monotonically decreasing as a function of x and it is greater than 1 for all $x < d_i$ it is sufficient to show that the inequality

$$\sum_{m=1}^{i} \frac{\left(1 + \frac{1}{i}\right)^m - 1}{m} \geq 1$$

holds for every $i \geq 1$.

By Theorem 1 we have

$$\sum_{m=1}^{i} \frac{\left(1 + \frac{1}{i}\right)^m - 1}{m} \geq \sum_{m=1}^{i} \frac{1 + \frac{m}{i} - 1}{m} = 1$$

which proves our assertion. $\square$

For an upper limit of d_i we refer to Gilbert and Mosteller (1966). Then, we get

$$\frac{i}{i+1} \leq d_i \leq \frac{i}{i+z} \tag{13}$$

where z is the unique solution of the equation

$$\int_0^z \frac{e^t - 1}{t} dt = 1$$

in the interval $(0, 1)$. $z \approx 0.804354$. Let us check some properties of the left hand side of the function from the Eq. (10). Let $f_i(x)$ denote a sequence of functions described as

$$f_i(x) = \sum_{m=1}^{i} \frac{x^{-m} - 1}{m}, \quad i = 1, 2, \ldots. \tag{14}$$

Let us write the formula in a recursive form

$$f_1(x) = \frac{1 - x}{x}$$

$$f_{i+1}(x) = f_i(x) + \frac{x^{-i-1} - 1}{i + 1}, \quad i = 2, 3, \ldots.$$

Now, it is easy to verify that

$$f_i(d_{i+1}) = \frac{i + 2 - d_{i+1}^{-i-1}}{i + 1},$$

$$f_{i+1}(d_i) = \frac{d_i^{-i-1} + i}{i + 1}.$$

3 The best choice problem with hint

3.1 The model

Suppose that except the decision maker (further denoted as DM) in the full-information best choice problem there is another player. However, he does not make a decision about stopping and choosing the best object since he has extra information about the best object, i.e. he knows exactly both position and the value of the current element. We will call him a prompter or a prophet (further denoted as PR). His aim is to sell this information in a proper moment and get for it as much as possible. PR must establish the price α for the hint before the beginning of the game and he can sell his knowledge only once during the game. The decision maker can accept this proposition, pay a fixed price and get information whether the current object is the best one or not. He can also reject the purchase option and then stop or continue observations.

The above game can be presented as a graph, i.e. as a game in an extensive form in each moment k and the actual value of the observed object x, i.e. in the state $\mathfrak{a} = (k, x)$. The payoff function for PR is written at the bottom of the graph.
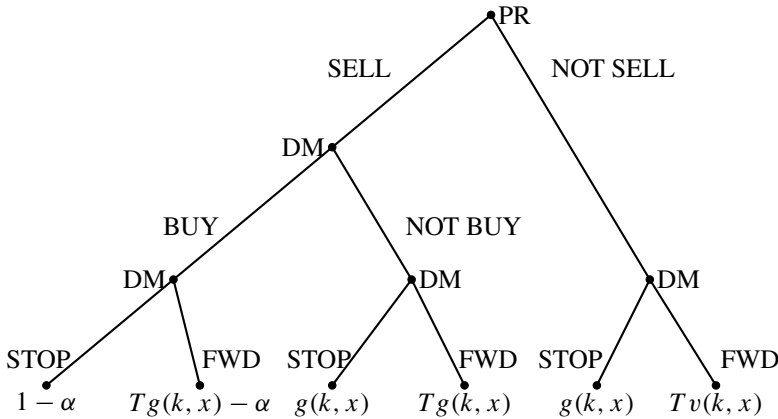
The goal is to establish the price of the hint $\alpha = \alpha(k, x)$. Consider a Markov chain $(\xi_k)_{k=1}^{n}$ observed by DM as in (4) in the state space $(\mathfrak{E}, \mathcal{E})$, transition probabilities (5) and $\mathcal{F}_k = \sigma(\xi_1, \ldots, \xi_k)$. Denote by ρ the strategy of PR, i.e. stopping moments with respect to the family $\mathcal{F}_k$. Let $\tau, \hat{\tau}$ denote the stopping moments of DM and let δ_k be a random variable which has value 1, if the proposition of the hint is accepted and 0 otherwise. If the offer is accepted, the history of observations will be enriched by the random variable H_k. If the event $\{\omega : \rho(\omega) = k\}$ occurs, the strategies of DM will change into two dimensional (δ_ρ, τ_ρ), where

$$\tau_\rho = \hat{\tau}\delta_\rho + (1 - \delta_\rho)\tau. \tag{15}$$

τ_ρ are stopping moments with respect to $\hat{\mathcal{F}}_k = \sigma(\mathcal{F}_k, \delta_1 \cdot H_1, \ldots, \delta_k \prod_{j=1}^{k-1}(1 - \delta_j) \cdot H_k)$. Let us introduce the concept of the hint. In fact, the hint is an indicator function of the absolutely maximal element in the observed sequence. We can denote it as

$$H_k := \mathbb{I}_{\{X_k=\max\{X_1,\ldots,X_n\}\}} = \begin{cases} 1, & X_k = \max\{X_1, \ldots, X_n\} \\ 0, & \text{otherwise} \end{cases} \quad k = 1, \ldots, n. \tag{16}$$

Suppose that we are in the state (k, x), i.e. in a moment k, we observe a locally maximal object X_k whose value is x, $x \in (0, 1)$. There are two possibilities: $x < d_{n-k}$ and $x \geq d_{n-k}$.

Consider the case $X_k = x$, $x < d_{n-k}$. Then, the optimal rule calls for continuing the observations, so the reward function (win probability) is

$$TV(k, x) = \sum_{j=k+1}^{n} x^{j-k-1} \left(\int_{x}^{d_{n-j} \vee x} V(j, y)dy + \int_{d_{n-j} \vee x}^{1} g(j, y)dy \right) \tag{17}$$

where $a \vee b = \max\{a, b\}$. In case of using the hint, the decision maker can get the information "this is the best object among all" with probability x^{n-k} or the opposite information with probability $1 - x^{n-k}$. In the first case the decision maker will stop and

choose the object. Otherwise, he will continue the observations in an optimal manner. Thus, the win probability is

$$E_{(k,x)}[\max\{H_k, TV(k, X_k)\}] = x^{n-k} + \left(1 - x^{n-k}\right) TV(k, x) \qquad (18)$$

We define the value of the hint v_1 as a difference between a reward with the hint and a reward without the hint, i.e the difference between (18) and (17)

$$v_1(k, x) = x^{n-k} (1 - Tv(k, x)) \qquad (19)$$

In case of $X_k = x$, $x \geq d_{n-k}$ the optimal rule calls for a stop immediately. The win probability is

$$g(k, x) = x^{n-k}. \qquad (20)$$

If the decision maker decides to use the hint, the payoff is

$$E_{(k,x)}[\max\{H_k, Tg(k, X_k)\}] = x^{n-k} + \left(1 - x^{n-k}\right) \sum_{j=1}^{n-k} x^{j-1} \int_x^1 y^{n-(k+j)} dy \qquad (21)$$

which gives the value of the hint: the difference between (21) and (20):

$$v_2(k, x) = x^{n-k} \left(1 - x^{n-k}\right) \sum_{m=1}^{n-k} \frac{x^{-m} - 1}{m}. \qquad (22)$$

Fact 1 *Let $x \leq d_{n-k}$ for $k \in \{1, \ldots, n\}$. The function $v_1(k, x)$ is an increasing function of x.*

Proof Since $Tv(k, x)$ is decreasing as x goes to the threshold d_{n-k} the whole function as a multiplication of increasing functions is increasing. □

For the function $v_2(i, x)$ we have

Fact 2 *Let $x \geq d_{n-k}$ for $k \in \{1, \ldots, n\}$. Then, the function $v_2(k, x)$ is a decreasing function of x.*

Proof Let us calculate the derivative of function v_2 with respect to x and let $i = n - k$. We obtain

$$\frac{d}{dx} v_2(n - i, x) = ix^{i-1} \left(1 - 2x^i\right) \sum_{m=1}^{i} \frac{x^{-m} - 1}{m} - \frac{1 - 2x^i + x^{2i}}{x(1 - x)}.$$

The derivative is negative if

$$x^i(i - ix) \left(1 - \left(\frac{x^i}{1 - x^i}\right)^2\right) \sum_{m=1}^{i} \frac{x^{-m} - 1}{m} \leq 1 \qquad (23)$$

for all x from the domain of the function $v_2(i, x)$. From the description of the problem we know that

$$\sum_{m=1}^{i} \frac{x^{-m} - 1}{m} \le 1$$

and $x^i \le 1$ and

$$\left(1 - \left(\frac{x^i}{1 - x^i}\right)^2\right) \le 1.$$

The inequality $i(1 - x) \le 1$ holds for all $x \ge 1 - \frac{1}{i} \ge d_i$ (see Lemma 1), so we conclude that the derivative is negative and the function $v_2(k, x)$ is the decreasing function of x. □

For the fixed index k the maximal value of v_1 is

$$\sup_{x \in [0, d_{n-k}]} v_1(k, x) = v_1(k, d_{n-k}) = d_{n-k}^{n-k} \left(1 - d_{n-k}^{n-k} \sum_{m=1}^{n-k} \frac{d_{n-k}^{-m} - 1}{m}\right)$$

and also

$$\sup_{x \in [0, d_{n-k}]} v_2(k, x) = v_2(k, d_{n-k}) = d_{n-k}^{n-k} \left(1 - d_{n-k}^{n-k}\right) \sum_{m=1}^{n-k} \frac{d_{n-k}^{-m} - 1}{m}.$$

and since the sum in the above formulas is equal to 1 (see 10) we get that

$$a_{n-k} := v_1(k, d_{n-k}) = v_2(k, d_{n-k}) = d_{n-k}^{n-k} \left(1 - d_{n-k}^{n-k}\right). \tag{24}$$

Lemma 2 *Let $i = n - k$. The sequence a_i is decreasing in i and*

$$\lim_{i \to \infty} a_i = e^{-z}\left(1 - e^{-z}\right) \approx 0.2472308,$$

where z is given by (13).

Proof A sequence $c_i = d_i^i$ is decreasing and converges to e^{-z}. (cf. Sakaguchi 1973). It is also bounded since $e^{-z} \le c_i \le 0.5$. Consider a function $f(x) = x(1 - x)$. For $x < 0.5$ it is increasing. Therefore, we get that a product $d_i^i(1 - d_i^i)$ is decreasing. The product is also bounded and converges to the product of the limits of sequences c_i and $1 - c_i$.

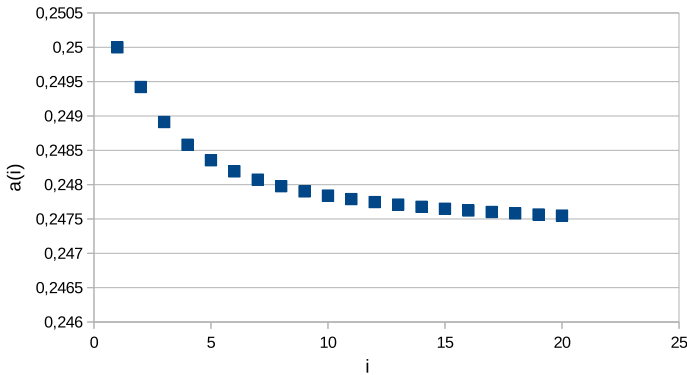$$\lim_{i \to \infty} d_i^i(1 - d_i^i) = e^{-z}(1 - e^{-z}) \approx 0.2472308.$$

□

Fig. 1 a_i for first 20 observations

Note that this value is greater than the value for the no-information case (where it is equal to $e^{-1}(1 - e^{-1}) \approx 0.232544$). Figure 1 shows first 20 values of a_i.

Let us recall the principle of optimality. An optimal policy has the property that whatever the initial state and the initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision (cf. Bellman 1957). To find the exact form of the value function $v_1(k, x)$ we can consider the remaining observations as the best choice problem for a random horizon. To be more specific, let us consider the following

Lemma 3 *Suppose in the full information best choice problem with finite horizon n the current state of the process is (k, x) for some $k \in \{1, 2, \ldots, n\}$ and $x \in (0, 1)$. Suppose that the process has not been stopped yet. Then, the optimal strategy is to stop at the first (if any) state $(k + m, u)$ such that $0 \leq m \leq n - k$ and $u \geq u_{n-k-m}(x)$, where $u_{n-k-m}(x)$ is the solution of the equation*

$$\sum_{j=m}^{n-k} \left(\frac{u - x}{1 - x}\right)^{j-m} \left(\frac{1 - x}{x}\right)^j \left[1 - \sum_{i=1}^{n-k-j} \binom{n - k}{i + j}\frac{(1 - x)^i - (u - x)^i}{i x^i}\right] = 0. \tag{25}$$

The win probability of using the optimal strategy is

$$P(win) = \sum_{m=1}^{n-k} \binom{n - k}{m}(1 - x)^m x^{n-k-m} \sum_{j=1}^m P_m(j), \tag{26}$$

where

$$P_0(0) = x^{n-k}, \tag{27a}$$

$$P_m(0) = 0,$$

$$P_m(1) = \frac{1 - u_{n-k-1}^m(x)}{m},$$

$$P_m(j+1) = \sum_{i=1}^{j} \frac{u_{n-k-i}^j(x)}{j(m-j)} - \sum_{i=1}^{j} \frac{u_{n-k-i}^m(x)}{m(m-j)} - \frac{u_{n-k-j+1}^m(x)}{m}$$

$$for\ 1 \leq j \leq m-1. \tag{27b}$$

Proof Since we observe the current object whose value is x and it is a relative maximum, we can truncate the further chain only in those observations that are greater than x. The probability that observations in moments $k+1, k+2, \ldots, n-k$ will be bigger than x is $1-x$. Therefore, from now, we consider a full information best choice problem with random horizon with observations from the uniform distribution on the interval $(x, 1)$. The horizon M is binomially distributed, i.e.

$$P(M=m) = \binom{n-k}{m}(1-x)^m x^{n-k-m}, m = 1, \ldots, n-k.$$

Consider the following sequence

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

From Porosiński (1987) we know that if the above sequence changes the sign K-times, then, the stopping region has no more than K stopping islands. However here $\{d(m,u)\}_{m=0}^{n-k}$ changes the sign at most one time. When k is close to n its value decreases to 0. So, the truncated problem is monotone and the optimal strategy is a threshold strategy. The thresholds $u_{n-k-m}(x)$ can be calculated directly from

$$\sum_{j=m}^{n-k} \left(\frac{u-x}{1-x}\right)^{j-m} \left(\frac{1-x}{x}\right)^j \left[1 - \sum_{i=1}^{n-k-j} \binom{n-k}{i+j} \frac{(1-x)^i - (u-x)^i}{ix^i}\right] = 0.$$
$$\tag{28}$$

(25). □

Remark 1 In the classical version of the best choice problem with random horizon provided by Porosiński, the payoff function forces the decision maker to make at least one step. It is not possible to stop at the very beginning or, in the language of Markov chains, at stage (0,0). However, here in the truncated problem, such a possibility exists because the payoff for the "initial" stage can be bigger than the expected one. The "initial" value x of the current object must be bigger than a threshold value. This threshold is $u_{n-k}(x)$. It can be calculated from (25) for $m=0$

$$\sum_{j=1}^{n-k} \binom{n-k}{j} \frac{(1-x)^j - (u-x)^j}{jx^j} = 1$$

Since in this special state $u = x$ we see that the $u_{n-k}(x)$ is the unique solution of

$$\sum_{j=1}^{n-k} \binom{n-k}{j} \frac{(1-x)^j}{jx^j} = 1$$

The same situation is in the case of (10) (this equivalence was shown in Samuels 1982). So, $u_{n-k}(x) = d_{n-k}$.

The prompter has two strategies to choose from: to sell the information or not to sell. Then, the DM has to choose either to buy the hint or not. However, if the price of the hint is less than the maximum value of the hint, the decision maker without a doubt will buy the hint. So, the PR has to decide before the game what the price for the hint will be. He has the following possibilities:

1. Set the constant price α during the whole game
2. Set the vector of the prices depending on the moment of the game: $\alpha = (\alpha_1, \ldots, \alpha_n)$
3. Set the price function depending on the value of the current observation: $\alpha = \alpha(x)$
4. Set the vector of the prices depending on the the moment of the game and the current value of the observed object: $\alpha = (\alpha_1(x), \ldots, \alpha_n(x))$

3.2 $\alpha = const$

Consider the following numbers

$$\underline{x}_k(\alpha) = \inf \{x \in (0, d_{n-k}) : v_1(k, x) \geq \alpha\}, \tag{29a}$$
$$\overline{x}_k(\alpha) = \sup \{x \in (d_{n-k}, 1) : v_2(k, x) \geq \alpha\}. \tag{29b}$$

There are three possibilities of the value of the price:

- $\alpha \geq 0.25$: then, the hint is not worth buying. The price is higher than its value.
- $e^{-z}(1 - e^{-z}) < \alpha < 0.25$: then, the hint is worth buying for $k = k^*, k^* + 1, k^* + 2, \ldots, n$, where
$$k^* = \min \{1 \leq k \leq n : a_{n-k} \geq \alpha\}. \tag{30}$$
- $\alpha < e^{-z}(1 - e^{-z})$: then, the hint is worth buying for $k = 1, 2, \ldots, n$ no matter how big is n. Using the previous symbol we can say that in this case $k^* = 1$.

Suppose that $\alpha < e^{-z}(1 - e^{-z})$. The hint will be sold if the current state of the process is in the set

$$S(\alpha) = \{(k, x) : 1 \leq k \leq n, \ \underline{x}_k(\alpha) \leq x \leq \overline{x}_k(\alpha)\}. \tag{31}$$

Suppose that in the moment k we observe a relatively maximal element whose value is x and it is worth buying a hint. Therefore, the probability of that event is given by

$$p_k(\alpha) = (\overline{x}_k(\alpha) - \underline{x}_k(\alpha)) \prod_{j=1}^{k-1} \left(\underline{x}_{j+1}(\alpha)\right)(1 - c_k(\alpha)), \quad k = 1, \ldots, n$$

where

$$c_k(\alpha) = \sum_{l=1}^{k} \sum_{i=1}^{l-1} \left(\prod_{j=1}^{i} \underline{x}_j(\alpha) \right)^{-1} \left[\frac{\underline{x}_k(\alpha) \left((\underline{x}_{l-1} \wedge \overline{x}_k)^i - (\underline{x}_l \wedge \overline{x}_k)^i \right)}{i} \right.$$
$$\left. - \frac{(\underline{x}_{l-1} \wedge \overline{x}_k)^{i+1} - (\underline{x}_l \wedge \overline{x}_k)^{i+1}}{i+1} \right],$$

where $a \wedge b = \min\{a, b\}$. The average payoff for the hint is equal to

$$g(\alpha) = \alpha \sum_{k=1}^{n} p_k(\alpha). \tag{32}$$

The optimal price is such a minimal number α that maximizes the Eq. (32):

$$\alpha^* = \inf \left\{ \alpha > 0 : g(\alpha) = \sup_{\alpha} g(\alpha) \right\} \tag{33}$$

4 The best choice problem with extortion

4.1 The model

In this case, the prompter who knows the exact value of the hint does not want to sell the knowledge as it was in Sect. 3. During the whole game he can block the current element once and demand from the second player to unlock the hidden element. The decision maker has two strategies: to pay an amount of money and stop at the unlocked element or do not pay and continue observations. The graph below presents the possible strategies of both players.
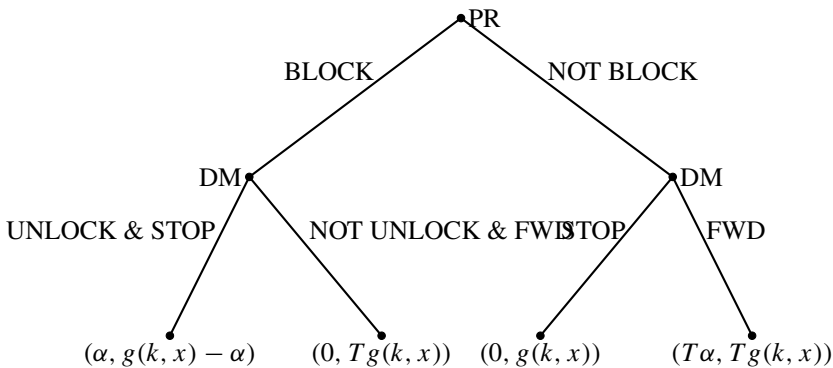
Suppose that we are in the state (k, x), i.e. in a moment k we observe an object X_k, whose value is x, $x \in (0, 1)$. There are two possibilities: $x < d_{n-k}$ and $x \geq d_{n-k}$.

Since the DM will not choose the object if $x < d_{n-k}$ let us consider the case when $x \geq d_{n-k}$. The PR can hide the object and demand a fixed price α. Therefore, his

payoff is α. The DM has two possibilities. The first is to pay the tribute and stop at the object. His payoff is in this case

$$\varphi_{1,\alpha}(k, x) = g(k, x) - \alpha = x^{n-k} - \alpha. \tag{34}$$

Otherwise, he will continue the observations and earn

$$\varphi_2(k, x) = Tg(k, x) = \sum_{j=1}^{n-k} x^{j-1} \int_x^1 y^{n-k-j} dy. \tag{35}$$

The DM will pay the tribute if inequality $\varphi_{1,\alpha}(k, x) \geq \varphi_2(k, x)$ holds. This is equivalent to

$$\alpha \leq x^{n-k} \left(1 - \sum_{m=1}^{n-k} \frac{x^{-m} - 1}{m} \right). \tag{36}$$

Note that the function on the right-hand side of the inequality is increasing as $d_{n-k} \leq x \leq 1$. The set of the states when it is worth to pay the money is defined as

$$\mathcal{T}(\alpha) = \{(k, x) : 0 \leq x \leq 1, \ k = 1, \ldots, n; \ x \geq t_{n-k}(\alpha)\}. \tag{37}$$

where $t_{n-k}(\alpha)$, $\alpha \in [0, 1]$ is the solution of the equation

$$\alpha = x^{n-k} \left(1 - \sum_{m=1}^{n-k} \frac{x^{-m} - 1}{m} \right)$$

in $[0, 1]$. $\mathcal{T}(\alpha) \subseteq D$, where D is defined in (9). The equality holds for $\alpha = 0$. It implies that $t_{n-k}(\alpha) \geq d_{n-k}$.

Let us assume that the DM does not know that the PR exists until he starts acting. He will pay the money if the observed chain of maximal elements falls into the set $\mathcal{T}(\alpha)$ but does not fall into the stopping set D earlier. The probability of that event is

$$p(\alpha) = \sum_{k=1}^{n} (1 - t_k(\alpha)) \prod_{l=1}^{k-1} (d_{n-l}) (1 - C_k(\alpha)) \tag{38}$$

where

$$C_k(\alpha) = \sum_{l=1}^{k} \sum_{i=1}^{l-1} \left(\prod_{j=1}^{i} \underline{x}_j(\alpha) \right)^{-1} \left[\frac{d_{n-i} \left(d_{n-l+1}^i - (d_{n-l+1} \wedge t_{n-k}(\alpha))^i \right)}{i} \right. $$
$$\left. - \frac{d_{n-l+1}^{i+1} - (d_{n-l+1} \wedge t_{n-k})^{i+1}}{i+1} \right].$$

The PR's expected payoff is

$$g(\alpha) = \alpha p(\alpha). \tag{39}$$

Theorem 2 *In the full-information best choice problem with a tribute the optimal strategy ρ^* for the prompter exists and*

$$\rho^* = \inf\{0 < k \le n : X_k = \max\{X_1, \ldots, X_k\}, \; X_k \ge t_{n-k}(\alpha^*)\} \qquad (40)$$

where

$$\alpha^* = \inf\{\alpha > 0 : g(\alpha) = \sup_a g(a)\}.$$

4.2 The limiting values

Let us analyze the properties of the payoff function for the PR as the number of observations tends to infinity. Suppose that $i \to \infty$ and write $x = 1 - \frac{t(i)}{i}$, $t(i) \in [0, i)$. We get

$$\lim_{i \to \infty} \varphi_{1,\alpha}(i, x) = \varphi_{1,\alpha}(t) = e^{-t} - \alpha \qquad (41)$$

$$\lim_{i \to \infty} \varphi_{2,\alpha}(i, x) = \varphi_{2,\alpha}(t) = e^{-t} \int_0^t \frac{e^u - 1}{u} du. \qquad (42)$$

Then, the price of the hint should satisfy the inequality

$$\alpha \le e^{-t}\left(1 - \int_0^t \frac{e^u - 1}{u} du\right), \quad 0 \le t \le z.$$

The threshold limit is

$$\lim_{i \to \infty} t_i(\alpha) = t_\alpha, \qquad (43)$$

where t_α is the unique solution of the equality

$$\alpha = e^{-t_\alpha}\left(1 - \int_0^{t_\alpha} \frac{e^u - 1}{u} du\right), \quad 0 \le t_\alpha \le z. \qquad (44)$$

The graph bellow presents the values of t_α as a function of parameter α (Fig. 2).

5 Conclusion

In the world around us, despite the widespread access to information, there are still cases where certain information is obscured and not accessible. Access to them can be extremely valuable. This is not always possible, but there may be a kind of special occasion to buy. In such cases, the profitability of the purchase and the decision should be seriously considered. As a result of these considerations, the above models were created. The aim of the work was to construct a mathematical model describing the mechanism of obtaining additional information in various market situations. Usually, such information is secret, and the possibility of obtaining it is difficult. Hence, in
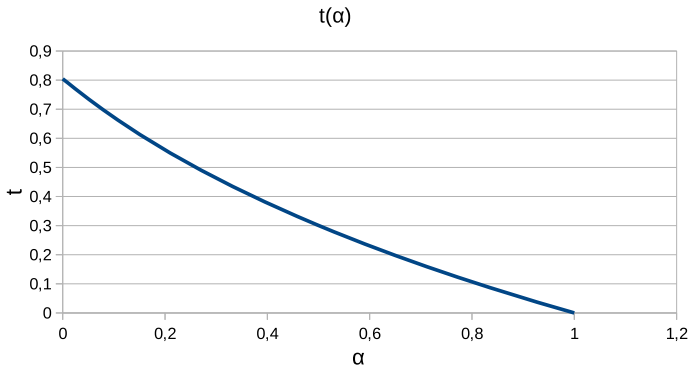
t(α)

Fig. 2 t_α as a function for different values of α

the model, there is one prompter that has exclusive information. The model can be expanded. One of the possibilities is to introduce more than one decision maker to the game. Then, the prompter decides which player will offer the most. Another possibility is the appearance of more people wanting to sell information. In any case, the knowledge about each other must be considered. In this model, we have found the formula for the optimal price for the hint. It has been shown that the value of the hint has its limits. In the second model, the prompter behaves more like a ripper and blocks the ability to stop. Also here, you can extend the game with additional players. In the game above we have found an equilibrium price and the optimal strategy for the prompter. The limit for the tribute as the number of the observations goes to infinity has been derived.

References

Bellman R (1957) Dynamic programming, 1st edn. Princeton University Press, Princeton

Bojdecki T (1978) On optimal stopping of a sequence of independent random variables probability maximizing approach. Stoch Process Appl 6(2):153–163. https://doi.org/10.1016/0304-4149(78)90057-1

Bullen P (2013) Handbook of means and their inequalities. Mathematics and its applications. Springer, Dordrecht. https://doi.org/10.1007/978-94-017-0399-4

Dotsenko SI, Marynych AV (2014) Hint, extortion, and guessing games in the best choice problem. Cybern Syst Anal 50(3):419–425. https://doi.org/10.1007/s10559-014-9630-8

Ferguson TS (1989) Who solved the secretary problem? Statist Sci 4(3):282–289. https://doi.org/10.1214/ss/1177012493

Gilbert JP, Mosteller F (1966) Recognizing the maximum of a sequence. J Am Stat Assoc 61(313):35–73

Gnedin AV (1996) On the full information best-choice problem. J Appl Probab 33(3):678–687

Gnedin AV, Miretskiy DI (2007) Winning rate in the full-information best-choice problem. J Appl Probab 44(2):560–565

Kuchta M (2017) Iterated full information secretary problem. Math Methods Oper Res 86(2):277–292. https://doi.org/10.1007/s00186-017-0594-0

Petruccelli JD (1982) Full-information best-choice problems with recall of observations and uncertainty of selection depending on the observation. Adv Appl Probab 14(2):340–358

Porosiński Z (1987) The full-information best choice problem with a random number of observations. Stoch Process Appl 24(2):293–307. https://doi.org/10.1016/0304-4149(87)90020-2

Porosinski Z (1992) The full-information best choice problem with two choices. In: Gritzmann P, Hettich R, Horst R, Sachs E (eds) Operations research '91: extended abstracts of the 16th symposium on operations research held at the University of Trier at September 9–11, 1991, pp 278–281. Physica-Verlag HD, Heidelberg. https://doi.org/10.1007/978-3-642-48417-9_77

Porosiński Z, Szajowski K (1996) On continuous-time two person full-information best choice problem with imperfect observation. Sankhyā Ser A 58(2):186–193

Sakaguchi M (1973) A note on the dowry problem. Rep Statist Appl Res Un Japan Sci Engrs 20(1):11–17

Sakaguchi M (1984) Best choice problems with full information and imperfect observation. Math Japon 29(2):241–250

Sakaguchi M, Szajowski K (1997) Single-level strategies for full-information best-choice problems. I Math Japon 45(3):483–495

Samuels SM (1982) Exact solutions for the full information best choice problem. Technical Report 82-17, Department of Statistics, Purdue University

Shiryayev AN (1978) Optimal stopping rules. Springer, New York

Tamaki M (1986) A full-information best-choice problem with finite memory. J Appl Probab 23(3):718–735