Optimal stopping in list-based decisions*

Saptarshi Mukherjee† Ruhi Sonal‡ Abhinaba Lahiri§

September 16, 2022

Abstract

We model a decision maker's choice from an exogenously given *list* of alternatives. In order to observe an alternative, the decision maker must observe all preceding. If observing alternatives is costly, then the decision maker may not observe all the alternatives in the list. We characterize an optimal stopping rule in this setting, the *stop-and-choose rule from lists (SCR-l)*. This rule provides an algorithm that describes when a rational decision maker will stop observing successive alternatives in a list, and determines which alternative will be chosen.

JEL classification: D81, D90

Keywords: Choice function, List, Satisficing, Stopping decision, Successive choice

1 Introduction

We consider a decision maker (DM) who encounters alternatives in an exogenously given *list*. In order to observe an alternative, the DM must observe all the preceding alternatives in the list. When observing alternatives is costly, the decision making problem is two-fold: when does the DM stop observing alternatives in the list and which alternative from the set that has been observed will the DM choose?

*Preliminary and incomplete.

†Department of Humanities and Social Sciences, IIT Delhi.

‡Department of Humanities and Social Sciences, IIT Jodhpur; Centre for Mathematical and Computational Economics, School of AI and Data Science, IIT Jodhpur.

§Centre for Mathematical and Computational Economics, School of AI and Data Science, IIT Jodhpur.

In this paper, we propose an optimal stop-and-choose rule in the above setting. We present the following examples of our setting:

1. *Netflix:* The popular OTT platform, Netflix categorizes the content on its platform: "popular on Netflix", "continue watching", "trending now" etc. are examples of such categories. The titles and thumbnails of each show within a category is displayed in a sequence. A viewer can view only a small number of thumbnails at a time and must scroll to the right to be able to observe the next available show in a category. The viewer decides when to stop scrolling and which show from the observed thumbnails to watch.

2. *Hiring a candidate:* Consider an HR manager who is interviewing candidates for a job. The candidates are interviewed in a pre-determined sequence. Conducting each interview is costly and the manager decides when to stop the search i.e., stop interviewing more candidates and which one of the interviewed candidates to select. Note that this problem is a modification of the classical *secretary problem*, which is a well known example of optimal search problems. In the secretary problem, the manager hires or rejects a candidate at the time of the interview. Therefore, the last candidate to be interviewed is hired. In our model, a candidate may be hired even if the manager continued the search for some time after interviewing the candidate.

Decision making from lists has previously been explored in Rubinstein and Salant (2006). Ishii et al. (2021) models choice from lists in a stochastic setting. Dimitrov et al. (2015) characterize a special class of choice rules: divide-and-choose rules from lists in which the decision maker may satisfice: an alternative is determined, after which the DM stops observing successive alternatives and chooses from the set observed. The paper closest to our approach is Weitzman (1979), in which a stop-and-choose rule is developed in a related setting: the decision maker chooses which alternatives to observe and when to stop observing more alternatives. In Weitzman (1979), the sequence in which alternatives are observed is determined by the DM. It is shown that the optimal strategy for a DM is to observe the alternative that has the highest "reservation price"[1] in the set of alternatives that have not been observed. The search stops when the reservation price for any alternative that has not been observed is lesser than the maximum payoff that an observed alternative yields. In our paper, the search problem faced by the DM is affected by the exogenous sequence in which the alternatives

[1] Weitzman (1979) interprets this as analogous to internal rate of return

appear: in order to observe an alternative that is at position k, the DM must observe the preceding $k-1$ alternatives. The stop-and-choose rule we model is as follows: the DM continues to observe the next alternative in the list if the maximum payoff from the set of observed alternatives is lower than the "reservation reward" of the next alternative. The reservation reward depends on the sequence i.e. it takes into account the $k-1$ alternatives that need to be observed before the DM can reach the kth alternative. If the maximum utility from an alternative in the set already observed exceeds the reservation reward of the next alternative, the DM stops the search and chooses the alternative that has the highest utility from the set that has been observed.

2 Model

Let A denote the set of all monetary rewards. For simplicity, assume that $A = \mathbb{R}_+$. The decision maker (DM) faces a predetermined sequence or list of $n \in \mathbb{N}$ boxes, each containing some potential monetary reward. Any such list is denoted by $X = (X_i)_{i=1}^n$. For any $n \in \mathbb{N}$, let $\mathbb{X}$ denote the set of all such lists of size n. For any $X \in \mathbb{X}$ and any $i \in \{1, 2, \dots, n\}$, X_i denote a random variable, whose realized values, denoted by x_i, are monetary rewards ($x_i \in A$). We assume that n is known by the DM. Given a list $X \in \mathbb{X}$, the DM believes that the potential rewards from any of the n boxes are distributed according to a n dimensional joint cumulative distribution function F, i.e., the DM believes that $X \sim F$. Here F is any n dimensional cumulative probability distribution function defined over A^n. Suppose $\mathbb{F}^n$ denote the set of all such n dimensional cumulative probability distribution function over A^n Given a list $X \in \mathbb{X}$, the DM, who wants to maximize his expected reward, has to determine which boxes to open and which reward to choose among the opened boxes. We assume that the DM cannot skip a box, i.e., given a list $X \in \mathbb{X}$, and any $i \in \{2, 3, \dots, n\}$, if the DM wants to open box i, then he has to open all previous boxes X_j, where $j \in \{1, 2, \dots, i-1\}$. Opening a box is costly for the DM. We denote by c_i the cumulative cost of opening boxes $X_1, X_2, \dots X_i$. For any $i < j$; $0 < c_i < c_j$. The DM may decide not to open any boxes. In such a case, he earns a default monetary reward $w \in A$.

2.1 Belief modification

For any list $X \in \mathbb{X}$, opening any box X_i is costly and in this model the cost accumulates as the DM moves through the list. As such, we assume that the DM uses F as a prior belief about the list and updates it based on observed rewards. In order to define this modification procedure, we introduce the following notations:

- Marginal Distribution: Given $X \in \mathbb{X}$ and the DM's belief that $X \sim F$, one can construct the DM's marginal beliefs as follows. For any $k \in \{1, 2, \dots, n\}$, and any $\{i_1, i_2, \dots, i_k\} \subseteq \{1, 2, \dots, n\}$, the marginal cumulative distribution of $(X(i_1), X(i_2), \dots, X(i_k))$ is defined as

$$F_{\{i_1,i_2,\dots,i_k\}}(.) = \int_{x(j)\in A \text{ for all } j\notin\{i_1,i_2,\dots,i_k\}} dF$$

  For example, given $X \sim F$, the DM's marginal belief about X_1, denoted by F_1, is defined as

$$F_1(x_1) = \int_{x_2\in A} \int_{x_3\in A} \dots \int_{x_n\in A} dF$$

- Conditional Beliefs: Given a list $X \in \mathbb{X}$ and an $i \in \{1, 2, \dots, n-1\}$, suppose the DM has opened all the boxes X_j, where $j \in \{1, 2, \dots, i\}$. Let the observed reward vector be $(x_1, x_2, \dots, x_i)$. Let the remaining list be denoted by X^i. Note that $X^i = (X_j)_{j=i+1}^n$. Now suppose that the DM believes that conditioned on observing $(x_1, x_2, \dots, x_i)$, $X^i \sim G_{(x_1,x_2,\dots,x_i)}$. Note that $G_{(x_1,x_2,\dots,x_i)} \in \mathbb{F}^{n-i}$ for any $(x_1, x_2, \dots, x_i) \in A^i$.

- Expected returns: Given a list $X \in \mathbb{X}$ and the DM's belief $X \sim F$, the expected return from opening box X_1 is

$$E(X_1) = \int_{x_1\in A} \int_{x_2\in A} \dots \int_{x_n\in A} x_1 dF = \int_{x_1\in A} x_1 dF_1 \in A$$

  As we assume that the DM cannot skip a box, so we assume that to calculate expected return from any box X_i, where $i \in \{2, 3, \dots, n\}$, he takes into account the observed reward vector $(x_1, x_2, \dots, x_{i-1})$. As such the expected reward from opening box X_i for any $i \in \{2, 3, \dots, n\}$ is

$$E(X_i|x_1, x_2, \dots, x_{i-1}) = \int_{x_i\in A} \int_{x_{i+1}\in A} \dots \int_{x_n\in A} x_i dG_{(x_1,x_2,\dots,x_i)}$$

Next, we introduce possible modification of the belief. We consider the following three methods of belief modification.

Independence : Here, DM disregards the observed reward vector, and assumes for any $i \in \{2, 3, \ldots, n\}$

$$G_{(x_1,x_2,\ldots,x_i)}(x_{i+1}, x_{i+2}, \ldots, x_n) = F_{\{i+1,i+2,\ldots,n\}}(x_{i+1}, x_{i+2}, \ldots, x_n)$$

In other words, under this modification scheme, given the list $X \in \mathbb{X}$, DM believes that each box X_i yields potential reward x_i following the cumulative distribution function F_i independent of any other box X_j.

Optimistic modification : Here the decision maker compares the observed reward with the expected reward as follows.

For any $i \in \{2, 3, \ldots, n\}$, if $x_1 \geq E(X_1), x_2 \geq E(X_2|x_1), \ldots, x_{i-1} \geq E(X_{i-1}|x_1, x_2, \ldots, x_{i-2})$, then the DM's belief about X^i conditioned on observing $(x_1, x_2, \ldots, x_{i-1})$, denoted by $G_{(x_1,x_2,\ldots,x_{i-1})}$, satisfy the following condition:

$$\int_{(x_i,x_{i+1},\ldots,x_n)\in A^{n-(i-1)}} x_i dG_{(x_1,x_2,\ldots,x_{i-1})} \geq \int_{(x_i,x_{i+1},\ldots,x_n)\in A^{n-(i-1)}} x_i dF^B_{(x_1,x_2,\ldots,x_{i-1})}$$

where $F^B_{(x_1,x_2,\ldots,x_{i-1})}(x_i, x_{i+1}, \ldots, x_n) = \frac{F(x_1,x_2,\ldots,x_n)}{F_{1,2,\ldots,i-1}(x_1,x_2,\ldots,x_{i-1})}$ is the usual conditional cumulative distribution function derived from F.

Otherwise, the DM assumes

$$G_{(x_1,x_2,\ldots,x_{i-1})}(x_i, x_{i+1}, \ldots, x_n) = F^B_{(x_1,x_2,\ldots,x_{i-1})}(x_i, x_{i+1}, \ldots, x_n)$$

In other words, under optimistic modification, if the DM observes that among the opened boxes, his realized returns are at least as good as the returns expected under his belief, then he can believe any distribution for the unobserved part of the list that results in a weakly higher expected return for the immediate next box as compared to the expected return calculated from his original belief. But if the DM does not observes that among the opened boxes, his realized returns are at least as good as the returns expected under his belief, then he sticks with his original belief.

Pessimistic modification : Here the decision maker compares the observed reward with the expected reward as follows.

For any $i \in \{2, 3, \ldots, n\}$, if $x_1 \leq E(X_1), x_2 \leq E(X_2|x_1), \ldots, x_{i-1} \leq E(X_{i-1}|x_1, x_2, \ldots, x_{i-2})$, then the DM's belief about X^i conditioned on observing $(x_1, x_2, \ldots, x_{i-1})$, denoted by $G_{(x_1,x_2,\ldots,x_{i-1})}$, satisfy the following condition:

$$\int_{(x_i,x_{i+1},\ldots,x_n)\in A^{n-(i-1)}} x_i dG_{(x_1,x_2,\ldots,x_{i-1})} \leq \int_{(x_i,x_{i+1},\ldots,x_n)\in A^{n-(i-1)}} x_i dF^B_{(x_1,x_2,\ldots,x_{i-1})}$$

Otherwise, the DM assumes

$$G_{(x_1,x_2,\ldots,x_{i-1})}(x_i,x_{i+1},\ldots,x_n)=F^B_{(x_1,x_2,\ldots,x_{i-1})}(x_i,x_{i+1},\ldots,x_n)$$

In other words, under pessimistic modification, if the DM observes that among the opened boxes, his realized returns are at most as good as the returns expected under his belief, then he can believe any distribution for the unobserved part of the list that results in a weakly lower expected return for the immediate next box as compared to the expected return calculated from his original belief. But if the DM does not observes that among the opened boxes, his realized returns are at most as good as the returns expected under his belief, then he sticks with his original belief.

The map $C : \mathbb{X} \longrightarrow A$ is a choice rule from lists. We denote the reservation reward associated with observing the ith box by v_i: if the DM is paid v_i, he is indifferent between observing or not observing X_i. Next we propose a class of choice rules as follows

Definition 1 *A choice rule from lists $C(\cdot)$ is a stop-and-choose rule from lists (SCRL) if for any $X \in \mathbb{X}$:*

(i) For any $i \in \{1,2,\ldots,n\}$, if $\max_{x_k\in A;k<i} -c_{i-1}+x_k \leq v_i$ then observe x_i. If not, stop.

(ii) If the decision maker has stopped after observing x_i, for some $i \leq n$, then $C(X)=\arg\max_{x_k\in A;k\leq i} -c_i+x_k$.

According to $SCRL$, the DM observes the next alternative in a list if the ex-ante expected utility from observing the alternative is higher than the maximum utility that the DM can obtain from the set of alternatives that have already been observed. Due to the list structure, the ex-ante expected utility from observing an alternative includes the opportunity for observing successive alternatives- if the DM wants to observe the kth alternative in a list, he must observe all the preceding $k-1$ alternatives. Once the DM stops, he chooses the utility maximizing alternative from the set observed.

We state and prove our main result: the characterization of $SCRL$.

3 Result

Theorem 2 *Assume that the DM follows independent belief modification. Then any choice rule C is an SCRL if and only if*

$$v_i = \max \begin{Bmatrix} -c_i + v_i \int_{-\infty}^{v_i} dF_i + \int_{v_i}^{\infty} x_i dF_i, \\ \\ -c_{i+1} + v_i \int_{-\infty}^{v_i} dF_{i+1} + \int_{v_i}^{\infty} x_{i+1} dF_{i+1}, \\ \\ \vdots \\ \\ -c_n + v_i \int_{-\infty}^{v_i} dF_n + \int_{v_i}^{\infty} x_n dF_n \end{Bmatrix}$$

Proof. We first show sufficiency. Consider a DM observing a list $X \in \mathbb{X}$. Suppose the reservation utility of the ith alternative in X is v_i, where

$$v_i = \max \begin{Bmatrix} -c_i + v_i \int_{-\infty}^{v_i} dF_i + \int_{v_i}^{\infty} x_i dF_i, \\ \\ -c_{i+1} + v_i \int_{-\infty}^{v_i} dF_{i+1} + \int_{v_i}^{\infty} x_{i+1} dF_{i+1}, \\ \\ \vdots \\ \\ -c_n + v_i \int_{-\infty}^{v_i} dF_n + \int_{v_i}^{\infty} x_n dF_n \end{Bmatrix}$$

Note that, irrespective of the definition of v_i, if the DM has observed i boxes, then he will choose the maximum reward from the opened boxes. So, if the decision maker has stopped after observing x_i, for some $i \leq n$, then $C(X) = \arg\max_{x_k \in A; k \leq i} -c_i + x_k$. Next, we need to show that for any $i \in \{1, 2, \ldots, n\}$, under the above definition of v_i, the DM will observe X_i if $\max_{x_k \in A; k<i} -c_{i-1} + x_k \leq v_i$. Suppose for contradiction that the DM observes X_i, but $\max_{x_k \in A; k<i} -c_{i-1} + x_k > v_i$. Note that v_i is the minimum utility that the DM requires in order to be indifferent to observing X_i. Note that $\max_{x_k \in A; k<i} -c_{i-1} + x_k$ denotes the maximum net reward the DM can achieve if the DM observed all the boxes up to $i-1$. Then $\max_{x_k \in A; k<i} -c_{i-1} + x_k > v_i$ shows that by observing box i, the DM is losing some reward, which contradicts reward maximizing behavior of the DM and concludes the proof of the sufficiency part, i.e, for the above definition of v_i, any choice rule must be a $SCRL$.

We now show necessity. For any $X \in \mathbb{X}$, let $C(X)$ be an $SCRL$. Suppose the DM stops after observing the $(i-1)$th alternative. By the definition of $SCRL$,

$$\max_{x_j \in A; j \leq (i-1)} -c_j + x_j > v_i$$

where v_i is the reservation utility for observing X_i. The maximum expected reward from $\{X_i, X_{i+1}, \ldots, X_n\}$ is

$$\max\left\{-c_i + \int_{-\infty}^{\infty} x_i dF_i, -c_{i+1} + \int_{-\infty}^{\infty} x_{i+1} dF_{i+1}, \ldots, -c_n + \int_{-\infty}^{\infty} x_n dF_n\right\}$$

v_i is set as the minimum utility that the DM requires in order to be indifferent to observing alternatives on wards from i. Since the DM cannot skip alternatives, v_i must contain information of the maximum expected utility that the DM can obtain by not stopping at $i-1$. Therefore, v_i is the solution of the dynamic programming problem (Weitzman (1979)):

$$v_i = \max\left\{-c_i + \int_{-\infty}^{\infty} x_i dF_i, -c_{i+1} + \int_{-\infty}^{\infty} x_{i+1} dF_{i+1}, \ldots, -c_n + \int_{-\infty}^{\infty} x_n dF_n\right\}$$

Note that in $[0, v_i]$ the DM is assured v_i. This implies that

$$v_i = \max\left\{\begin{array}{c} -c_i + v_i \int_{-\infty}^{v_i} dF_i + \int_{v_i}^{\infty} x_i dF_i, \\ \\ -c_{i+1} + v_i \int_{-\infty}^{v_i} dF_{i+1} + \int_{v_i}^{\infty} x_{i+1} dF_{i+1}, \\ \\ \vdots \\ \\ -c_n + v_i \int_{-\infty}^{v_i} dF_n + \int_{v_i}^{\infty} x_n dF_n \end{array}\right\}$$

■

Theorem 2 shows that the reservation utility characterizes $SCRL$. When the DM encounters alternatives in a list, he compares the expected utility from observing each subsequent alternative with the maximum utility he can obtain from the set of alternatives that have already been observed. The formulation of reservation utility takes into account the list structure: for each position in the list considers not only the opportunity of observing the alternative at that position, but also the opportunity of observing all successive alternatives. The characterization under different methods of belief modification is still in progress.

References

Dimitrov, D., S. Mukherjee, and N. Muto (2015): "'Divide-and-choose' in list-based decision problems," *Theory Decis*, 81, 17–31.

Ishii, Y., M. Kovach, and L. Ülkü (2021): "A model of stochastic choice from lists," *Journal of Mathematical Economics*, 96.

Rubinstein, A. and Y. Salant (2006): "A model of choice from lists," *Theoretical Economics*, 1, 3–17.

Weitzman, M. L. (1979): "Optimal search for the best alternative," *Econometrica*, 47.