Estimated Prophet: The Prophet Inequality in Fair, Sample, and Multi-Selection Settings

Quinn McIntyre, Ezra Steinberg, Peter Westbrook, Ethan Zhang

August 2023

Abstract

The prophet inequality is a classic model of online decision-making that has been extensively studied for nearly half a century. With the aim of making more realistic assumptions, many variations of this model have been explored. Some recent work has focused on generalizing the prophet inequality to situations where multiple selections need to be made, or some explicit notion of fairness needs to be satisfied, or full knowledge of the distributions is unavailable. In this paper, after presenting several classic proofs of the prophet inequality, we extend these recent advancements in several ways.

Arsenis and Kleinberg [8] introduced the notions of identity-independent fairness (IIF) and time-independent fairness (TIF), and showed that it is possible to construct an algorithm satisfying these properties that is still optimally competitive. In this paper, we make explicit their algorithm, and discuss a linear program that may help generalize their result to the $k > 1$ selections setting.

We also show that any offline IIF algorithm is at best $\frac{25}{27}$-competitive, and prove that any offline IIF algorithm which is only allowed to select the maximum of the realized variables is at best $\frac{1}{2}$-competitive.

Arsenis and Kleinberg also present a $\frac{1}{9}$-competitive IIF and TIF algorithm that makes decisions based on access to only two samples from each distribution. We analyze their algorithm more precisely, and demonstrate that it is in fact $\frac{1}{6}$-competitive.

Finally, we resolve a recent conjecture of Pashkovich and Sayutina [10], and demonstrate that in the setting of making $k > 1$ selections with access to only a single sample from each distribution, the algorithm that sets the k^{th} largest sample as a threshold is $1/2$-competitive.

Keywords: Online algorithms, prophet inequalities, fairness, k-select, single sample, linear programming, contention resolution

Contents

1 Introduction

Suppose that you're an employer and you're looking to hire someone for a position. You have a list of candidates, each of whom sent in an application, so you have a rough sense of their abilities and how they'll perform on the job. However, you'll gain a lot more information when you're able to interview each candidate.

In this setup, after each interview, you have to make an irreversible decision: to hire or to pass. You can only choose one candidate for the position, so once you hire someone the process is over. Additionally, once you pass on a candidate, there's no opportunity to revoke your decision after further interviews. What is your strategy? How does the uncertainty of future candidates impact each decision that you make?

This is an example of an online decision-making problem, where you don't start with all of the information at once. Instead, you gain additional information after every step of the process. By contrast, an offline decision-making problem is one in which you begin with all the information necessary to make a decision.

One way to model the situation mathematically is as follows. We have a fixed number of independent positive random variables $X_1, X_2, \ldots, X_n$, and we know beforehand all of their distributions and the order in which they appear. One by one, we observe the realization of the random variable and decide whether we accept or reject this value. Similar to the hiring problem from earlier, both decisions are irreversible, and an acceptance ends the process. Our goal is to create an algorithm that maximizes the expected value of our pick.

A reasonable way to measure the quality of such an algorithm is to take the ratio of the expected value of our algorithm and the expected value of a *prophet*, who is able to see all of the realized values and choose the largest one. Essentially, we're comparing how our pick does on average to the true maximum of the realized values (also averaged out over all possible realizations). The *prophet inequality*, introduced by Krengel and Sucheston [1], states that for any random variables $X_1, \ldots, X_n$, we can create an algorithm such that

$$\frac{\mathbb{E}[\text{ALG}]}{\mathbb{E}[\text{PROPHET}]} \geq \frac{1}{2}.$$

In this paper, we begin with modern overviews of several different proofs of the prophet inequality. We then focus on two variations of the prophet inequality.

The first variation explores the notion of fairness in the context of prophet inequalities. Arsenis and Kleinberg [8] introduced two ideas of fairness for these decision-making algorithms: identity-independent fairness (IIF) and time-independent fairness (TIF). Identity-independent fairness ensures that two candidates will have the same chance of being selected given they have the same value. Time-independent fairness ensures that a candidate's chance of being selected is independent of the time at which they arrive. In this paper, we analyze various IIF and TIF algorithms, in both the online and offline settings, and evaluate how well they perform.

The second variation explores making k selections; the goal is now to choose a fixed number k of the random variables, rather than just one, and to maximize their sum. We investigate this in the context of having access to only a single sample from each distribution, and resolve a recent conjecture of Pashkovich and Sayutina [10].

1.1 Our results

To formally state our results, we start by defining two notions of fairness, IIF (identity-independent fairness) and TIF (time-independent fairness), the former which is defined for an algorithm ALG and the latter a family of algorithms $\{ALG^\pi\}_{\pi \in S_n}$:

Definition 1 (IIF [8])**.** *An algorithm ALG for a given instance* $\mathcal{I} = ((X_i)_{i=1}^n, \pi)$ *supported on a set* S *is said to satisfy identity-independent fairness if there exists a function* $p : S \to [0,1]$ *such that:*

$$Pr[ALG \text{ hires } i | X_i = x] = p(x), \forall i \in [n], x \in S, \pi \in S_n.$$

We note that this definition works for both the online and offline settings. In the offline setting, we know the realizations of all the random variables from the beginning and must select one.

Definition 2 (TIF [8])**.** *A family of algorithms* $\{ALG^\pi\}_{\pi \in S_n}$*, one for each arrival ordering* π*, for an instance* $\mathcal{I} = (X_i)_{i=1}^n$ *supported on a set* S *is said to satisfy time-independent fairness if there exists a function* $p : [n] \times S \to [0,1]$ *such that:*

$$Pr[ALG^\pi \text{ hires } i | X_i = x] = p(i, x), \forall i \in [n], x \in S.$$

Arsenis and Kleinberg [8] prove via linear programming that there exists a $\frac{1}{2}$-competitive algorithm that is both IIF and TIF. In our paper, we discuss an explicit version of their algorithm that achieves this $\frac{1}{2}$ competitive ratio:

Algorithm 1: Explicit IIF and TIF Algorithm

Parameters: $\mathcal{I} = ((X_i)_{i=1}^n, \pi)$
for *t=1, . . . , n* **do**
  $i \leftarrow \pi(t)$.
  Inspect X_i and let $x \leftarrow X_i$.
  $Q_t \leftarrow 1 - \sum_{k=1}^{t-1} \sum_{y \in \mathcal{S}} f_{\pi(k)}(y) \frac{\Pr[X_{max}=y]}{2 \sum_{i=1}^n \Pr[X_i=y]}$
  Flip a coin with Heads probability equal to $q_t(x) = \frac{\Pr[X_{max}=x]}{2Q_t \sum_{i=1}^n \Pr[X_i=x]}$
  if *coin comes up Heads* **then**
    Hire i and halt.
  else
    Reject i and proceed.
  end
end

Theorem 1. *Algorithm 1 is an explicit IIF and TIF online algorithm that is* $\frac{1}{2}$*-competitive.*

Additionally, we also explore the possibility of algorithms making $k > 1$ selections while being IIF and TIF. Finding an algorithm that maintains fairness whilst competitively selecting k variables is challenging, so we find a linear program such that a solution to the linear program gives an online implementable algorithm. We conjecture that we can use this linear program to develop an IIF and TIF algorithm for making k selections.

The next two theorems discuss how the IIF condition affects algorithms in the offline setting. Arsenis and Kleinberg [8] showed that $\mathbb{E}$[OPTIMAL OFFLINE IIF ALG] $\geq \frac{1}{2}\mathbb{E}$[PROPHET] for any random variables, but they do not have a tight example and leave as an open question whether this constant of $\frac{1}{2}$ can be improved. Theorems 2 and 3 attempt to make progress on that open question.

Theorem 2. *We can find random variables such that, for any offline IIF algorithm, we have*

$$\mathbb{E}[ALG] \leq \frac{25}{27}\mathbb{E}[PROPHET]$$

Theorem 3. *When we consider an offline IIF algorithm, along with the condition that it can only either accept the maximum or not make a decision at all, it can be at best $\frac{1}{2}$-competitive.*

In their paper which introduced the notions of IIF and TIF, Aresnis and Kleinberg [8] provide a $\frac{1}{2}$-competitive offline IIF algorithm. Their algorithm was restricted to either accepting the maximum or not making a decision at all, which means the bound presented in Theorem 4 is tight. Their algorithm uses a sampling approach, which does not actually need full information about the distributions, but instead only needs to obtain a single sample from each random variable.

An algorithm that relies on one sample of realizations is called a *single sample* algorithm, while an algorithm that relies on two sample of realizations is called a *double sample* algorithm. Arsenis and Kleinberg [8] also presented the following online, double-sample algorithm which is both IIF and TIF:

Algorithm 2: Double-sample Online IIF and TIF Algorithm

Data: $Y_i, Z_i \sim f_i, \pi \in S_n$
for *t=1, . . . , n* **do**
    Observe $X_{\pi(t)} \sim f_{\pi(t)}$.
    if $X_{\pi(t)} > Y_{\max}$ *and* ($X_{\pi(s)} < Y_{\max}$ *for all* $s < t$) *and* ($Z_{\pi(s)} < Y_{\max}$ *for all* $s \geq t$) **then**
        Hire $X_{\pi(t)}$.
    end
end

In their paper, this algorithm was proven to be $\frac{1}{9}$-competitive, worse than the $\frac{1}{2}$ resulting from placing either the online or sampled based constraints alone on the algorithm. However, this bound is not tight, and with the following result

Theorem 4. *Algorithm 2 is IIF and TIF. Moreover,*

$$\mathbb{E}[ALG] \geq \frac{1}{6}\mathbb{E}[X_{\max}]$$

we improve the bound to $\frac{1}{6}$. In Arsenis and Kleinberg's work, all definitions and algorithms were constructed assuming the random variables are discrete. In our proof, it will be convenient for us to extend these definitions to the continuous case, where we introduce the following equivalent definitions:

Definition 3 (Continuous IIF)**.** *For each i, let $S_i = X_i \cdot \mathbb{1}(ALG \text{ accepts } X_i)$. Then an algorithm is IIF if and only if S_i is continuous and $\frac{s_i(x)}{f_i(x)}$ is independent of i for each $i = 1, \ldots, n$, where s_i is the pdf of S_i.*

Definition 4 (Continuous TIF)**.** *Define $\{ALG^\pi\}_{\pi \in S_n}$ as before. For each i and $ALG^\pi \in \{ALG^\pi\}_{\pi \in S_n}$, let $S_{i,\pi} = X_i \cdot \mathbb{1}(ALG^\pi \text{ accepts } X_i)$. Then $\{ALG^\pi\}_{\pi \in S_n}$ is TIF if and only if $A_{i,\pi}(x)$ is differentiable and $\frac{s_{i,\pi}(x)}{f_i(x)}$ is independent of π for each $\pi \in S_n$ and $i = 1, \ldots, n$, where $s_{i,\pi}(x)$ is the pdf of $S_{i,\pi}$.*

We claim that in the continuous case, algorithm 2 remains IIF and TIF, and satisfies the $\frac{1}{6}$ bound.

Lastly, we explore what happens with $k \geq 2$ selections, without any regard for fairness, given that we have access to a single sample from each distribution. Define $X^1 = X_{\max}$, $X^k = \max(\{X_1, \ldots, X_n\} - \{X^1, \ldots, X^{k-1}\})$. Define Y^i analogously. We study the algorithm which sets the k^{th} largest sample as a threshold, and show that it is 1/2-competitive against a prophet who always picks the k largest X-values, defining X-values as realizations of the random variables $X_1 \ldots X_n$:

Algorithm 3: Single-Sample k-select Algorithm

```
Data: Y_i ~ X_i
j ← 0
for t=1, ..., n do
    if X_t > Y^k and j < k then
        Hire X_t.
        j ← j + 1
    end
end
```

Theorem 5. *For any k, Algorithm 3, for continuous random variables X_i, is $\frac{1}{2}$-competitive, that is,*

$$\mathbb{E}[ALG] \geq \frac{1}{2}\mathbb{E}[\sum_{i=1}^{k} X^k]$$

Furthermore, this lower bound is tight for all k.

Theorem 5 had previously been known for $k = 1$ [7]. Pashkovich and Sayutina very recently proved this result for $k = 2$ [10], and conjectured that it was true for higher values of k[1]. Theorem 5 resolves their conjecture. Note that setting the k^{th} largest sample as a threshold is not necessarily the best choice, and other algorithms for the problem perform better for large values of k [5], but the simplicity of this algorithm makes it interesting to analyze.

1.2 Our techniques

Our linear program that corresponds to online implementable algorithms making k selections is inspired by [8] and [11].

[1] [10] implicitly assumes that the random variables are continuous, because they assume that various random variables don't equal each other. Without this assumption, we can also recover the result using randomized tie-breaking whenever two variables have the same realization. However, Pashkovich and Sayutina are interested specifically in deterministic algorithms, so the randomized tie-breaking strategy is unavailable to them.

To prove theorems 2 and 3, we systematically searched for and optimized possible adversarial examples.

In algorithm 2, all comparisons are strict. Thus, although the algorithm was originally defined for discrete random variables, Arsenis and Kleinberg assigned to each random variable a random number uniformly chosen from $[0, 1]$, which is the value compared in the event of a tie [8]. To make their discussion more formal, we use continuous random variables. In order to improve the bound from the paper and prove theorem 4, we explicitly compute the expected value of the algorithm in terms of the probability density functions of the random variables.

To prove Theorem 5, we first transform the algorithm into a uniformly worse-performing one that always selects the k lowest X-values above the threshold Y^k. We then condition on the combined, sorted list of X and Y-values $\{a_1, \cdots, a_{2n}\}$, and we express both the algorithm's and prophet's performance in terms of the expected number of picks each makes above a_j for each j. The remainder of the proof involves casework for small values of j and k, and simple lower bounds for high values of j and k. This is largely done by considering the probability that each X-value is above a_j and is selected by the algorithm/prophet and bounding this using a key lemma due to Nuti and Vondrák [12].

2 A Few Proofs of the Prophet Inequality

2.1 A proof based on dynamic programming

In this section, we present one of the very first proofs of the prophet inequality, a classic proof due to ideas of A. Dvoretzky, published by Krengel and Sucheston [1]. Consider a strategy of "optimal play" that at any decision point makes the choice with the highest expected value conditioned on our current information. Let Φ_i be the expected return from optimal play on the random variables $X_i, \ldots, X_n$. Then $\Phi_{n+1} = 0$, and $\Phi_i = \mathbb{E}[\max\{X_i, \Phi_{i+1}\}]$. This is because Φ_{n+1} is the result of optimal play with 0 variables, and optimal play at variable i obtains the maximum of the current realization and the expected value of optimal play on the remaining variables (if we reject the current one).

The original inductive version of the proof goes as follows. Let $S_i = \max\{X_i, \ldots, X_n\}$. The proof shows that

$$\Phi_i \geq \mathbb{E}[(S_i - \Phi_i)_+]$$

by (reverse) induction. Clearly the statement is true for $i = n + 1$. If we assume the statement holds for k, then

$$\begin{aligned}
\Phi_{k-1} &= \mathbb{E}[\max(X_{k-1}, \Phi_k)] \\
&= \Phi_k + \mathbb{E}[(X_{k-1} - \Phi_k)_+] \\
&\geq \mathbb{E}[(S_k - \Phi_k)_+] + \mathbb{E}[(X_{k-1} - \Phi_k)_+] \\
&\geq \mathbb{E}[(S_{k-1} - \Phi_k)_+] \\
&\geq \mathbb{E}[(S_{k-1} - \Phi_{k-1})_+]
\end{aligned}$$

establishing the inequality for $k - 1$. This implies the inequality

$$\Phi_1 \geq \mathbb{E}[(S_1 - \Phi_1)_+] \implies \Phi_1 \geq \mathbb{E}[S_1] - \Phi_1$$

so we can conclude that $\Phi_1 \geq \frac{1}{2}\mathbb{E}[S_1]$, where $\mathbb{E}[S_1] = \mathbb{E}[X_{max}]$ represents the expected value of the prophet.

2.1.1 A variation: competing against the the top k picks of the prophet

DP Kennedy [3] generalizes the result to a situation where the prophet must select k random variables, and we must select one. If we define $M_i, 1 \leq i \leq n$, to be i^{th} order statistic, then Kennedy provides the following upper bound on the sum of the top k order statistics, that is, the prophet's choices:

$$\mathbb{E}[\sum_{i=1}^{k} M_i] \leq (k+1)\Phi_1$$

where the case of $k = 1$ represents the usual prophet inequality.

We will use backwards induction to show that defining $M_{i,j}$ to be the ith largest order statistic from X_1 to X_j (with M_i implicitly meaning $M_{i,n}$), then:

$$\mathbb{E}[\sum_{i=1}^{k} M_i] \leq \mathbb{E}[\sum_{i=1}^{k} \max(\Phi_{j+2}, M_{i,j}) + \Phi_{j+1}]$$

for $k \leq j \leq n$. Clearly, for $j = n$, the (in)equality is trivial since $\Phi_{n+1} = \Phi_{n+2} = 0$. We have moreover that

$$\begin{aligned}
\mathbb{E}[\sum_{i=1}^{k} \max(\Phi_{j+2}, M_{i,j}) + \Phi_{j+1}] &\leq \mathbb{E}[\sum_{i=1}^{k} \max(\Phi_{j+1}, M_{i,j}) + \Phi_{j+1}] \\
&= \mathbb{E}[\sum_{i=1}^{k-1} \max(\Phi_{j+1}, M_{i,j-1}) + \max(\Phi_{j+1}, M_{k,j-1}, X_j) + \Phi_{j+1}]
\end{aligned}$$

The final step there can be seen by splitting into cases: if X_j is not among the k highest of the first j realizations, then we can ignore it and we have the same sum with the $M_{i,j}$ replaced with $M_{i,j-1}$, which makes no difference for the highest k realizations. If X_j is among the k highest of the first j realizations, then $X_j > M_{k,j-1}$, and so both sums are $\max(\Phi_{j+1}, X_j)$ added to sum of $\max(\Phi_{j+1}, M_{i,j-1})$ for the highest $k-1$ realizations up to $j-1$.

We have also that

$$\max(\Phi_{j+1}, M_{k,j-1}, X_j) + \Phi_{j+1} \leq \max(\Phi_{j+1}, M_{k,j-1}) + \max(X_j, \Phi_{j+1})$$

since, in fact, $\max(a, b, c) + a \leq \max(a, b) + \max(c, a)$ for all real numbers a, b, c. Therefore,

$$\begin{aligned}
\mathbb{E}[\sum_{i=1}^{k} \max(\Phi_{j+2}, M_{i,j}) + \Phi_{j+1}] &\leq \mathbb{E}[\sum_{i=1}^{k-1} \max(\Phi_{j+1}, M_{i,j-1}) + \max(\Phi_{j+1}, M_{k,j-1}) + \max(X_j, \Phi_{j+1})] \\
&= \mathbb{E}[\sum_{i=1}^{k-1} \max(\Phi_{j+1}, M_{i,j-1}) + \max(\Phi_{j+1}, M_{k,j-1}) + \Phi_j] \\
&= \mathbb{E}[\sum_{i=1}^{k} \max(\Phi_{j+1}, M_{i,j-1}) + \Phi_j]
\end{aligned}$$

Therefore, the inequality

$$\mathbb{E}[\sum_{i=1}^{k} M_i] \leq \mathbb{E}[\sum_{i=1}^{k} \max(\Phi_{j+2}, M_{i,j}) + \Phi_{j+1}]$$

holds for $j = n$, and it holding for a given j between $k+1$ and n implies that it holds for $j-1$. Therefore, it holds for $j = k$, and so

$$\mathbb{E}[\sum_{i=1}^{k} M_i] \leq \mathbb{E}[\sum_{i=1}^{k} \max(\Phi_{k+2}, M_{i,k}) + \Phi_{k+1}] = \mathbb{E}[\sum_{i=1}^{k} \max(\Phi_{k+2}, X_i) + \Phi_{k+1}]$$

Since $\max(\Phi_{k+2}, X_i) \leq \max(X_i, \Phi_{i+1}) = \Phi_i$ for $i \leq k$, we have

$$\mathbb{E}[\sum_{i=1}^{k} M_i] \leq \mathbb{E}[\sum_{i=1}^{k} \Phi_i + \Phi_{k+1}] = \mathbb{E}[\sum_{i=1}^{k+1} \Phi_i] \leq (k+1)\Phi_1$$

which concludes the proof.

2.2 A proof based on induction on the number of random variables

In this section, we present another proof of the prophet inequality, which involves changing the random variables and reducing their number, while increasing the ratio of the prophet's pick to the algorithm's pick (thus putting us at a worse position than before). More specifically, we are going to perform the changes:

$$X_1, \ldots, X_n \to \Phi_2, X_2, \ldots, X_n$$

and

$$\Phi_2, X_2, \ldots, X_n \to \Phi_2, X_2, \ldots, X_{n-2}, L_p.$$

for some chancy variable L_p and deterministic Φ_2 (defined as in the DP proof). This proof is due to Hill and Kertz [2], and we aim here to present it in a more accessible and modern fashion.

In the first change, the prophet and algorithm's expected pick both worsen in an amount so that the overall ratio increases. In the second change, we will construct L_p so that the algorithm's expected pick remains the same and the prophet's expected pick increases, leading to an overall ratio increase. This reduces the number of random variables by one, simplifying the problem to the case where only two random variables remain.

Let $X_1, \ldots, X_n$ be the random variables in the prophet inequality. Let ALG denote the best algorithm to select one variable from the list and maximize the expected gain. Then the prophet inequality claims that for $n \geq 2$,

$$\mathbb{E}[\text{ALG}] \geq \frac{1}{2}\mathbb{E}[X_{\max}]$$

where $X_{\max}$ is the largest value realized amongst $X_1, , \ldots, , X_n$. We aim to show this by induction. First, assume $n = 2$. Then, note that

$$\mathbb{E}[\text{ALG}] \geq \max(\mathbb{E}[X_1], \mathbb{E}[X_2])$$

since one algorithm would be to choose the larger of $\mathbb{E}[X_1], \mathbb{E}[X_2]$, meaning the best algorithm achieves at least that much. Thus,

$$\begin{aligned} 2\mathbb{E}[\text{ALG}] &\geq 2\max(\mathbb{E}[X_1], \mathbb{E}[X_2]) \\ &\geq \mathbb{E}[X_1] + \mathbb{E}[X_2] \end{aligned}$$

$$\geq \mathbb{E}[X_{\max}]$$

and

$$\mathbb{E}[\text{ALG}] \geq \frac{1}{2}\mathbb{E}[X_{\max}]$$

completing the base case. Now, assume the prophet inequality is true for all m less than n, where $n \geq 3$ is a positive integer. We proceed with the inductive step.

For each $i = 1, \ldots, n$, let Φ_i be the expected return from optimal play on the random variables $X_i, \ldots, X_n$. Moreover, let $X_{\max[i,j]}$ denote the largest realization of X amongst $X_i, , \ldots, , X_j$ inclusive, where $i \leq j$ are integers in $\{1, , \ldots, , n\}$, and $(X)_+ = \max(X, 0)$ the positive part of a random variable X. Now, we claim that

$$\mathbb{E}[X_{\max}] \leq \mathbb{E}[\max(\Phi_2, X_{\max[2,n]})] + \mathbb{E}[(X_1 - \Phi_2)_+] \tag{1}$$

This follows from the fact

$$\begin{aligned}
\mathbb{E}[X_{\max}] &\leq \mathbb{E}[\max(\Phi_2, X_{\max})] \\
&= \mathbb{E}[\max(\Phi_2, X_{\max[2,n]}) + (X_1 - \max(\Phi_2, X_{\max[2,n]}))_+] \\
&= \mathbb{E}[\max(\Phi_2, X_{\max[2,n]})] + \mathbb{E}[(X_1 - \max(\Phi_2, X_{\max[2,n]}))_+] \\
&\leq \mathbb{E}[\max(\Phi_2, X_{\max[2,n]})] + \mathbb{E}[(X_1 - \Phi_2)_+]
\end{aligned}$$

Note that we used linearity of expectation here, in addition to the fact $\Phi_2 \leq \max(\Phi_2, X_{\max[2,n]})$. Next, we have the following simple proposition:

Proposition 1. *For all $a \geq b > 0$ and $\delta \geq 0$, we have $\frac{a+\delta}{b+\delta} \leq \frac{a}{b}$.*

Proof. Indeed, $ab + b\delta \leq ab + a\delta$, which can be rearranged into the desired inequality. □

Now, for random variables $Y_1, , \ldots, , Y_m$ let $\text{ALG}(Y_1, , \ldots, , Y_m)$ denote the best expected value of an algorithm. Then since $\mathbb{E}[\max(\Phi_2, X_{\max[2,n]}) \geq \text{ALG}(\Phi_2, X_2, \ldots, , X_n) > 0$ (because for any specific case, the true maximum is always at least what the algorithm gives, which is nonzero unless we are in the trivial case in which all the random variables are zero), and $\mathbb{E}[(X_1 - \Phi_2)_+] \geq 0$,

$$\frac{\mathbb{E}[\max(\Phi_2, X_{\max[2,n]})]}{\text{ALG}(\Phi_2, X_2, \ldots, , X_n)} \geq \frac{\mathbb{E}[\max(\Phi_2, X_{\max[2,n]})] + \mathbb{E}[(X_1 - \Phi_2)_+]}{\text{ALG}(\Phi_2, X_2, \ldots, , X_n) + \mathbb{E}[(X_1 - \Phi_2)_+]}$$

by the proposition. Then since $\Phi_i = \Phi_{i+1} + \mathbb{E}[(X_i - \Phi_{i+1})_+]$ as discussed in the DP proof, $\text{ALG}(\Phi_2, X_2, , \ldots, , X_n) = \text{ALG}(X_2, , \ldots, , X_n)$ and

$$\begin{aligned}
\frac{\mathbb{E}[\max(\Phi_2, X_{\max[2,n]})] + \mathbb{E}[(X_1 - \Phi_2)_+]}{\text{ALG}(\Phi_2, X_2, \ldots, , X_n) + \mathbb{E}[(X_1 - \Phi_2)_+]} &= \frac{\mathbb{E}[\max(\Phi_2, X_{\max[2,n]})] + \mathbb{E}[(X_1 - \Phi_2)_+]}{\text{ALG}(X_1, X_2, \ldots, , X_n)} \\
&= \frac{\mathbb{E}[\max(\Phi_2, X_{\max[2,n]})] + \mathbb{E}[(X_1 - \Phi_2)_+]}{\Phi_1}
\end{aligned}$$

and by (1) we have

$$\frac{\mathbb{E}[\max(\Phi_2, X_{\max[2,n]})] + \mathbb{E}[(X_1 - \Phi_2)_+]}{\Phi_1} \geq \frac{\mathbb{E}[X_{\max}]}{\Phi_1}$$

It follows that

$$\frac{\mathbb{E}[\max(\Phi_2, X_{\max[2,n]})]}{\mathrm{ALG}(\Phi_2, X_2, \ldots, , X_n)} \geq \frac{\mathbb{E}[X_{\max}]}{\Phi_1} \tag{2}$$

Now, consider a random variable L_p such that $\Pr(L_p = \Phi_{n-1}/p) = p > 0$, and $L_p = 0$ otherwise for some small p to be determined (such a variable with high probability of being zero and low probability of being a high value is called a *long shot.*) Now, it is easy to verify $\Phi_{n-1} = \mathbb{E}[L_p]$, which means that

$$\mathrm{ALG}(\Phi_2, X_2, .., X_{n-2}, L_p) = \mathrm{ALG}(\Phi_2, X_2, .., X_n) \tag{3}$$

Taking the limit from the right as $p \to 0$, note that

$$\mathbb{E}[\max(\Phi_2, X_{\max[2,n-2]}, L_p)] \to \mathbb{E}[\max(\Phi_2, X_{\max[2,n-2]}) + L_p] = \mathbb{E}[\max(\Phi_2, X_{\max[2,n-2]})] + \mathbb{E}[L_p]$$

Indeed,

$$\begin{aligned}\max(\Phi_2, X_{\max[2,n-2]}, L_p) - \max(\Phi_2, X_{\max[2,n-2]}) - L_p &= \max(\Phi_2, X_{\max[2,n-2]}) - \max(\Phi_2, X_{\max[2,n-2]}) \\ &= 0\end{aligned}$$

when $L_p = 0$ and

$$\begin{aligned}& \max(\Phi_2, X_{\max[2,n-2]}, L_p) - \max(\Phi_2, X_{\max[2,n-2]}) - L_p \\ = \quad & \max\left(\Phi_2, X_{\max[2,n-2]}, \frac{1}{p}\right) - \max(\Phi_2, X_{\max[2,n-2]}) - \frac{1}{p}\end{aligned}$$

when $L_p = \frac{1}{p}$, revealing

$$\begin{aligned}& \mathbb{E}[\max(\Phi_2, X_{\max[2,n-2]}, L_p) - \max(\Phi_2, X_{\max[2,n-2]}) - L_p]] \\ = \quad & (1-p) \cdot \mathbb{E}[0] + p \cdot \mathbb{E}\left[\max\left(\Phi_2, X_{\max[2,n-2]}, \frac{1}{p}\right) - \max(\Phi_2, X_{\max[2,n-2]}) - \frac{1}{p}\right] \\ = \quad & \mathbb{E}[\max(p\Phi_2, pX_{\max[2,n-2]}, 1) - 1] - p\mathbb{E}[\max(\Phi_2, X_{\max[2,n-2]})] \to 0\end{aligned}$$

as $p \to 0$, as needed. Now, from a simple calculation with the fact that $\Phi_i = \Phi_{i+1} + \mathbb{E}[(X_i - \Phi_{i+1})_+]$, note that

$$\mathbb{E}[\max(\Phi_2, X_{\max[2,n-2]})] + \mathbb{E}[L_p] = \mathbb{E}[\max(\Phi_2, X_{\max[2,n-2]})] + \mathbb{E}[X_n] + \mathbb{E}[(X_{n-1} - \mathbb{E}[X_n])_+]$$

As any algorithm for the random variables $X_2, , \ldots, , X_n$ can choose at least $\mathbb{E}[X_n]$ on expectation, it follows $\mathbb{E}[X_n] \leq \Phi_2$ and

$$\begin{aligned}& \mathbb{E}[\max(\Phi_2, X_{\max[2,n-2]})] + \mathbb{E}[X_n] + \mathbb{E}[(X_{n-1} - \mathbb{E}[X_n])_+] \\ \geq \quad & \mathbb{E}[\max(\Phi_2, X_{\max[2,n-2]})] + \mathbb{E}[X_n] + \mathbb{E}[(X_{n-1} - \max(\Phi_2, X_{\max[2,n]}))_+] \\ = \quad & \mathbb{E}[\max(\Phi_2, X_{\max[2,n-1]})] + \mathbb{E}[X_n] \\ = \quad & \mathbb{E}[\max(\Phi_2, X_{\max[2,n-1]}) + X_n] \\ > \quad & \mathbb{E}[\max(\Phi_2, X_{\max[2,n]})]\end{aligned}$$

Combined with inequality 2 and equation 3, we obtain for sufficiently small p that

$$\mathbb{E}[\max(\Phi_2, X_{\max[2,n-2]}, L_p)] \geq \mathbb{E}(\max(\Phi_2, X_{\max[2,n]})]$$

and

$$\frac{\mathbb{E}[\max(\Phi_2, X_{\max[2,n-2]}, L_p)]}{\text{ALG}(\Phi_2, X_2, , \ldots, , X_{n-2}, L_p)} = \frac{\mathbb{E}[\max(\Phi_2, X_{\max[2,n]})] + \mathbb{E}[(X_1 - \Phi_2)_+]}{\text{ALG}(X_1, X_2, \ldots, , X_n)}$$
$$\geq \frac{\mathbb{E}(\max(\Phi_2, X_{\max[2,n]})]}{\text{ALG}(\Phi_2, X_2, .., X_n)}$$
$$\geq \frac{\mathbb{E}[X_{\max}]}{\Phi_1}$$

for sufficiently small p. By the inductive hypothesis on the $n-1$ variables $\Phi_2, X_2, , \ldots, , X_{n-2}, L_p$, note that

$$2 \geq \frac{\mathbb{E}[\max(\Phi_2, X_{\max[2,n-2]}, L_p)]}{\text{ALG}(\Phi_2, X_2, , \ldots, , X_{n-2}, L_p)}$$

It follows

$$\Phi_1 = \mathbb{E}[\text{ALG}] \geq \frac{1}{2}\mathbb{E}[X_{\max}]$$

completing the inductive step. The prophet inequality for all $n \geq 2$ follows, as desired.

2.3 A proof based on the idea of an online contention resolution scheme

In this section, we present another proof ot the prophet inequality, closely related to the idea of a contention resolution scheme. For this method, we pick a separate threshold for each X_i and accept the first variable that beats its threshold. We'll show that, if the thresholds are chosen correctly, then the algorithm is $\frac{1}{2}$-competitive. This proof is adapted from the work of Alaei [4].

We let $p_i = \Pr[X_i = X_{\max}]$, so we have $\sum_i p_i = 1$. Additionally, we define each τ_i such that $\Pr[X_i \geq \tau_i] = p_i$.

We begin by proving an important inequality. In preparation, note that

$$\begin{aligned}\Pr[X_i = X_{\max} \text{ and } X_i < \tau_i] &= \Pr[X_i = X_{\max}] - \Pr[X_i = X_{\max} \text{ and } X_i \geq \tau_i] \\ &= \Pr[X_i \geq \tau_i] - \Pr[X_i = X_{\max} \text{ and } X_i \geq \tau_i] \\ &= \Pr[X_i < X_{\max} \text{ and } X_i \geq \tau_i]\end{aligned}$$

Next, for each X_i, we see that

$$\begin{aligned} & \mathbb{E}[X_i | X_i = X_{\max}] \\ = \;& \frac{1}{\Pr[X_i = X_{\max}]}\big(\Pr[X_i = X_{\max} \text{ and } X_i \geq \tau_i]\mathbb{E}[X_i | X_i = X_{\max} \text{ and } X_i \geq \tau_i] + \\ & \Pr[X_i = X_{\max} \text{ and } X_i < \tau_i]\mathbb{E}[X_i | X_i = X_{\max} \text{ and } X_i < \tau_i]\big) \\ \leq \;& \frac{1}{\Pr[X_i \geq \tau_i]}\big(\Pr[X_i = X_{\max} \text{ and } X_i \geq \tau_i]\mathbb{E}[X_i | X_i = X_{\max} \text{ and } X_i \geq \tau_i] + \\ & \Pr[X_i < X_{\max} \text{ and } X_i \geq \tau_i]\mathbb{E}[X_i | X_i < X_{\max} \text{ and } X_i \geq \tau_i]\big) \\ = \;& \mathbb{E}[X_i | X_i \geq \tau_i]\end{aligned}$$

We can calculate the expected value of the prophet's choice, which comes out to

$$\mathbb{E}[\text{PROPHET}] = \sum_{i=1}^{n} \Pr[X_i = X_{\max}]\mathbb{E}[X_i | X_i = X_{\max}] \leq \sum_{i=1}^{n} p_i \mathbb{E}[X_i | X_i \geq \tau_i]$$

Now we can get to the main idea behind the proof. We denote by r_i the probability that we reach element X_i in our decision making process, and we note that $r_1 = 1$ and $r_1 \geq r_2 \geq \cdots \geq r_n$. We define a new threshold Θ_i such that $\Pr[X_i \geq \Theta_i] = \frac{p_i}{2r_i}$, and we accept X_i if it beats this threshold. We're certain this is defined for Θ_1 because $r_1 = 1$. We'll use induction to prove that $r_i \geq \frac{1}{2}$ for $i > 1$, which will demonstrate why Θ_i is well-defined.

Assume that $r_i \geq \frac{1}{2}$ for $1 \leq i \leq k$. For each i, we have

$$r_{i+1} = r_i \left(1 - \frac{p_i}{2r_i}\right) = r_i - \frac{p_i}{2}$$

Summing these equations for $1 \leq i \leq k$, we get

$$r_{k+1} = r_1 - \sum_{i=1}^{k} \frac{p_i}{2} \geq r_1 - \frac{1}{2}\sum_{i=1}^{n} p_i = 1 - \frac{1}{2} = \frac{1}{2}$$

so we know that $r_{k+1} \geq \frac{1}{2}$, so we can define Θ_{k+1}. Thus, we can do this for $1 \leq i \leq n$.

Clearly, we must have $\Theta_i \geq \tau_i$ and $\mathbb{E}[X_i|X_i \geq \Theta_i] \geq \mathbb{E}[X_i|X_i \geq \tau_i]$. Now, we have

$$\begin{aligned}
\mathbb{E}[\text{ALG}] &= \sum_{i=1}^{n} r_i \Pr[X_i \geq \Theta_i]\mathbb{E}[X_i|X_i \geq \Theta_i] \\
&= \sum_{i=1}^{n} r_i \cdot \frac{p_i}{2r_i} \cdot \mathbb{E}[X_i|X_i \geq \Theta_i] \\
&= \frac{1}{2}\sum_{i=1}^{n} p_i \mathbb{E}[X_i|X_i \geq \Theta_i] \\
&\geq \frac{1}{2}\sum_{i=1}^{n} p_i \mathbb{E}[X_i|X_i \geq \tau_i] \geq \frac{1}{2}\mathbb{E}[\text{PROPHET}]
\end{aligned}$$

2.3.1 A variation: graph matching with edge arrivals

We now consider a variant of the prophet inequality, and use a similar approach based on the idea contention resolution to create an algorithm for this variant. This proof is adapted from the work of Ezra, Feldman, Gravin, and Tang [6].

Consider a graph, and we denote the edge connecting vertices u, v as (uv). Each edge has a positive random variable assigned to it, and we'll call that X_{uv}. In the edge arrival problem, the values of each edge are revealed to us one at a time, in a fixed order, and we make the decision whether to accept the edge (which can only be done if both vertices are currently free) or to move on. Our goal is to create an algorithm that maximizes the expected value of the sum of our selected edges. We will provide an algorithm that gives $\mathbb{E}[\text{ALG}] \geq \frac{1}{3}\mathbb{E}[\text{PROPHET}]$.

First, a quick bit of more notation. We say that $(u'v') < (uv)$ if the value of $X_{u'v'}$ is revealed to us before the value of X_{uv}.

We denote p_{uv} as the probability that (uv) is included in the optimal choice (i.e. the prophet's choice). Because each vertex can have at most one edge selected, for any fixed vertex v we know that $\sum_{u'} p_{u'v} \leq 1$. We can then define τ_{uv} such that $\Pr[X_{uv} \geq \tau_{uv}] = p_{uv}$, and using similar reasoning as the last section we can show that

$$\mathbb{E}[X_{uv}|(uv) \in \text{ optimal}] \leq \mathbb{E}[X_{uv}|X_{uv} \geq \tau_{uv}]$$

Now, we'll get into the algorithm. We denote r_{uv} as the probability that both vertices u and v are available when the value of X_{uv} is revealed to us, so this value is 1 for the first edge that is revealed. Additionally, we define Θ_{uv} such that $\Pr[X_{uv} \geq \Theta_{uv}] = \frac{p_{uv}}{3r_{uv}}$. The algorithm selects (uv) if

1. both u, v are available when we reach (uv) and
2. the realized value of X_{uv} is at least Θ_{uv}

so we can see that every edge is selected with probability $\frac{p_{uv}}{3}$.
We'll now show that Θ_{uv} can always be defined (i.e. $r_{uv} \geq \frac{1}{3} \forall u, v$). The reasoning is similar to the original proof; we see (inducting on the order in which edges arrive) that

$$\Pr[u \text{ is unavailable when we reach } (uv)] = \sum_{v' | uv' < uv} \Pr[(uv') \text{ is selected}] = \sum_{v' | uv' < uv} \frac{p_{uv'}}{3} \leq \frac{1}{3}$$

which means that

$$r_{uv} = \Pr[u, v \text{ both available when we reach } (uv)] \geq 1 - \Pr[u \text{ is unavailable}] - \Pr[v \text{ is unavailable}] \geq \frac{1}{3}$$

so Θ_{uv} can always be defined, and clearly $\Theta_{uv} \geq \tau_{uv}$ and $\mathbb{E}[X_{uv} | X_{uv} \geq \Theta_{uv}] \geq \mathbb{E}[X_{uv} | X_{uv} \geq \tau_{uv}]$. We find that

$$\mathbb{E}[\text{PROPHET}] = \sum_{u,v} \Pr[(uv) \in \text{optimal}] \mathbb{E}[X_{uv} | (uv) \in \text{optimal}] \leq \sum_{u,v} p_{uv} \mathbb{E}[X_{uv} | X_{uv} \geq \tau_{uv}]$$

and that

$$\begin{aligned} \mathbb{E}[\text{ALG}] &= \sum_{u,v} r_{uv} \Pr[X_{uv} \geq \Theta_{uv}] \mathbb{E}[X_{uv} | X_{uv} \geq \Theta_{uv}] \\ &= \sum_{u,v} \frac{p_{uv}}{3} \mathbb{E}[X_{uv} | X_{uv} \geq \Theta_{uv}] \\ &\geq \frac{1}{3} \sum_{u,v} p_{uv} \mathbb{E}[X_{uv} | X_{uv} \geq \tau_{uv}] \geq \frac{1}{3} \mathbb{E}[\text{PROPHET}] \end{aligned}$$

completing the proof.

3 An Explicit Algorithm for a Fair Prophet Inequality

Arsenis and Kleinberg [8] provide a proof that there exists an IIF 1 and TIF 2 algorithm that is $\frac{1}{2}$-competitive using the method of linear programming. Here, inspired by the quantities appearing in the contention resolution proof presented above, we provide an exposition with an explicit version of the algorithm given by Arsenis and Kleinberg [8], rather than using the output of a black-box linear program.

3.1 The explicit algorithm

To define the algorithm, let us write

$$r(x) := \frac{\Pr[X_{max} = x]}{2 \sum_{i=1}^{n} \Pr[X_i = x]},$$

and additionally, for convenience of notation, let us write

$$f_i(x) := \Pr[X_i = x].$$

Recall algorithm 1, equivalent to the algorithm considered in [8], but explicit:

Algorithm: Explicit IIF and TIF Algorithm

Parameters: $\mathcal{I} = ((X_i)_{i=1}^n, \pi)$

for *t=1, . . . , n* **do**

  $i \leftarrow \pi(t)$.

  Inspect X_i and let $x \leftarrow X_i$.

  $Q_t \leftarrow 1 - \sum_{k=1}^{t-1} \sum_{y \in \mathcal{S}} f_{\pi(k)}(y) r(y)$

  Flip a coin with Heads probability equal to $q_t(x) = \frac{r(x)}{Q_t}$

  if *coin comes up Heads* **then**

    Hire i and halt.

  else

    Reject i and proceed.

  end

end

3.2 Algorithm is well-defined

For this algorithm to be well-defined, the coin flip must always be heads with a valid probability. This means that $0 \leq q_t(x) \leq 1$ for all t. $r(x)$ is always positive, so this condition will be satisfied as long as $Q_t \geq r(x)$, because then Q_t will be positive, so $0 \leq \frac{r(x)}{q_t} \leq 1$.
First consider the inequality

$$r(x) + \sum_{k=1}^{n} \sum_{y \in \mathcal{S}} r(y) \cdot f_{\pi(k)}(y) \leq 1.$$

To show that this is true, plug in the definition of $r(x)$ and $f_{\pi(k)}(y)$ into the right hand side of the expression. This gives

$$\frac{\Pr[X_{max} = x]}{2 \sum_{i=1}^{n} \Pr[X_i = x]} + \sum_{k=1}^{n} \sum_{y \in \mathcal{S}} \frac{\Pr[X_{max} = y]}{2 \sum_{i=1}^{n} \Pr[X_k = y]} \cdot \Pr[X_k = y].$$

Rearranging the sum in the right hand side of the expression gives

$$\begin{aligned} & \sum_{k=1}^{n} \sum_{y \in \mathcal{S}} \frac{\Pr[X_{\max} = y]}{2 \sum_{i=1}^{n} \Pr[X_i = y]} \cdot \Pr[X_i = y] \\ = \quad & \sum_{y \in \mathcal{S}} \Pr[X_{\max} = y] \cdot \frac{\sum_{k=1}^{n} \Pr[X_k = y]}{2 \sum_{i=1}^{n} \Pr[X_i = y]} \\ = \quad & \frac{1}{2} \sum_{y \in \mathcal{S}} \Pr[X_{\max} = y] \end{aligned}$$

$$= \frac{1}{2}.$$

To bound the left hand side of the expression, consider that

$$\mathbb{1}(X_{max} = x) \leq \sum_{i=1}^{n} \mathbb{1}(X_i = x).$$

This is because if $X_{max} = x$, then for some i, $X_i = x$. Taking the expectation of these terms gives that

$$\Pr[X_{max} = x] \leq \sum_{i=1}^{n} \Pr[X_i = x].$$

This implies that

$$\frac{\Pr[X_{max} = x]}{2\sum_{i=1}^{n} \Pr[X_i = x]} \leq \frac{1}{2}.$$

This proves that

$$r(x) + \sum_{k=1}^{n} \sum_{y \in \mathcal{S}} r(y) \cdot f_{\pi(k)}(y) \leq \frac{1}{2} + \frac{1}{2} = 1.$$

This implies that for all π,

$$r(x) + \sum_{k=1}^{n-1} \sum_{y \in \mathcal{S}} r(y) \cdot f_{\pi(k)}(y) \leq 1.$$

This in turn implies that

$$r(x) + \sum_{k=1}^{t-1} \sum_{y \in \mathcal{S}} r(y) \cdot f_{\pi(k)}(y) \leq 1, \quad \forall t \in [n],$$

which means that for all t, $Q_t \geq r(x)$, proving that this algorithm is well defined.

3.3 Algorithm is IIF and TIF

To demonstrate that the algorithm is IIF and TIF, we will show that

$$\Pr[\text{ALG hires } \pi(i) | X_{\pi(i)} = x] = r(x).$$

This can be shown by induction. For the base case, take the first random variable that the algorithm looks at. In this case $Q_1 = 1$, because $t - 1 = 0$, so the probability that the algorithm selects $X_{\pi(1)}$ conditioned on $X_{\pi(1)} = x$ is $r(x)$. Therefore, $\Pr[\text{ALG hires } \pi(1)|X_{\pi(1)} = x] = r(x)$. Now, the inductive step will use strong induction and will assume that for all $i \leq t-1$, $\Pr[\text{ALG hires } \pi(i)|X_{\pi(i)} = x] = r(x)$. Now, this means that

$$\sum_{k=1}^{t-1} \sum_{y \in \mathcal{S}} f_{\pi(k)}(y)p(y) = \sum_{k=1}^{t-1} \Pr[\text{ALG hires } \pi(k)] = \Pr[\text{ALG hires something before } t].$$

Therefore,

$$\Pr[\text{ALG reaches time } t] = 1 - \Pr[\text{ALG hires something before } t] = 1 - \sum_{k=1}^{t-1} \sum_{y \in \mathcal{S}} f_{\pi(k}(y)p(y) = Q_t.$$

Now, the probability that $X_{\pi(t)}$ is selected is the probability that the algorithm reaches variable t, and then the coin toss comes up heads. These events are, by construction, independent, so

$$\Pr[\text{ALG hires } \pi(t) | X_{\pi(t)} = x] = Q_t \cdot \frac{r(x)}{Q_t} = r(x).$$

This completes the inductive proof. The function $r(x)$ does not depend on i, meaning that the probability of selection of a given random variable only depends on the value of the realization, x, nor does it depend on when is the sequence that random variable is seen. Therefore this algorithm is IIF and TIF.

3.4 Algorithm is $\frac{1}{2}$-competitive

This algorithm can be shown to be $\frac{1}{2}$-competitive by calculating the expected value of the algorithm.

Theorem 6. *Algorithm 1 satisfies the property that*

$$\mathbb{E}[ALG] = \frac{1}{2}\mathbb{E}[PROPHET]$$

Proof. As a first step,

$$\begin{aligned}\mathbb{E}[\text{ALG}] = \sum_{i=1}^{n} \sum_{x \in S} & \mathbb{E}[\text{ALG} | \text{ALG hires } X_i \text{ and } X_i = x] \\ & \cdot \Pr[\text{ALG hires } X_i \text{ and } X_i = x].\end{aligned}$$

Now,

$$\mathbb{E}[\text{ALG} | \text{ALG hires } X_i \text{ and } X_i = x] = x,$$

and

$$\begin{aligned}\Pr[\text{ALG hires } X_i \text{ and } X_i = x] &= \Pr[\text{ALG hires } X_i | X_i = x] \Pr[X_i = x] \\ &= \frac{\Pr[X_{max} = x]}{2\sum_{i=1}^{n} \Pr[X_i = x]} \Pr[X_i = x].\end{aligned}$$

The final equality comes from the proof in the prior section. Now, plugging these expressions into the expected value of the algorithm gives

$$\mathbb{E}[\text{ALG}] = \sum_{i=1}^{n} \sum_{x \in S} x \cdot \frac{\Pr[X_{max} = x]}{2\sum_{i=1}^{n} \Pr[X_i = x]} \Pr[X_i = x].$$

Rearranging the sum gives the expression

$$\mathbb{E}[\text{ALG}] = \sum_{x \in S} x \cdot \Pr[X_{max} = x] \cdot \frac{\sum_{i=1}^{n} \Pr[X_i = x]}{2\sum_{i=1}^{n} \Pr[X_i = x]} = \frac{1}{2} \sum_{x \in S} x \cdot \Pr[X_{max} = x] = \frac{1}{2}\mathbb{E}[\text{PROPHET}].$$

This proves that the expected value of this algorithm is exactly $\frac{1}{2}$ the expected value of the prophet's pick. ☐

4 Linear Programming to Solve k-Select Prophet Inequalities

Constructing fair k-selection algorithms which have a competitive ratio close to 1 directly is very difficult, so we consider instead a linear programming based approach. To set up this approach, it is necessary to find a linear program such that its solutions are in bijection with online implementable algorithms. We would then like to understand what corresponds to the expected value of the online implementable algorithm. In this section, inspired by [8] and [11], we investigate these questions. We conjecture that we can use our linear program to develop an IIF and TIF algorithm for making k selections.

4.1 Linear program for k selections

Consider the following linear program, with variables $h_{i,j}(x)$.

$$0 \leq h_{i,0}(x) \leq \Pr[X_i = x]\left(1 - \sum_x \sum_{i'<i} h_{i',0}(x)\right)$$
$$0 \leq h_{i,j}(x) \leq \Pr[X_i = x]\left(\sum_x \sum_{i'<i} h_{i',j-1}(x) - \sum_x \sum_{i'<i} h_{i',j}(x)\right)$$
$$h_{i,j}(x) = 0 \text{ for } j \geq k \text{ or } j \leq -1 \text{ or } i \leq 0.$$

The final condition in the case $j \leq -1$ or $i \leq 0$ is largely for convenience of the following proofs.

4.1.1 Bijection between algorithms and linear program

The next step is to establish a bijection between solutions to this linear program and online implementable k-selection algorithms.

Theorem 7. *Any online implementable algorithm gives a valid solution to this linear program where*

$$h_{i,j}(x) = \Pr[\textit{ALG selects } X_i \text{ \& } j \textit{ selections have been made} \text{ \& } X_i = x].$$

Proof. Consider for an online algorithm

$$h_{i,j}(x) = \Pr[\text{ALG selects } X_i \text{ \& } j \text{ selections have been made \& } X_i = x].$$

Using the fact that

$$0 \leq \Pr[\text{ALG selects } X_i \mid j \text{ selections have been made \& } X_i = x] \leq 1,$$

it is easy to check that $h_{i,j}(x)$ must satisfy the linear program. □

Now, let us consider a solution to the linear program, and construct the corresponding Algorithm 4.

This algorithm is automatically well-defined because, by definition, $0 \leq p_{i,j}(x) \leq 1$.

Now define:

Algorithm 4: Online k-Selection Algorithm

```
Parameters: I = ((X_i)_{i=1}^n, π)
j ← 0
for t = 1, ..., n do
    i ← π(t)
    Inspect X_i and let x ← X_i
    if j = 0 then
        p ← p_{0,j}(x) = h_{i,0}(x) / (Pr[X_i=x](1 − Σ_x Σ_{i<i'} h_{i,0}(x)))
    else
        p ← p_{i,j}(x) = h_{i,j}(x) / (Pr[X_i=x](Σ_x Σ_{i'<i} h_{i',j−1}(x) − Σ_x Σ_{i'<i} h_{i',j}(x)))
    end
    Flip a coin with Heads probability equal to p
    if coin comes up Heads then
        Hire i
        j ← j + 1
    else
        Reject i
    end
    if j = k then
        Halt
    end
end
```

Definition 5. *Let*

$$Q_{ij} := \begin{cases} 1 - \sum_x \sum_{i'<i} h_{i',0}(x) & \text{if } j = 0, \\ \sum_x \sum_{i'<i} h_{i',j-1}(x) - \sum_x \sum_{i'<i} h_{i',j}(x) & \text{if } j \geq 1. \end{cases}$$

To prove the bijection, it is necessary to prove the solutions $h_{i,j}(x)$ to the linear program that appear in Algorithm 4 are equivalent to the probabilities Pr[ALG selects X_i & j selections have been made & $X_i = x$]. To establish this result, it is necessary to prove Lemma 1

Lemma 1. *$Q_{i,j}$ is the probability of reaching variable i having made j selections.*

To prove Lemma 1, it is necessary to prove two intermediate lemmas. First, define

Definition 6.

$$p_{i,j} := \sum_{x \in S_i} p_{i,j}(x) \cdot Pr\left[X_i = x\right],$$

where $S_i = \text{Supp}[X_i]$. The next step is to prove the probabilistic interpretation of this quantity.

Lemma 2. *$p_{i,j} = Pr$[Selecting $X_i | j$ selections before]*

Proof.

$$\sum_{x \in S_i} p_{j,i}(x) \Pr\left[X_i = x\right] = \sum_{x \in S_i} \Pr[\text{Selecting } X_i | j \text{ selections before and } X_i = x]$$

$$\cdot \frac{\Pr\left[X_i = x \text{ and } j \text{ selections before}\right]}{P\left[j \text{ selections before}\right]}.$$

This equality follows because the realization of each random variable is independent of the picks of the algorithm up to this point. Now this becomes

$$\frac{\Pr[\text{Selecting } X_i \ \& \ j \text{ selections before}]}{\Pr[j \text{ selections before}]} = \Pr[\text{Selecting } X_i | j \text{ selections before}].$$

□

Lemma 3.

$$\begin{aligned}
&\Pr[\textit{reaching variable } i \textit{ having made } j \textit{ selections}] \\
&\quad = p_{i-1,j-1} \cdot \Pr[\textit{reaching variable } i-1 \textit{ having made } j-1 \textit{ selections}] \\
&\qquad + (1 - p_{i-1,j}) \cdot \Pr[\textit{reaching variable } i-1 \textit{ having made } j \textit{ selections}].
\end{aligned}$$

Proof. No selections can have been made before the first variable is realized, giving the base case. Next, there are two ways that j selections have been made before the ith variable, either $j-1$ selections were made at the $i-1$-th variable and it selects the $i-1$-th variable, or j selections were made at the $i-1$-th variable and it does not select the $i-1$-th variable. □

The next step is to prove that $Q_{i,j}$ is the probability of reaching variable i having made j selections. This can be done by first proving two lemmas.

Lemma 4. *$Q_{i,0}$ is the probability of reaching variable i having made 0 selections.*

Proof. This proof will be done inductively. For the base case $Q_{1,0} = 1$, because

$$Q_{1,0} = 1 - \sum_x \sum_{i'<1} h_{i,0}(x) = 1.$$

The probability of reaching the first variable having made zero selections is trivially one, so $Q_{1,0}$ is the probability of reaching variable 1 having made 0 selections.

Now, for the inductive step, assume that $Q_{i,0}$ is the probability of reaching variable i having made 0 selections.

$$\begin{aligned}
Q_{i+1,0} &= 1 - \sum_x \sum_{i'<i+1} h_{i',0}(x) \\
&= 1 - \sum_x \sum_{i'<i} h_{i',0}(x) - \sum_x h_{i,0}(x) \\
&= 1 - \sum_x \sum_{i'<i} h_{i',0}(x) - \sum_x \frac{h_{i,0}(x)}{\Pr[X_i = x]\left(1 - \sum_x \sum_{i'<i} h_{i',0}(x)\right)} \\
&\cdot \Pr[X_i = x]\left(1 - \sum_x \sum_{i'<i} h_{i',0}(x)\right) \\
&= Q_{i,0} - \left(\sum_x p_{i,0}(x) \Pr[X_i = x]\right) Q_{i,0}
\end{aligned}$$

$$= \left(1 - \sum_x p_{i,0}(x) \Pr[X_i = x]\right) Q_{i,0}.$$

Now, by Definition 6, this becomes

$$Q_{i+1,0} = (1 - p_{i,0})\, Q_{i,0},$$

so, by the inductive hypothesis, this becomes

$$Q_{i+1,0} = (1 - p_{i,0}) \Pr[\text{reaching variable } i \text{ having made 0 selections}].$$

Finally, by applying Lemma 3 and recognizing that the other term in the expression will be 0, because it is impossible to make negative selections, this gives the final result:

$$Q_{i+1,0} = \Pr[\text{reaching variable } i+1 \text{ having made 0 selections}],$$

completing the inductive step. □

Lemma 5. *$Q_{i,j}$ is the probability of reaching variable i having made j selections for $j \geq 1$.*

Proof. This proof will be done inductively. For the base case: $Q_{1,j} = 0$ for all $j \geq 1$ by Definition 5.

For the inductive step, assume that $Q_{i,j}$ is the probability of making it to variable i having made j selections for all j. Now, the goal is to prove that $Q_{i+1,j}$ is the probability of reaching the $i+1$-st variable having made j selections. Beginning with the definition of $Q_{i+1,j}$:

$$\begin{aligned}
Q_{i+1,j} &= \sum_x \sum_{i'<i+1} h_{i',j-1}(x) - \sum_x \sum_{i'<i+1} h_{i',j}(x) \\
&= \sum_x h_{i,j-1}(x) - \sum_x h_{i,j}(x) + \sum_x \sum_{i'<i} h_{i',j-1}(x) - \sum_x \sum_{i'<i} h_{i',j}(x) \\
&= \sum_x \frac{h_{i,j-1}(x)}{\Pr[X_i = x] \cdot \left(\sum_x \sum_{i'<i} h_{i',j-2}(x) - \sum_x \sum_{i'<i} h_{i',j-1}(x)\right)} \\
&\quad \cdot \Pr[X_i = x] \left(\sum_x \sum_{i'<i} h_{i',j-2}(x) - \sum_x \sum_{i'<i} h_{i',j-1}(x)\right) \\
&\quad - \sum_x \frac{h_{i,j}(x)}{\Pr[X_i = x] \cdot \left(\sum_x \sum_{i'<i} h_{i',j-1}(x) - \sum_x \sum_{i'<i} h_{i',j}(x)\right)} \\
&\quad \cdot \Pr[X_i = x] \left(\sum_x \sum_{i'<i} h_{i',j-1}(x) - \sum_x \sum_{i'<i} h_{i',j}(x)\right) \\
&\quad + \sum_x \sum_{i'<i} h_{i',j-1}(x) - \sum_x \sum_{i'<i} h_{i',j}(x) \\
&= \left(\sum_x p_{i,j-1}(x) \Pr[X_i = x]\right) Q_{i,j-1} + \left(1 - \sum_x p_{i,j}(x) \Pr[X_i = x]\right) Q_{i,j}.
\end{aligned}$$

Now, by Definition 6, this becomes

$$Q_{i+1,j} = p_{i,j-1} Q_{i,j-1} + (1 - p_{i,j}) Q_{i,j}.$$

Next, by applying Lemma 4 in cases where $Q_{i,0}$ appears and applying the inductive hypothesis for all $Q_{i,j}$ where $j \geq 1$ gives the following expression:

$$Q_{i+1,j} = p_{i,j-1} \cdot \Pr[\text{reaching variable } i \text{ having made } j-1 \text{ selections}] \\ + (1 - p_{i,j}) \cdot \Pr[\text{reaching variable } i \text{ having made } j \text{ selections}].$$

Now, by Lemma 3, this becomes

$$Q_{i+1,j} = \Pr[\text{reaching variable } i+1 \text{ having made } j \text{ selections}],$$

completing the inductive step. □

Proof. Proof of Lemma 1. This theorem follows directly from Lemmas 4 and 5. □

Now, to complete the bijection, show:

Theorem 8. *Any solution to this linear program gives a valid online implementable algorithm where*

$$h_{i,j}(x) = \Pr[ALG \text{ selects } X_i \text{ \& } j \text{ selections have been made \& } X_i = x].$$

Proof. This follows directly from substitution into the definition of $p_{i,j}(x)$ and Lemma 1. □

Having shown both directions, this gives:

Theorem 9. *There exists a bijection between solutions to this linear program and online implementable k-select algorithms where*

$$h_{i,j}(x) = \Pr[ALG \text{ selects } X_i \text{ \& } j \text{ selections have been made \& } X_i = x].$$

Proof. This follows from Theorems 7 and 8. □

4.1.2 The expected value of the algorithm

Now, we can derive that an expression for the expected value of the algorithm, giving us our objective function for the linear program.

Theorem 10.

$$\mathbb{E}[ALG] = \sum_x \sum_{i=1}^{n} \sum_{j=0}^{k} x \cdot h_{i,j}(x).$$

Proof. Immediately we have that

$$\Pr[\text{Selecting } X_i | X_i = x] = \sum_{j=0}^{k} p_{i,j}(x) Q_{i,j} = \sum_{j=0}^{k} \frac{h_{i,j}(x)}{\Pr\left[X_i = x\right] Q_{i,j}} Q_{i,j}.$$

This follows directly from Lemma 1.

$$\begin{aligned}
\mathbb{E}[ALG] &= \sum_{i=1}^{n}\sum_{x} x \cdot \Pr[\text{Selecting } X_i | X_i = x] \cdot \Pr\left[X_i = x\right] \\
&= \sum_{i=1}^{n}\sum_{x}\sum_{j=0}^{k} x \cdot \frac{h_{i,j}(x)}{\Pr\left[X_i = x\right]} \cdot \Pr\left[x_i = x\right] \\
&= \sum_{x}\sum_{i=1}^{n}\sum_{j=0}^{k} x \cdot h_{i,j}(x).
\end{aligned}$$

This completes the proof of theorem 10.

□

Having calculated the expected value of the algorithm, this gives the complete linear program:

$$\begin{aligned}
\text{Maximize} \quad & \sum_{i=1}^{n}\sum_{x}\sum_{j=0}^{k} x \cdot h_{i,j}(x) \\
\text{subject to} \quad & 0 \leq h_{i,0}(x) \leq \Pr\left[X_i = x\right]\left(1 - \sum_{x}\sum_{i'<i} h_{i',0}(x)\right), \\
& 0 \leq h_{i,j}(x) \leq \Pr\left[X_i = x\right]\left(\sum_{x}\sum_{i'<i} h_{i',j-1}(x) - \sum_{x}\sum_{i'<i} h_{i',j}(x)\right), \\
& h_{i,j}(x) = 0 \text{ for } j \geq k \text{ or } j \leq -1 \text{ or } i \leq 0.
\end{aligned}$$

5 An Exploration of Offline IIF Algorithms

Up until now, we've focused on variations of the prophet inequality, where the values of the random variables are realized one at a time in an online fashion and we have to choose to accept or reject at each realization.

However, in this section we'll discuss an offline version of the problem, where we know the realizations of all the random variables and must choose one of them. Again, the goal is to maximize the expected value of our pick.

Clearly, without any fairness restrictions, we could always choose the largest value and obtain $\mathbb{E}[\text{ALG}] = \mathbb{E}[\text{PROPHET}]$. However, we'll see that when we restrict ourselves to using IIF algorithms, then this isn't always the case.

5.1 Example where any offline IIF algorithm is at best $\frac{25}{27}$-competitive

In this section, we'll give the explicit random variables promised in Theorem 2. We'll prove that any offline IIF algorithm on these variables will be at best $\frac{25}{27}$-competitive.
We provide the following example:

$$X_1 = \begin{cases} 0 & \text{w.p. } \frac{7}{9} \\ 1 & \text{w.p. } \frac{2}{9} \end{cases} \qquad X_2 = \begin{cases} 1 & \text{w.p. } \frac{1}{3} \\ \frac{5}{2} & \text{w.p. } \frac{2}{3} \end{cases}$$

Proof. In order to get this $\frac{25}{27}$ bound, we must first define some notation. We let

$$f_{a,b} = \Pr[\text{ALG choose } X_1 | X_1 = a \text{ AND } X_2 = b]$$

Similarly, we can define

$$g_{a,b} = \Pr[\text{ALG choose } X_2 | X_1 = a \text{ AND } X_2 = b]$$

In order for this to be a valid offline selection (i.e. choosing at most one element in any given run), we must have

$$\begin{aligned} f_{a,b}, g_{a,b} &\geq 0 \quad \forall a, b \\ f_{a,b} + g_{a,b} &\leq 1 \quad \forall a, b \end{aligned}$$

Now, let's figure out how to fit the IIF constraint. All we need to ensure is that

$$\Pr[\text{choose } X_1 | X_1 = 1] = \Pr[\text{choose } X_2 | X_2 = 1]$$

We calculate that

$$\begin{aligned} \Pr[\text{choose } X_1 | X_1 = 1] &= \Pr[X_2 = 1]\, f_{1,1} + \Pr\left[X_2 = \frac{5}{2}\right] f_{1,\frac{5}{2}} \\ &= \frac{1}{3} f_{1,1} + \frac{2}{3} f_{1,\frac{5}{2}} \end{aligned}$$

Similarly, we can calculate that

$$\Pr[\text{choose } X_2 | X_2 = 1] = \frac{7}{9} g_{0,1} + \frac{2}{9} g_{1,1}$$

Thus, the IIF condition imposes the following equality:

$$\frac{1}{3} f_{1,1} + \frac{2}{3} f_{1,\frac{5}{2}} = \frac{7}{9} g_{0,1} + \frac{2}{9} g_{1,1}$$

For reasons that will become clear later, we divide this equation by 3, resulting in

$$\frac{1}{9} f_{1,1} + \frac{2}{9} f_{1,\frac{5}{2}} = \frac{7}{27} g_{0,1} + \frac{2}{27} g_{1,1}$$

Next, we calculate the expected value of our algorithm. We see that

$$\begin{aligned} & \mathbb{E}[\text{ALG}] \\ = \; & \Pr[X_1 = 0, X_2 = 1](0 \cdot f_{0,1} + 1 \cdot g_{0,1}) + \Pr\left[X_1 = 0, X_2 = \frac{5}{2}\right](0 \cdot f_{0,\frac{5}{2}} + \frac{5}{2} \cdot g_{0,\frac{5}{2}}) \\ & + \Pr[X_1 = 1, X_2 = 1](1 \cdot f_{1,1} + 1 \cdot g_{1,1}) + \Pr\left[X_1 = 1, X_2 = \frac{5}{2}\right](1 \cdot f_{1,\frac{5}{2}} + \frac{5}{2} \cdot g_{1,\frac{5}{2}}) \\ = \; & \frac{7}{27} g_{0,1} + \frac{35}{27} g_{0,\frac{5}{2}} + \frac{2}{27} f_{1,1} + \frac{2}{27} g_{1,1} + \frac{4}{27} f_{1,\frac{5}{2}} + \frac{10}{27} g_{1,\frac{5}{2}} \\ = \; & \left(\frac{7}{27} g_{0,1} + \frac{2}{27} g_{1,1}\right) + \left(\frac{35}{27} g_{0,\frac{5}{2}} + \frac{2}{27} f_{1,1} + \frac{4}{27} f_{1,\frac{5}{2}} + \frac{10}{27} g_{1,\frac{5}{2}}\right) \\ = \; & \left(\frac{1}{9} f_{1,1} + \frac{2}{9} f_{1,\frac{5}{2}}\right) + \left(\frac{35}{27} g_{0,\frac{5}{2}} + \frac{2}{27} f_{1,1} + \frac{4}{27} f_{1,\frac{5}{2}} + \frac{10}{27} g_{1,\frac{5}{2}}\right) \end{aligned}$$

$$= \frac{35}{27} g_{0,\frac{5}{2}} + \frac{5}{27} f_{1,1} + \frac{10}{27} f_{1,\frac{5}{2}} + \frac{10}{27} g_{1,\frac{5}{2}}$$

Due to the restrictions $g_{0,\frac{5}{2}} \leq 1$, $f_{1,1} \leq 1$ and $f_{1,\frac{5}{2}} + g_{1,\frac{5}{2}} \leq 1$, we have $\mathbb{E}[\text{ALG}] \leq \frac{35}{27} + \frac{5}{27} + \frac{10}{27} = \frac{50}{27}$. To figure out how competitive this is, we can calculate the prophet's expected value: $\mathbb{E}[X_{\max}]$. We see that $X_{\max} = X_2$, and we have

$$\mathbb{E}[X_{\max}] = \frac{1}{3} \cdot 1 + \frac{2}{3} \cdot \frac{5}{2} = 2$$

Thus, for any IIF algorithm, we have

$$\frac{\mathbb{E}[\text{ALG}]}{\mathbb{E}[\text{PROPHET}]} \leq \frac{50/27}{2} = \frac{25}{27}$$

□

5.2 Considering offline IIF algorithms which only choose the maximum

In this section, we will prove Theorem 3. For reference, we copy it below:

Theorem. *When we consider an offline IIF algorithm, along with the condition that it can only either accept the maximum or not make a decision at all, then it can be at best $\frac{1}{2}$-competitive.*

Proof. To prove this, we'll consider the following random variables:

$$X_1 = \begin{cases} 0 & \text{w.p. } 1-\epsilon \\ U(1, 1+\epsilon) & \text{w.p. } \epsilon \end{cases} \quad X_2 = \begin{cases} 0 & \text{w.p. } \epsilon \\ U(1, 1+\epsilon) & \text{w.p. } 1-\epsilon \end{cases}$$

for some $\epsilon > 0$, where $U(1, 1+\epsilon)$ is the continuous uniform distribution between 1 and $1+\epsilon$. That is, $U(1, 1+\epsilon)$ has the pdf

$$f(x) = \begin{cases} \frac{1}{\epsilon} & \text{for } 1 \leq x \leq 1+\epsilon \\ 0 & \text{for } x < 1 \text{ or } x > 1+\epsilon \end{cases}$$

For some $1 \leq x \leq 1+\epsilon$, let's consider the scenario where $X_1 = x$. The algorithm is only allowed to choose X_1 when $X_2 < x$, which happens with probability

$$\Pr[X_2 < x] = \Pr[X_2 = 0] + \Pr[X_2 > 0] \cdot \Pr[X_2 < x | X_2 > 0] = \epsilon + (1-\epsilon)\left(\frac{x-1}{\epsilon}\right)$$

Because X_1 can only be chosen if it is larger than X_2, we obtain

$$\Pr[\text{ALG choose } X_1 | X_1 = x] \leq \Pr[X_2 < x] = \epsilon + (1-\epsilon)\left(\frac{x-1}{\epsilon}\right)$$

The IIF condition then implies that we have

$$\Pr[\text{ALG choose } X_2 | X_2 = x] \leq \epsilon + (1-\epsilon)\left(\frac{x-1}{\epsilon}\right)$$

Now, we obtain bounds for $\mathbb{E}[\text{PROPHET}]$ and $\mathbb{E}[\text{ALG}]$. For the prophet, we have

$$\mathbb{E}[\text{PROPHET}] = \mathbb{E}[X_{\max}] \geq \mathbb{E}[X_2] = (1-\epsilon)\left(1 + \frac{\epsilon}{2}\right) \geq 1-\epsilon$$

For the algorithm, we take

$$\begin{aligned}
&\mathbb{E}[\text{ALG}] \\
= \; & \int_1^{1+\epsilon} x\,\epsilon f(x) \Pr[\text{ALG choose } X_1 | X_1 = x]\,dx + \int_1^{1+\epsilon} x(1-\epsilon)f(x) \Pr[\text{ALG choose } X_2 | X_2 = x]\,dx \\
\leq \; & \int_1^{1+\epsilon} x\,\epsilon f(x)\,dx + \int_1^{1+\epsilon} x(1-\epsilon)f(x) \Pr[\text{ALG choose } X_2 | X_2 = x]\,dx \\
\leq \; & \int_1^{1+\epsilon} x(\epsilon)\left(\frac{1}{\epsilon}\right)dx + \int_1^{1+\epsilon} x(1-\epsilon)\left(\frac{1}{\epsilon}\right)\left(\epsilon + (1-\epsilon)\left(\frac{x-1}{\epsilon}\right)\right)dx \\
= \; & \int_1^{1+\epsilon} x\,dx + \frac{1-\epsilon}{\epsilon}\int_1^{1+\epsilon} x\left(\epsilon + (1-\epsilon)\left(\frac{x-1}{\epsilon}\right)\right)dx \\
= \; & \frac{1}{2}\epsilon(\epsilon+2) + \left(\frac{1-\epsilon}{\epsilon}\right)\left(\frac{\epsilon(\epsilon^2+5\epsilon+3)}{6}\right) \\
= \; & \frac{\epsilon(\epsilon+2)}{2} + \frac{(1-\epsilon)(\epsilon^2+5\epsilon+3)}{6}
\end{aligned}$$

As we take $\epsilon \to 0$, these inequalities become $\mathbb{E}[\text{PROPHET}] \geq 1$ and $\mathbb{E}[\text{ALG}] \leq \frac{1}{2}$. Thus, we conclude that for every $\epsilon > 0$ we have

$$\frac{\mathbb{E}[\text{ALG}]}{\mathbb{E}[\text{PROPHET}]} \leq \frac{1}{2} + \epsilon$$

so the algorithm is at best $\frac{1}{2}$-competitive. □

6 A Double-Sample Online Algorithm Satisfying IIF and TIF

Without any consideration of fairness, the prophet inequality has a $\frac{1}{2}$ bound. In the literature, Arsenis and Kleinberg [8] introduce two notions of fairness, IIF and TIF, and provide an online double sample algorithm satisfying both. The bound, however, worsens to $\frac{1}{9}$. In this section, we analyze whether this $\frac{1}{9}$ bound can be improved upon.

Let $f_1, \ldots, f_n$ be probability density functions. Let $X_1, \ldots, X_n, Y_1, \ldots, Y_n, Z_1, \ldots, Z_n$ be continuous random variables such that $X_i, Y_i, Z_i \sim f_i$ for each $i = 1, \ldots, , n$. In the following algorithm due to Arsenis and Kleinberg, introduced earlier as algorithm 2, the variables being selected are $X_1, \ldots, X_n$, which are processed in an online fashion. The realizations of $Y_1, \ldots, , Y_n, Z_1, \ldots, Z_n$ are revealed offline, before the actual algorithm starts.

Algorithm: Double-sample Online IIF and TIF Algorithm

Data: $Y_i, Z_i \sim f_i, \pi \in S_n$

for *t=1, ..., n* **do**

  Observe $X_{\pi(t)} \sim f_{\pi(t)}$.

  if $X_{\pi(t)} > Y_{\max}$ *and* ($X_{\pi(s)} < Y_{\max}$ *for all* $s < t$) *and* ($Z_{\pi(s)} < Y_{\max}$ *for all* $s \geq t$) **then**

    Hire $X_{\pi(t)}$.

  end

end

Note that the order of the X_i's can vary. Moreover, all comparisons made can be safely assumed to be strict, since all random variables are continuous.

Arsenis and Kleinberg have showed that for discrete random variables (tie-breaking mechanism), the algorithm is IIF and TIF [8]. Moreover, they proved that

$$\mathbb{E}[ALG] \geq \frac{1}{9}\mathbb{E}[X_{\max}]$$

[8]. Here, we seek to work with continuous random variables, allowing us to formalize Arsenis and Kleinberg's tie-breaking mechanism and to demonstrate a better bound of $\frac{1}{6}$. That is, we claim that

$$\mathbb{E}[ALG] \geq \frac{1}{6}\mathbb{E}[X_{\max}].$$

This improves upon the original result, proving the IIF and TIF algorithm Arsenis and Kleinberg provided does better than claimed.

6.1 $\frac{1}{6}$ bound with continuous random variables

In this section, we prove theorem 4, which we recall is

Theorem. *Algorithm 2 is IIF and TIF. Moreover,* $\mathbb{E}[ALG] \geq \frac{1}{6}\mathbb{E}[X_{\max}]$.

Proof. First, we convert to continuous random variables, and then, using the continuous definitions of IIF and TIF, show that the algorithm is still IIF and TIF. Afterwards, we compute the expected value of the algorithm and show that it is at least $\frac{1}{6}$ the prophet's pick.

Now, let's prove that the algorithm (which in this section is ALG) is still IIF and TIF. For each $\pi \in S_n$ and $i = 1, , \dots, , n$, let $i = \pi(t)$. Then we have

$$\begin{aligned}\Pr(S_{i,\pi} < x) &= \Pr(ALG^{\pi} \text{ accepts } X_i \text{ \& } X_i < x) \\ &= \Pr(Y_{\max} < X_i \leq x \text{ \& } X_{\pi(s),s<t} < Y_{\max} \text{ \& } Z_{\pi(s),s\geq t} < Y_{\max})\end{aligned}$$

By the independence of $X_1, \dots, , X_n$ and $Z_1, , \dots, , Z_n$, we can make an independent copies $X'_1, , \dots, , X'_n$ of the random variables $X_1, , \dots, , X_n$, call their maximum $X'_{\max}$, and write

$$\begin{aligned}\Pr(S_{i,\pi} < x) &= \Pr(x > X_i > Y_{\max} \text{ \& } X'_{\max} < Y_{\max}) \\ &= \mathbb{E}[\mathbb{1}(x > X_i > Y_{max} > X'_{\max})] \\ &= \int_0^\infty \int_0^\infty \int_0^\infty \mathbb{1}(x > w > z > y) f_i(w) f_{X'_{\max}}(y) f_{Y_{\max}}(z)\, dw dy dz\end{aligned}$$

by the independence of $X_i, X'_{\max}, Y_{max}$. Note that $X'_{\max}$ and $Y_{\max}$ have the same distribution. Denote their common pdf by $f(x)$, and common cdf by $F(x)$. Then

$$\begin{aligned}\Pr(S_{i,\pi} < x) &= \int_0^\infty \int_0^\infty \int_0^\infty \mathbb{1}(x > w > z > y) f_i(w) f(y) f(z)\, dw dy dz \\ &= \int_0^x \int_0^w \int_0^z f_i(w) f(y) f(z)\, dy dz dw \\ &= \int_0^x \int_0^w f_i(w) F(z) f(z)\, dz dw\end{aligned}$$

By the definition of a continuous random variable, S_i, being the integral of a continuous function, is continuous. Moreover,

$$s_{i,\pi}(x) = \int_0^x f_i(x)F(z)f(z)\,dz$$

and

$$\begin{aligned}\frac{s_{i,\pi}(x)}{f_i(x)} &= \frac{\int_0^x f_i(x)F(z)f(z)\,dz}{f_i(x)} \\ &= \int_0^x F(z)f(z)\,dz \\ &= \int_0^x F(z)F'(z)\,dz \\ &= \frac{1}{2}F(x)^2\end{aligned}$$

As this expression does not depend on π or i, the algorithm is IIF and TIF, as desired.

Having proven the algorithm is IIF and TIF, we compute $\mathbb{E}[ALG^{\pi}]$ and prove the $\frac{1}{6}$ bound. Before we do so, observe that

$$\begin{aligned}\Pr[Y_{max} \leq y] &= \Pr[Y_1, , \ldots, , Y_n \leq y] \\ &= \prod_{i=1}^{n} F_i(y)\end{aligned}$$

so that by the product rule

$$\begin{aligned}f(y) &= \Pr[Y_{max} \leq y]' \\ &= \sum_{i=1}^{n} \frac{f_i(y)}{F_i(y)} \cdot F_1(y)F_2(y)\cdots F_n(y) \\ &\leq \sum_{i=1}^{n} \frac{f_i(y)}{F_i(y)} \cdot F_i(y) \\ &= \sum_{i=1}^{n} f_i(y)\end{aligned}$$

Thus

$$\begin{aligned}\mathbb{E}[ALG^{\pi}] &= \sum_{i=1}^{n} \mathbb{E}[S_{i,\pi}] \\ &= \sum_{i=1}^{n} \int_0^{\infty} s_{i,\pi}(x)x\,dx \\ &= \sum_{i=1}^{n} \int_0^{\infty} \frac{1}{2}F(x)^2 \cdot f_i(x)x\,dx \\ &= \int_0^{\infty} \frac{1}{2}F(x)^2 \cdot x \sum_{i=1}^{n} f_i(x)\,dx\end{aligned}$$

$$\geq \int_0^\infty \frac{1}{2} F(x)^2 \cdot x f(x)\, dx$$

Noting that $(F(x)^3)' = 3F(x)^2 f(x)$, observe that

$$\int_0^\infty \frac{1}{2} F(x)^2 \cdot x f(x)\, dx = \frac{1}{6} \int_0^\infty (F(x)^3)' \cdot x\, dx$$

which by the Tail Integral Formula implies

$$\begin{aligned}
\frac{1}{6} \int_0^\infty (F(x)^3)' \cdot x\, dx &= \frac{1}{6} \int_0^\infty (1 - F(x)^3)\, dx \\
&\geq \frac{1}{6} \int_0^\infty (1 - F(x))\, dx \\
&= \frac{1}{6} \int_0^\infty f(x) x\, dx \\
&= \frac{1}{6} \mathbb{E}[X_{\max}]
\end{aligned}$$

Thus

$$\mathbb{E}[ALG] \geq \frac{1}{6} \mathbb{E}[X_{\max}]$$

and the result follows. □

7 Single Sample Prophet Inequalities with k Selections

7.1 Introduction

In this section, we will establish the following result (Theorem 5) for algorithm 3:

Theorem. *For any k, Algorithm 3 is $\frac{1}{2}$-competitive, that is,*

$$\mathbb{E}[ALG] \geq \frac{1}{2} \mathbb{E}[PROPHET]$$

where the prophet's performance is defined as the sum of the k highest elements of $\{X_1, \ldots, X_n\}$. This has proven been proven for the case where $k = 1$ [7] and $k = 2$ [10], but not for higher values of k. Here, we will prove the result for all $k \geq 2$. We will further prove that for any value of k, this lower bound in expectation is tight.

We will actually prove this theorem for a slightly different algorithm (not actually implementable in online decision-making), algorithm 5.

In the following, we will implictly let $X^m = -\infty$ for $m > n$, and similarly $a_i = -\infty$ for $i > 2n$ and we will refer to the realizations of the random variables $X_1 \ldots X_n$ as X-values, or simply X's. Since this algorithm always selects the worst (up to) k X's that could have been chosen by Algorithm 3, we have clearly that Algorithm 3 will always perform better than Algorithm 5, and so it suffices to prove

Algorithm 5: Worse Single-Sample k-select Algorithm

Data: $Y_i \sim X_i$
Let $Y^1 = Y_{\max}$, $Y^k = \max(\{Y_1, \ldots, Y_n\} - \{Y^1, \ldots, Y^{k-1}\})$
for t=1, ..., n **do**
  if $X^t > Y^k$ *and* $Y^k > X^{t+1}$ **then**
    Hire $\{X^{\max(1,t+1-k)}, \ldots, X^t\}$.
  end
end

Theorem 11. *For any k, Algorithm 5 is $\frac{1}{2}$-competitive, that is,*

$$\mathbb{E}[ALG] \geq \frac{1}{2}\mathbb{E}[PROPHET]$$

Hereafter, "algorithm" will be used generically to refer to Algorithm 5.
In order to prove the theorem, note that we can express the algorithm's expected performance as

$$\sum_{i=1}^{k} \mathbb{E}[A_i]$$

where A_i denotes the algorithm's ith largest pick, if any. Thus the algorithm's expected performance is

$$\sum_{i=1}^{k} \int_0^\infty \Pr(A_i \geq x)dx$$

And the prophet's becomes

$$\sum_{i=1}^{k} \mathbb{E}[X^i] = \sum_{i=1}^{k} \int_0^\infty \Pr(X^i \geq x)dx$$

where X^i is defined analagously to Y^i as the ith largest order statistic from the X's. Conditioning on $\{a_1 > a_2 > \ldots > a_{2n}\}$, the ranked list of the realizations $\{X_1, \ldots, X_n, Y_1, \ldots, Y_n\}$, it suffices to show that for any j,

$$\sum_{i=1}^{k} \Pr(A_i \geq a_j) \geq \frac{1}{2} \sum_{i=1}^{k} \Pr(X^i \geq a_j)$$

The former sum is the expected number of algorithm picks above a_j. This can be expressed as the sum of probabilities from $i = 1$ to n that $X^i \geq a_j$ and the algorithm picks X^i. It is guaranteed that $X^i \geq a_j$ and the algorithm picks X^i when $X^i > Y^{j+1-i}$, $Y^k > X^{i+k}$ and $X^i > Y^k$. Therefore,

$$\begin{aligned}
\sum_{i=1}^{k} \Pr(A_i \geq a_j) &\geq \sum_{i=1}^{n} \Pr(X^i > Y^{j+1-i}, Y^k > X^{i+k}, X^i > Y^k) \\
&\geq \sum_{i=1}^{j+1-k} \Pr(X^i > Y^k > X^{i+k}) + \sum_{i=j+2-k}^{k} \Pr(X^i > Y^{j+1-i}, Y^k > X^{i+k})
\end{aligned}$$

We now have the following important lemma which will let us provide a simple analysis of the second sum above:

Lemma 6. *If $i > j - k$, then*

$$\Pr(X^i > Y^{j+1-i}, Y^k > X^{i+k}) \geq \frac{1}{2}\Pr(X^i > Y^{j+1-i}).$$

Proof. Consider the quantity $\Pr(X^i > Y^{j+1-i}, Y^k > X^{i+k}) = \Pr(X^i > Y^{j+1-i})\Pr(Y^k > X^{i+k}|X^i > Y^{j+1-i})$. Let us concentrate on the conditional probability. We know that there is a certain number θ of Y's, $\theta \leq j - i$, greater than X^i. We have also that $Y^k > X^{i+k}$ if the sequence $a_1, \ldots, a_{i+2k-1}$ contains at least k Y's. This will hold if $a_{i+\theta+1}, \ldots, a_{i+2k-1}$ contains at least $k - \theta$ Y's. Consider revealing the θ X's sampled from the same distribution as $Y^1, \ldots, Y^\theta$. At most θ spots amongst $a_{i+\theta+1}, \ldots, a_{i+2k-1}$ have been taken up by these X's. The remaining spots, of which there are $2k - 2\theta - 1$ each have at least a $\frac{1}{2}$ chance of being a Y value, and we would like $k - \theta$ of these spots to be Y's. If each spot had exactly a $\frac{1}{2}$ chance of being a Y, then the chance of $k - \theta$ Y's would be exactly $\frac{1}{2}$, so we conclude $\Pr(Y^k > X^{i+k}|X^i > Y^{j+1-i}) \geq \frac{1}{2}$. □

The lemma lets us conclude that,

$$\sum_{i=1}^{k}\Pr(A_i \geq a_j) \geq \sum_{i=1}^{j+1-k}\Pr(X^i > Y^k > X^{i+k}) + \frac{1}{2}\sum_{i=j+2-k}^{k}\Pr(X^i > Y^{j+1-i})$$

Meanwhile, for the prophet,

$$\frac{1}{2}\sum_{i=1}^{k}\Pr(X^i \geq a_j) = \frac{1}{2}\sum_{i=1}^{\min(j+1-k,k)}\Pr(X^i > Y^{j+1-i}) + \frac{1}{2}\sum_{i=j+2-k}^{k}\Pr(X^i > Y^{j+1-i})$$

Therefore, it suffices to show that

$$\sum_{i=1}^{j+1-k}\Pr(X^i > Y^k > X^{i+k}) \geq \frac{1}{2}\sum_{i=1}^{\min(j+1-k,k)}\Pr(X^i > Y^{j+1-i}) \tag{4}$$

for any k, j.

For the remainder of this section, we will concentrate on establishing this inequality.

We will use the following lemma due to Nuti and Vondrák [12] in future sections:

Lemma 7. *If $a \leq b$, then $\Pr(X^a < Y^b)$ or equivalently $\Pr(Y^a < X^b)$ is at most $\Pr(Z < a)$, where Z is a binomial random variable with $n = a + b - 1$ and $p = \frac{1}{2}$. This is equivalent to what the probability that $X^a < Y^b$ would be if each of $a_1, \ldots, a_{2n}, \ldots$ were independently and uniformly at random assigned to be either X or Y, which we will hereafter refer to as "the binomial case".*

7.2 $k = 2$

When $j = 1$ this holds trivially as both sides of (4) vanish.

7.2.1 $j = 2$

Here we must show

$$\Pr(X^1 > Y^2 > X^3) \geq \frac{1}{2}\Pr(X^1 > Y^2)$$

which follows by direct application of Lemma 6 with $i = 1, j = k = 2$.

7.2.2 $j = 3$

Here we show that

$$\sum_{i=1}^{2} \Pr(X^i > Y^2 > X^{i+2}) \geq \frac{1}{2} \sum_{i=1}^{2} \Pr(X^i > Y^{4-i})$$

Since $\Pr(X^2 > Y^2) = \frac{1}{2}$ by symmetry, we have that the right hand side is at most $\frac{3}{4}$. The left hand side can be expanded into

$$1 - \Pr(Y^2 > X^1) - \Pr(X^3 > Y^2) + 1 - \Pr(Y^2 > X^2) - \Pr(X^4 > Y^2)$$

which by Lemma 7 is at least $\frac{3}{4}$.

7.2.3 $j = 4$

To show

$$\sum_{i=1}^{3} \Pr(X^i > Y^2 > X^{i+2}) \geq \frac{1}{2} \sum_{i=1}^{2} \Pr(X^i > Y^{5-i})$$

we can rearrange the LHS into $\Pr(X^1 > Y^2 > X^5) + \Pr(X^2 > Y^2 > X^4)$. By Lemma 7, this is at least $\frac{73}{64} - \Pr(X^4 > Y^2)$. By symmetry, $\Pr(X^4 > Y^2) \leq \Pr(Y^3 > X^2)$, so the RHS is at most equal to $1 - \frac{1}{2} \Pr(X^4 > Y^2)$. Rearranging, we need only show that

$$\frac{73}{64} - \frac{1}{2} \Pr(X^4 > Y^2) \geq 1$$

which can be proven by using Lemma 7 to show $\Pr(X^4 > Y^2) \leq \frac{3}{16}$

7.2.4 $j \geq 5$

When $j \geq 5$, we have

$$\sum_{i=1}^{2} \Pr(A^i \geq a_j) \geq \sum_{i=1}^{2} \Pr(A_i \geq a_5) \geq \sum_{i=1}^{4} \Pr(X^i > Y^2 > X^{2+i})$$

This sum can be rearranged to $\Pr(X^1 > Y^2 > X^6) + \Pr(X^2 > Y^2 > X^5)$. By Lemma 7 this is at least equal to $\frac{69}{64}$, and for any j,

$$\frac{69}{64} > 1 > \frac{1}{2} \sum_{i=1}^{2} \Pr(X^i > Y^{j+1-i})$$

7.3 $k = 3$

For $j \in \{1, 2\}$, the inequality (4) vanishes. Further, we have using Lemma 7 that

$$\Pr(X^1 > Y^3 > X^4) \geq \frac{17}{32}$$
$$\Pr(X^2 > Y^3 > X^5) \geq \frac{59}{128}$$
$$\Pr(X^3 > Y^3 > X^6) \geq \frac{91}{256}$$

7.3.1 $j = 3$

By Lemma 6, applied with $i = 1, j = k = 3$ we have

$$\Pr(X^1 > Y^3 > X^4) \geq \frac{1}{2}\Pr(X^1 > Y^3)$$

7.3.2 $j = 4$

Making use of Lemma 7,

$$\begin{aligned}\sum_{i=1}^{2}\Pr(X^i > Y^3 > X^{3+i}) &\geq \frac{167}{128} - \Pr(Y^3 > X^2)\\ &\geq 1 - \frac{1}{2}\Pr(Y^3 > X^2)\\ &\geq \frac{1}{2}\sum_{i=1}^{2}\Pr(X^i > Y^{5-i})\end{aligned}$$

given the fact that $\Pr(Y^3 > X^2) \leq \frac{5}{16}$.

7.3.3 $j = 5$

$$\begin{aligned}\sum_{i=1}^{3}\Pr(X^i > Y^3 > X^{3+i}) &\geq \frac{345}{256}\\ &> \frac{1}{2}(2 + \frac{1}{2})\\ &\geq \frac{1}{2}(\sum_{i=1}^{2}\Pr(X^i > Y^{6-i}) + \Pr(X^3 > Y^3))\end{aligned}$$

7.3.4 $j \geq 6$

$$\begin{aligned}\sum_{i=1}^{3}\Pr(A_i \geq a_j) &\geq \sum_{i=1}^{3}\Pr(A_i \geq a_6)\\ &\geq \sum_{i=1}^{4}\Pr(X^i > Y^3 > X^{3+i})\\ &= \sum_{i=1}^{3}\Pr(X^i > Y^3 > X^{8-i})\\ &\geq \frac{205}{128}\\ &> \frac{3}{2}\\ &\geq \frac{1}{2}\sum_{i=1}^{3}\Pr(X^i > Y^{j+1-i})\end{aligned}$$

7.4 $k = 4$

Again, (4) vanishes when $j < 4$. For higher values, we have by Lemma 7

$$\Pr(X^1 > Y^4 > X^5) \geq \frac{147}{256}$$
$$\Pr(X^2 > Y^4 > X^6) \geq \frac{143}{256}$$
$$\Pr(X^3 > Y^4 > X^7) \geq \frac{124}{256}$$
$$\Pr(X^4 > Y^4 > X^8) \geq \frac{99}{256}$$

This implies that for $4 \leq j \leq 7$,

$$\sum_{i=1}^{j-3} \Pr(X^i > Y^4 > X^{4+i}) \geq \frac{j-3}{2} \geq \frac{1}{2} \sum_{i=1}^{j-3} \Pr(X^i > Y^{j+1-i})$$

and consequently for $j \geq 8$,

$$\begin{aligned}\sum_{i=1}^{j-3} \Pr(X^i > Y^4 > X^{4+i}) &\geq \sum_{i=1}^{4} \Pr(X^i > Y^4 > X^{4+i}) \\ &\geq 2 \\ &> \frac{1}{2} \sum_{i=1}^{4} \Pr(X^i > Y^{j+1-i})\end{aligned}$$

7.5 $k \geq 5$

Consider $k \geq 5$. For $j < k$, both sides of (4) vanish, as we saw previously.

For values of j in the range $[k, 2k-2]$, it suffices to show that

$$\frac{\sum_{i=1}^{j+1-k} \Pr(X^i > Y^k > X^{i+k})}{j+1-k} \geq \frac{1}{2}$$

which implies

$$\sum_{i=1}^{j+1-k} \Pr(X^i > Y^k > X^{i+k}) \geq \frac{j+1-k}{2} \geq \frac{1}{2} \sum_{i=1}^{\min(j+1-k,k)} \Pr(X^i > Y^{j+1-i})$$

For $j \geq 2k-1$, it suffices also to show

$$\frac{\sum_{i=1}^{k} \Pr(X^i > Y^k > X^{i+k})}{k} \geq \frac{1}{2}$$

which implies

$$\sum_{i=1}^{j+1-k} \Pr(X^i > Y^k > X^{i+k}) \geq \sum_{i=1}^{k} \Pr(X^i > Y^k > X^{i+k}) \geq \frac{k}{2} \geq \frac{1}{2} \sum_{i=1}^{\min(j+1-k,k)} \Pr(X^i > Y^{j+1-i})$$

Combined, we must show that for $k \leq j \leq 2k-1$,

$$\frac{\sum_{i=1}^{j+1-k} \Pr(X^i > Y^k > X^{i+k})}{j+1-k} \geq \frac{1}{2}$$

that is, for any r such that $1 \leq r \leq k$,

$$\frac{\sum_{i=1}^{r} \Pr(X^i > Y^k > X^{i+k})}{r} \geq \frac{1}{2}$$

Since we can write the probability $\Pr(X^i > Y^k > X^{i+k})$ as $1 - \Pr(Y^k > X^i) - \Pr(Y^k < X^{k+i})$, and $1 \leq k < k+i$, we can lower bound the probability using Lemma 7 by $\Pr(X^i > Y^k > X^{i+k})$ in the binomial case. Therefore, for the rest of the section, we will assume that all probabilities refer to probabilities in the binomial case.

To establish our desired inequality, we first establish the following lemma:

Lemma 8. *The series* $\Pr(X^i > Y^k > X^{i+k})$ *is concave for* $1 \leq i \leq k$.

Proof. In the binomial case, for $1 \leq i \leq k-1$,

$$\begin{aligned}\Pr(X^{i+1} > Y^k > X^{i+k+1}) - \Pr(X^i > Y^k > X^{i+k}) &= \Pr(X^{i+k} > Y^k > Y^{i+k+1}) - \Pr(X^i > Y^k > X^{i+1}) \\ &= \frac{\binom{i+2k-1}{k-1}}{2^{i+2k}} - \frac{\binom{i+k-1}{k-1}}{2^{i+k}}\end{aligned}$$

Consider the difference

$$\begin{aligned}&(\Pr(X^{i+2} > Y^k > X^{i+k+2}) - \Pr(X^{i+1} > Y^k > X^{i+k+1})) \\ &\qquad - (\Pr(X^{i+1} > Y^k > X^{i+k+1}) - \Pr(X^i > Y^k > X^{i+k}))\end{aligned}$$

where $1 \leq i \leq k-2$. We have that this evaluates to

$$\begin{aligned}\left(\frac{\binom{i+2k}{k-1}}{2^{i+2k+1}} - \frac{\binom{i+k}{k-1}}{2^{i+k+1}}\right) - \left(\frac{\binom{i+2k-1}{k-1}}{2^{i+2k}} - \frac{\binom{i+k-1}{k-1}}{2^{i+k}}\right) &= \frac{i+2k}{2(i+k+1)}\frac{\binom{i+2k-1}{k-1}}{2^{i+2k}} - \frac{\binom{i+2k-1}{k-1}}{2^{i+2k}} \\ &\quad + \frac{\binom{i+k-1}{k-1}}{2^{i+k}} - \frac{i+k}{2i+2}\frac{\binom{i+k-1}{k-1}}{2^{i+k}} \\ &\leq \frac{\binom{i+2k-1}{k-1}}{2^{i+2k}} - \frac{\binom{i+2k-1}{k-1}}{2^{i+2k}} + \frac{\binom{i+k-1}{k-1}}{2^{i+k}} - \frac{\binom{i+k-1}{k-1}}{2^{i+k}} \\ &= 0\end{aligned}$$

Thus the series is concave for $1 \leq i \leq k$. □

Lemma 9. *For a concave series* $x_1, \ldots, x_n$,

$$\frac{\sum_{i=1}^{n} x_i}{n} \geq \frac{x_1 + x_n}{2}$$

Proof. We have that by the concavity, for any $i \leq n-2$,

$$(x_{i+2} - x_{i+1}) - (x_{i+1} - x_i) \leq 0$$

implying that

$$\sum_{i=1}^{n-2} (x_i + x_{i+2}) \leq \sum_{i=1}^{n-2} 2x_{i+1} \implies x_1 + x_n \leq x_2 + x_{n-1}$$

By induction on this argument, we have that $x_1 + x_n \leq x_{1+i} + x_{n-i}$ for $i \leq n-1$. Therefore,

$$2\sum_{i=1}^{n} x_i = \sum_{i=0}^{n-1} (x_{1+i} + x_{n-i}) \geq n(x_1 + x_n) \implies \frac{x_1 + x_n}{2} \leq \frac{\sum_{i=1}^{n} x_i}{n}$$

□

We have therefore that for $1 \leq r \leq k$,

$$\frac{\sum_{i=1}^{r} \Pr(X^i > Y^k > X^{i+k})}{r} \geq \frac{\Pr(X^1 > Y^k > X^{1+k}) + \Pr(X^r > Y^k > X^{r+k})}{2}$$

If $\Pr(X^r > Y^k > X^{r+k}) \geq \Pr(X^k > Y^k > X^{2k})$, then we have that

$$\frac{\Pr(X^1 > Y^k > X^{1+k}) + \Pr(X^r > Y^k > X^{r+k})}{2} \geq \frac{\Pr(X^1 > Y^k > X^{1+k}) + \Pr(X^k > Y^k > X^{2k})}{2}$$

so it would suffice to show that

$$\frac{\Pr(X^1 > Y^k > X^{1+k}) + \Pr(X^k > Y^k > X^{2k})}{2} \geq \frac{1}{2}.$$

If $\Pr(X^r > Y^k > X^{r+k}) \leq \Pr(X^k > Y^k > X^{2k})$, then by concavity we must have that $\Pr(X^r > Y^k > X^{r+k}) \geq \Pr(X^1 > Y^k > X^{1+k})$. Therefore,

$$\frac{\Pr(X^1 > Y^k > X^{1+k}) + \Pr(X^r > Y^k > X^{r+k})}{2} \geq \Pr(X^1 > Y^k > X^{1+k})$$

so it would suffice to show that

$$\Pr(X^1 > Y^k > X^{1+k}) \geq \frac{1}{2}$$

Therefore, if we can show that

$$\Pr(X^1 > Y^k > X^{1+k}) \geq \frac{1}{2}$$
$$\frac{\Pr(X^1 > Y^k > X^{1+k}) + \Pr(X^k > Y^k > X^{2k})}{2} \geq \frac{1}{2}$$

for all $k \geq 5$, then we would have the $\frac{1}{2}$-competitiveness for $k \geq 5$, completing our proof.
To show the first inequality, we have

$$\Pr(X^1 > Y^k > X^{k+1}) = 1 - \Pr(Y^k > X^1) - \Pr(Y^k < X^{k+1}) = 1 - \frac{1}{2^k} - \left(\frac{1}{2} - \frac{\binom{2k}{k}}{2^{2k+1}}\right) = \frac{1}{2} + \frac{\binom{2k}{k}}{2^{2k+1}} - \frac{1}{2^k}$$

Now we know that

$$\binom{2k}{k} \geq \frac{4^k}{2k+1}.$$

It follows that the above is at least

$$\frac{1}{2} + \frac{1}{4k+2} - \frac{1}{2^k}$$

which is above $\frac{1}{2}$ for $k \geq 5$. To show the second inequality we need, we have

$$\begin{aligned}\Pr(X^1 > Y^k > X^{k+1}) + \Pr(X^k > Y^k > X^{2k}) &= 1 - \Pr(Y^k > X^1) - \Pr(X^{k+1} > Y^k) \\ &\quad + 1 - \Pr(Y^k > X^k) - \Pr(X^{2k} > Y^k) \\ &= 1 + \frac{\binom{2k}{k}}{2^{2k+1}} - \frac{1}{2^k} - \Pr(X^{2k} > Y^k)\end{aligned}$$

We have the following lower bounds for the above for $k \in [5, 15]$:

k	LB
5	1.0020
6	1.0254
7	1.0392
8	1.0477
9	1.0530
10	1.0564
11	1.0585
12	1.0598
13	1.0605
14	1.0609
15	1.0608

By the Chernoff-Hoeffding bound,

$$\Pr(X^{2k} > Y^k) = \Pr(\geq 2kX's \in \{a_1, \ldots, a_{3k-1}\}) \leq \left(\left(\frac{3}{4}\right)^{\frac{2}{3}}\left(\frac{3}{2}\right)^{\frac{1}{3}}\right)^{3k-1} < 0.94495^{3k-1}$$

At $k = 16$, we have that

$$1 + \frac{\binom{2k}{k}}{2^{2k+1}} - \frac{1}{2^k} - 0.94495^{3k-1} \geq 1.0001$$

and for $k \geq 16$, when k increases by 1,

$$\frac{\binom{2k}{k}}{2^{2k+1}}$$

is multiplied by at least $\frac{(2k+2)(2k+1)}{4(k+1)^2} \geq \frac{33}{34}$ while

$$\frac{1}{2^k} + 0.94495^{3k-1}$$

is multiplied by at most $0.94495^3 < \frac{29}{34}$. Thus for $k \geq 5$,

$$\left(1 + \frac{\binom{2k}{k}}{2^{2k+1}} - \frac{1}{2^k} - \Pr(X^{2k} > Y^k)\right) \geq 1 \implies \frac{\Pr(X^1 > Y^k > X^{1+k}) + \Pr(X^k > Y^k > X^{2k})}{2} \geq \frac{1}{2}$$

As was shown above, this establishes $\frac{1}{2}$-competitiveness for $k \geq 5$.

7.6 Tightness

We also prove that the lower bound of $\frac{1}{2}$ is tight for all k. Consider the following random variables:

$$X_1, \dots, X_N \sim \text{Unif}[0,1], \quad X_{N+1} \sim \begin{cases} 0 & \text{w.p. } \frac{N-1}{N} \\ N^2 & \text{w.p. } \frac{1}{N} \end{cases}$$

We have that as N goes to infinity, and k remains fixed, the prophet's expectation is asymptotically $\frac{1}{N} \cdot N^2 = N$. Meanwhile, the algorithm's expectation is asymptotically N^2 times the probability $X_{N+1} = N^2$ and the algorithm selects it. If we define Y^i and X^i as the ith order statistic out of $Y_1, \dots, Y_N$ and $X_1, \dots, X_N$, respectively, then

$$\mathbb{E}[\text{ALG}] = \frac{1}{N} \cdot N^2 \cdot \left(\frac{1}{N} \Pr(Y^{k-1} > X^k) + \frac{N-1}{N} \Pr(Y^k > X^k) \right) \leq N^2 \cdot \frac{1}{N} \cdot \left(\frac{1}{N} + \frac{N-1}{2N} \right)$$

which asymptotically is equal to $\frac{N}{2}$. This establishes that the single-sample threshold k-select algorithm achieves, at best $\frac{1}{2}$ the prophet's expectation.

8 Open Questions

Our work does leave several open questions. First, is there a simple single-threshold algorithm for the prophet inequality with k selections using just a single sample that is $1-O\left(\frac{\log k}{\sqrt{k}}\right)$-competitive? We cannot expect better, since even with full knowledge of the distributions, we cannot do better than $1 - O\left(\frac{\log k}{\sqrt{k}}\right)$ [9].

Another direction of further research concerns the worst-case random variables for the offline IIF case. We already know that $\frac{\mathbb{E}[\text{OPTIMAL OFFLINE IIF ALG}]}{\mathbb{E}[\text{PROPHET}]} \geq \frac{1}{2}$ for any random variables, and it was previously posed as an open question if this $1/2$ could be improved [8]. Based on our work, we are certain that this ratio cannot be larger than $25/27$, however it's still unclear if this $\frac{1}{2}$ is tight or not.

Another open question is if the k-select linear program can be used to generate an online IIF and TIF algorithm for k selections that achieves a competitive ratio close to 1 as k goes to infinity. Perhaps it is possible to do this through some sort of 'smoothing' to turn a solution into a fair solution, similar to [8], and this may involve creating a relaxation of the linear program outlined.

Finally, it is unknown whether the bound of $\frac{1}{6}$ for the online IIF and TIF algorithm using two samples is tight, or if it can be improved. It is also unknown whether there is a constant competitive IIF and TIF algorithm using just a single sample.

9 Acknowledgements

We would like to express our thanks to our mentor, Pranav Nuti, for all of his guidance and support with this project. He went above and beyond as a mentor and helped all of us become familiar with the process of math research and directed us towards interesting and fruitful problems.

We are also grateful to the SURIM program for providing the platform and resources to conduct this research. Special thanks go to Dr. Lernik Asserian, the program organizer, for her unwavering dedication to the program.

Finally we would like to acknowledge the VPUE grant we received that made this research possible.

References

[1] Ulrich Krengel and Louis Sucheston. "On semiamarts, amarts, and processes with finite value". In: *Probability on Banach spaces* 4 (1978), pp. 197–266.

[2] T. P. Hill and R. P. Kertz. "Ratio comparisons of supremum and stop rule expectations". In: *Z. Wahrscheinlichkeitstheorie verw Gebiete* 56 (1981), pp. 283–285.

[3] Douglas P Kennedy. "Optimal stopping of independent random variables and maximizing prophets". In: *The Annals of Probability* (1985), pp. 566–571.

[4] Saeed Alaei. "Bayesian Combinatorial Auctions: Expanding Single Buyer Mechanisms to Many Buyers". In: *SIAM Journal on Computing* 43.2 (2014), pp. 930–972. eprint: `https://doi.org/10.1137/120878422`.

[5] Pablo D. Azar, Robert Kleinberg, and S. Matthew Weinberg. "Prophet Inequalities with Limited Information". In: *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms*. SODA '14. Portland, Oregon: Society for Industrial and Applied Mathematics, 2014, pp. 1358–1377.

[6] Tomer Ezra et al. "Online Stochastic Max-Weight Matching: Prophet Inequality for Vertex and Edge Arrival Models". In: *Proceedings of the 21st ACM Conference on Economics and Computation*. EC '20. Virtual Event, Hungary: Association for Computing Machinery, 2020, pp. 769–787.

[7] Aviad Rubinstein, Jack Wang, and Matthew Weinberg. "Optimal Single-Choice Prophet Inequalities from Samples". In: *11th Innovations in Theoretical Computer Science Conference, ITCS 2020* 60 (Nov. 2020).

[8] Makis Arsenis and Robert Kleinberg. "Individual Fairness in Prophet Inequalities". In: *Proceedings of the 23rd ACM Conference on Economics and Computation*. EC '22. Boulder, CO, USA: Association for Computing Machinery, 2022, p. 245.

[9] Jiashuo Jiang, Will Ma, and Jiawei Zhang. "Tightness without Counterexamples: A New Approach and New Results for Prophet Inequalities". In: *Proceedings of the 24th ACM Conference on Economics and Computation*. EC '23. London, United Kingdom: Association for Computing Machinery, 2023, p. 909.

[10] Kanstantsin Pashkovich and Alice Sayutina. "Single Sample Prophet Inequality for Uniform Matroids of Rank 2". In: (June 2023), pp. 1–14.

[11] Jiashuo Jiang, Will Ma, and Jiawei Zhang. "Tight Guarantees for Multi-unit Prophet Inequalities and Online Stochastic Knapsack". In: *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pp. 1221–1246. eprint: `https://epubs.siam.org/doi/pdf/10.1137/1.9781611977073.51`.

[12] Pranav Nuti and Jan Vondrák. "Secretary Problems: The Power of a Single Sample". In: *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pp. 2015–2029. eprint: `https://epubs.siam.org/doi/pdf/10.1137/1.9781611977554.ch77`.