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

Quinn McIntyre Ezra Steinberg Peter Westbrook Ethan Zhang

Stanford University, Department of Mathematics

## Introduction: Online Decision Making

As an employer hiring for a position, you have applications from multiple candidates. You have a rough estimate of their abilities from the applications, but you will not know for sure until you interview them. After interviewing each one, you must decide immediately: hire or pass. You can only hire one person, and once you pass on a candidate, there's no going back. How do you strategize considering the uncertainty of future candidates?

This problem can be formalized with a sequence of $n$ random variables $((X_i)_{i=1}^n)$, where we (the employer) must select one of the realizations of the random variables (candidates).

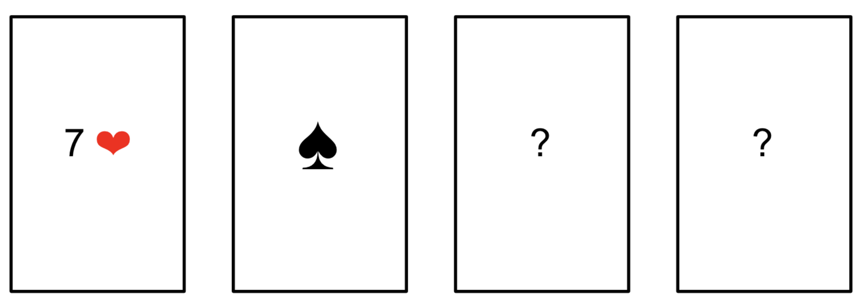


Figure 1. Try it yourself with our card demo!

It has been known for nearly 50 years [1] that there exists an algorithm such that

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

where the prophet selects the maximum while knowing the outcome of all of the random variables. This is known as the prophet inequality.

## Background: Types of Algorithms

There are several types of algorithms that we study:

- An **online** algorithm is an algorithm for a situation in which information is revealed gradually instead of all at once. The prophet inequality is about an online algorithm.
- An algorithm for a situation in which we we know all the information beforehand is called **offline**.
- Instead of the standard assumption that one knows the distribution of the random variables, if we assume instead that we only have access to samples from the random variables, then we may use **sample based** algorithms.

## Background: Fairness

Arsenis and Kleinberg [3] introduced two ideas of fairness for these decision-making algorithms: identity-independent fairness (IIF) and time-independent fairness (TIF):

- IIF ensures that two candidates of the same value will have the same chance of being selected. TIF ensures that the candidates' order doesn't affect each person's chance of being selected.
- Consider what happens if we know $X_1$ will be a certain value $x$. Based on the distributions of the other variables, as well as how our algorithm works, we can say that $X_1$ will be selected with certain probability.
- For an IIF algorithm, if we now condition on any variable $X_i$ being $x$, the chance that we pick it is the same probability as it was for $X_1$.
- For a TIF algorithm, each variable's chance of being selected (conditional on their realized value) is independent of the order of the random variables. This is helpful to consider in the hiring example; a TIF algorithm would give no preferential treatment to a candidate based on their position in the order.

## Background: $k$-Select

The $k$-select prophet inequality is a generalization of the standard prophet inequality where $k$ random variables can be selected. In this setting, the objective becomes maximizing the sum of the $k$ variables selected.

## Exploring Offline IIF Algorithms

In the offline case, we see the realizations of all the random variables and then make our decision, again with the goal of maximizing the expected value of our pick. Without any fairness constraints, we could clearly always choose the largest option and obtain $\mathbb{E}[\text{ALG}] = \mathbb{E}[\text{PROPHET}]$. With the IIF constraint, we have two new results:

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.

The proof of this statement considers the random variables:

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

When $X_1$ takes small non-zero values (i.e. much closer to 1 than $1+\epsilon$), then it's unlikely to be the maximum and its chance of being chosen is small. When $X_2$ takes the same value, although it's almost certain to be the maximum, the IIF condition prohibits us from accepting it most of the time. This means we can only make a pick about half of the time.

Even without this maximum-only restriction, an offline IIF algorithm cannot always match the prophet.

We found a distribution where the best possible algorithm is only $\frac{25}{27}$-competitive.

## Double Sample Algorithm that is IIF and TIF

Arsenis and Kleinberg discovered a sample based IIF + TIF algorithm relying on two samples.

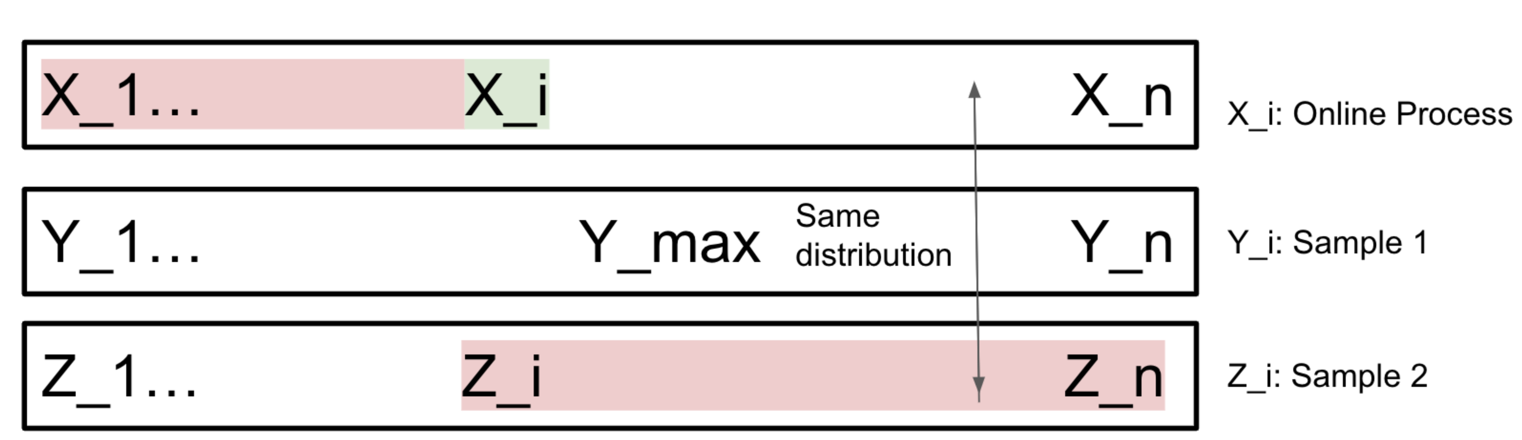


Figure 2. Visual representation of the algorithm. We accept $X_i$ if and only if it larger than $Y_{\max}$ and everything that is colored red is less than $Y_{\max}$. There is at most one $X_i$ that satisfies this property.

Previously, the competitive ratio of this algorithm was shown to be at least $\frac{1}{9}$. However, this bound is not tight, and with the following result

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

we improve the bound to $\frac{1}{6}$. This shows that online sample based IIF and TIF algorithms can achieve a competitive ratio of at least $\frac{1}{6}$.

## Single-Sample Algorithm for Making $k$-Selections

Consider the following algorithm for $k$-select:

- Let the samples we have access to be $Y_1, \ldots, Y_n \sim X_1, \ldots, X_n$.
- Set a threshold equal to $Y^k$, the $k^{\text{th}}$ largest realization out of $\{Y_1, \ldots, Y_n\}$.
- Select each of $X_1, \ldots, X_n$ that we see that is higher than the threshold until we have made $k$ selections or have decided on each of $X_1, \ldots, X_n$.

We proved that this algorithm is $\frac{1}{2}$-competitive, that is:

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

This had been shown in the literature for $k = 1$ [2] and $k = 2$ [4], but we were able to generalize this to all natural numbers $k$. Moreover, we showed that this lower bound of $\frac{1}{2}$ is tight. While there exist more complicated sample-based algorithms that can achieve asymptotically better competitiveness for large values of $k$, this is still meaningful as it establishes that a very simple algorithm is able to achieve the $\frac{1}{2}$ lower bound.

## $k$-Select Through Linear Programming

Directly finding an algorithm that maintains fairness whilst competitively selecting $k$ variables is challenging, so an alternate approach is to find an equivalent linear program such that a solution to the linear program gives an online implementable algorithm.

$$0 \leq h_{i,0}(x) \leq \Pr[X_i = x]\left(1 - \sum_x \sum_{i'<i} h_{i,0}(x)\right)$$

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

$$h_{i,j}(x) = 0 \text{ for } j \geq k \text{ or } j \leq -1 \text{ or } i \leq 0.$$

This linear program may look nothing like an algorithm, but we proved that we can construct an algorithm such that the variables $h_{i,j}(x)$ have a probabilistic interpretation:

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

We conjecture that we can use this linear program to develop an IIF + TIF $k$-select algorithm.

## Acknowledgements

We would like to express our thanks to our mentor, Pranav Nuti and to the SURIM program and its organizer Lernik Asserian. Under their support, this research was made possible.

## References

[1] Ulrich Krengel and Louis Sucheston. "Semiamarts and finite values". In: *Bulletin of The American Mathematical Society* 83 (July 1977), pp. 745–747.

[2] 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).

[3] 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.

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