Prophet Inequality from Samples*

Tomer Ezra†

Abstract

We study a variant of the single-choice prophet inequality problem where the decision-maker does not know the underlying distributions and has only access to samples from the distributions. Rubinstein et al. [2020] showed that the optimal competitive ratio of $\frac{1}{2}$ can surprisingly be obtained by observing a single sample from each distribution. In this paper, we give an alternative proof of this result.

1 Introduction

The prophet inequality [Krengel and Sucheston, 1977, 1978] captures scenarios where a decision-maker needs to select a reward among a sequence of rewards in an online manner. The rewards are stochastic and are distributed according to known distributions. Each time, one reward is revealed to the decision-maker which needs to decide immediately and irrevocably whether to accept it (which ends the process) or discard it (hoping to obtain a more valuable reward in the future). The performance of the decision-maker is measured against the offline optimum (or a "prophet" that can see into the future, and always selects the maximum reward). The prophet inequality framework (and its variants) captures many online settings such as dynamic pricing, matching markets, and resource allocation.

Traditionally, it is assumed that the distributions of the random variables are known to the decision-maker. This assumption, while convenient for theoretical analysis, raises practical concerns: Where do these priors come from, and how realistic is it to assume that they are known with precision? In many real-world scenarios, exact knowledge of the underlying distributions is not available, and instead, decision-makers must rely on samples drawn from these distributions.

Azar et al. [2014] proposed addressing this challenge by considering the prophet inequality in a setting where only samples, rather than complete knowledge of the distributions, are available. This approach by Azar et al. [2014] sparked a line of research that explores the prophet inequality framework under various conditions where the online algorithm has access only to samples [Azar et al., 2014, Rubinstein et al., 2020, Dütting et al., 2021, Caramanis et al., 2022, Correa et al., 2022, Gravin et al., 2022, Kaplan et al., 2022, Cristi and Ziliotto, 2024, Correa et al., 2024, Fu et al., 2024].

Among this line of works, Rubinstein et al. [2020] considered the classical setting of single-choice prophet inequality from samples and showed that it is possible to achieve a competitive ratio of $\frac{1}{2}$ even when only a single sample is provided from each distribution.

*This paper is based on a result from [Ezra, 2026].

†Tel Aviv University, Tel Aviv, Israel. Email: tomerezra@tauex.tau.ac.il

Our contribution. Our result is an alternative proof of the main result of Rubinstein et al. [2020] where we show that for the single-sample case, the threshold algorithm that sets the maximum sample as a threshold achieves a competitive ratio of $\frac{1}{2}$. We show this result by establishing that for each value x, the probability that the algorithm selects at least x is at least half of the probability that the maximum is at least x. Thus, this algorithm not only achieves half of the expectation of the prophet, but also, the distribution of this algorithm $\frac{1}{2}$-stochastically dominates the distribution of the prophet. Our proof relies on establishing two lower bounds that compare the probability of selecting at least x to the probability that the maximum is at least x. We then integrate over our constructed lower bounds and derive the approximation of half.

1.1 Related Work

The study of prophet inequalities with limited prior information was initiated by Azar et al. [2014], who demonstrated that a single sample from each distribution is sufficient to achieve an asymptotically optimal competitive ratio for the k-capacity prophet inequality. Rubinstein et al. [2020] focused on the single-choice prophet inequality and showed that the optimal competitive ratio of $\frac{1}{2}$ can be attained using just one sample per distribution. Additionally, they showed that when the rewards' values are drawn from identical distributions, $O(\frac{n}{\epsilon^6})$ samples suffice to achieve a loss of only an ϵ-fraction compared to the best online algorithm with full distributional information. Dütting et al. [2021] explored the single-sample prophet inequality under a matching feasibility constraint and established a competitive ratio of $\frac{1}{16}$. Caramanis et al. [2022] extended the single-sample prophet inequality framework to various settings, including different types of matroids and matching problems under diverse arrival models. Gravin et al. [2022] considered the prophet inequality with fewer than one sample per distribution, where each distribution is sampled with probability p, and demonstrated that the optimal competitive ratio achievable in this setting is $\frac{p}{p+1}$. Correa et al. [2024] improved the sample complexity results of Rubinstein et al. [2020] for the i.i.d. prophet inequality, showing that $\frac{n}{\epsilon}$ samples are sufficient to achieve a loss of an ϵ-fraction. Recently, Cristi and Ziliotto [2024] introduced a unified framework showing that $O\left(\frac{1}{\epsilon^5}\right)$ of samples from all distributions is sufficient to achieve the optimal competitive ratio (up to a factor of ϵ) for three variants of the prophet inequality: prophet secretary (where the rewards arrive in random order), free order (where the decision-maker can choose the order of the observed rewards), and the i.i.d. case. Fu et al. [2024] examined the prophet inequality under matroid feasibility constraints and proved that a polylogarithmic number of samples per distribution is enough to achieve a constant competitive ratio.

2 Preliminaries

We consider a setting with n boxes. Every box i contains some value v_i drawn from an underlying independent distribution F_i. We denote the product distribution $F = F_1 \times \ldots \times F_n$. The values are revealed sequentially in an online fashion, and the decision-maker needs to select one of them in an immediate and irrevocable manner without observing future values. We consider algorithms that do not know the distributions F_i, but instead will have access to samples from them. We assume that prior to the online decision-making process where v_i are revealed to the algorithm, one sample s_i from each distribution F_i is given to the algorithm, where we denote by $S = (S_1, \ldots, S_n)$. We denote an online algorithm by ALG, and given a sequence of samples $S = (S_1, \ldots, S_n)$ and online values $v = (v_1, \ldots, v_n)$, we denote by $\mathrm{ALG}(S, v)$ the (possibly random) value chosen by the algorithm ALG on v and S.

We measure the performance of an algorithm ALG by the competitive ratio measure, which is

$$\alpha(\mathrm{ALG}) := \inf_F \frac{\mathbb{E}_{S\sim F}\left[\mathbb{E}_{v\sim F}[\mathrm{ALG}(S,v)]\right]}{\mathbb{E}_{v\sim F}[\max_i v_i]},$$

where the expectation in the numerator is also taken over the randomness of the algorithm.

We assume for simplicity of the analysis that all distributions are continuous, therefore there are no ties between different values and samples. This assumption is made without loss of generality by employing standard techniques, such as using randomized tie-breaking when values or samples are identical (see e.g., [Ezra et al., 2018, Rubinstein et al., 2020]).

Lastly, we compare different distributions through the notion of approximate stochastic domination.

Definition 2.1 (Stochastic Domination)**.** *For two distributions over $\mathbb{R}_{\geq 0}$ and a parameter $\gamma \in [0, 1]$, we say that D_1 γ-stochastically dominates the distribution D_2 if for every $x \in \mathbb{R}_{\geq 0}$ it holds that*

$$\Pr_{X\sim D_1}[X \geq x] \geq \gamma \cdot \Pr_{X\sim D_2}[X \geq x].$$

In particular, our proof relies on the fact that if the distribution of the chosen value of an algorithm γ-stochastically dominates the distribution of the prophet, it implies a competitive ratio of γ since

$$\mathbb{E}[\mathrm{ALG}(v)] = \int_0^\infty \Pr[\mathrm{ALG}(v) \geq x]dx \geq \int_0^\infty \gamma \cdot \Pr[\max_i v_i \geq x]dx = \gamma \cdot \mathbb{E}[\max_i v_i]. \tag{1}$$

3 Single-Sample Prophet Inequality

In this section, we provide our result of an alternative proof for the result of Rubinstein et al. [2020], showing that an optimal competitive ratio of $1/2$ can be obtained when the online algorithm is given only a single sample from each distribution. Consider the algorithm ALG that sets the maximum sample as a threshold. Let F_i be the CDF of box i. Let $F = \prod_i F_i$, be the product of the CDFs (which is the CDF of the maximum), and let f be the corresponding PDF of the maximum. We show that ALG $\frac{1}{2}$-stochastically dominates the prophet, i.e., for every value $x \in \mathbb{R}_{\geq 0}$, it holds that

$$\Pr[\mathrm{ALG}(S, v) \geq x] \geq \frac{1}{2} \Pr[\max_i v_i \geq x]. \tag{2}$$

Equation (2) together with Equation (1) imply that $\alpha(\mathrm{ALG}) \geq \frac{1}{2}$.

Proof of Inequality (2). We first observe that $\Pr[\max_i v_i \geq x] = 1 - F(x)$.

On the other hand, if the maximum sample, which we denote by s, satisfies that $s \leq x$, then:

$$
\begin{aligned}
\Pr[\mathrm{ALG}(S,v) \geq x \mid \max(S) = s] &= \sum_{i=1}^{n}(1-F_i(x)) \underbrace{\left(\prod_{j<i} F_j(s)\right)}_{\text{Reaching } i} \\
&= \sum_{i=1}^{n}(1-F_i(x)) \left(\prod_{j<i} F_j(x) \cdot \frac{F_j(s)}{F_j(x)}\right) \\
&\geq \frac{F(s)}{F(x)} \sum_{i=1}^{n}(1-F_i(x)) \left(\prod_{j<i} F_j(x)\right) \\
&= \frac{F(s)}{F(x)}(1-F(x)), \qquad (3)
\end{aligned}
$$

where the inequality is since we multiplied each term by $\prod_{j\geq i} \frac{F_j(s)}{F_j(x)}$ which is at most 1 since $s \leq x$.

If the maximum sample satisfies that $s > x$, then:

$$
\Pr[\mathrm{ALG}(S,v) \geq x \mid \max(S) = s] = \sum_{i=1}^{n}(1-F_i(s)) \left(\prod_{j<i} F_j(s)\right) = 1 - F(s). \qquad (4)
$$

Overall,

$$
\begin{aligned}
\Pr[\mathrm{ALG}(S,v) \geq x] &= \int_0^\infty f(s) \cdot \Pr[\mathrm{ALG}(S,v) \geq x \mid \max(S) = s] ds \\
&\geq \int_0^x \frac{F(s)\cdot(1-F(x))}{F(x)} f(s) ds + \int_x^\infty (1-F(s)) f(s) ds \\
&= \frac{1-F(x)}{F(x)} \left[\frac{1}{2}F(s)^2\right]_0^x + \left[F(s) - \frac{1}{2}F(s)^2\right]_x^\infty \\
&= \frac{1-F(x)}{F(x)} \cdot \frac{F(x)^2}{2} + 1 - \frac{1}{2} - F(x) + \frac{1}{2}F(x)^2 \\
&= \frac{1-F(x)}{2} = \frac{\Pr[\max_i v_i \geq x]}{2},
\end{aligned}
$$

where the first equality is by integrating over the maximum sample; the first inequality is by partitioning to values of s that are smaller and larger than x and by Equations (3) and (4).

This concludes the proof of Equation (2). $\square$

References

P. D. Azar, R. Kleinberg, and S. M. Weinberg. Prophet inequalities with limited information. In *Proceedings of the twenty-fifth annual ACM-SIAM symposium on Discrete algorithms*, pages 1358–1377. SIAM, 2014.

C. Caramanis, P. Dütting, M. Faw, F. Fusco, P. Lazos, S. Leonardi, O. Papadigenopoulos, E. Pountourakis, and R. Reiffenhäuser. Single-sample prophet inequalities via greedy-ordered selection. In *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1298–1325. SIAM, 2022.

J. Correa, A. Cristi, B. Epstein, and J. Soto. The two-sided game of googol. *Journal of Machine Learning Research*, 23(113):1–37, 2022.

J. Correa, A. Cristi, B. Epstein, and J. A. Soto. Sample-driven optimal stopping: From the secretary problem to the iid prophet inequality. *Mathematics of Operations Research*, 49(1):441–475, 2024.

A. Cristi and B. Ziliotto. Prophet inequalities require only a constant number of samples. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pages 491–502, 2024.

P. Dütting, F. Fusco, P. Lazos, S. Leonardi, and R. Reiffenhäuser. Prophet inequalities for matching with a single sample. *arXiv preprint arXiv:2104.02050*, 2021.

T. Ezra. Prophet inequality from samples: Is the more the merrier? In *Proceedings of the 2026 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2099–2112. Society for Industrial and Applied Mathematics, 2026.

T. Ezra, M. Feldman, and I. Nehama. Prophets and secretaries with overbooking. In *Proceedings of the 2018 ACM Conference on Economics and Computation*, pages 319–320, 2018.

H. Fu, P. Lu, Z. G. Tang, H. Wu, J. Wu, and Q. Zhang. Sample-based matroid prophet inequalities. In *Proceedings of the 25th ACM Conference on Economics and Computation*, pages 781–781, 2024.

N. Gravin, H. Li, and Z. G. Tang. Optimal prophet inequality with less than one sample. In *International Conference on Web and Internet Economics*, pages 115–131. Springer, 2022.

H. Kaplan, D. Naori, and D. Raz. Online weighted matching with a sample. In *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1247–1272. SIAM, 2022.

U. Krengel and L. Sucheston. Semiamarts and finite values. *Bulletin of the American Mathematical Society*, 83(4):745–747, 1977.

U. Krengel and L. Sucheston. On semiamarts, amarts, and processes with finite value. *Probability on Banach spaces*, 4:197–266, 1978.

A. Rubinstein, J. Z. Wang, and S. M. Weinberg. Optimal single-choice prophet inequalities from samples. In *11th Innovations in Theoretical Computer Science Conference*, 2020.