Prophet Inequalities for Cost Minimization

Vasilis Livanos  Ruta Mehta

September 19, 2022

Abstract

Prophet inequalities for rewards maximization are fundamental results from optimal stopping theory with several applications to mechanism design and online optimization. We study the cost minimization counterpart of the classical prophet inequality due to Krengel, Sucheston, and Garling [KS77], where one is facing a sequence of costs $X_1, X_2, \ldots, X_n$ in an online manner and must “stop” at some point and take the last cost seen. Given that the $X_i$’s are independent random variables drawn from known distributions, the goal is to devise a stopping strategy $S$ (online algorithm) that minimizes the expected cost. The best cost possible is $\mathbb{E}[\min_i X_i]$ (offline optimum), achievable only by a prophet who can see the realizations of all $X_i$’s upfront. We say that strategy $S$ is an $\alpha$-approximation to the prophet ($\alpha \geq 1$) if $\mathbb{E}[S] \leq \alpha \cdot \mathbb{E}[\min_i X_i]$.

We first observe that if the $X_i$’s are not identically distributed, then no strategy can achieve a bounded approximation, no matter if the arrival order is adversarial or random, even when restricted to $n = 2$ and distributions with support size at most two. This leads us to consider the case where the $X_i$’s are independent and identically distributed (I.I.D.). For the I.I.D. case, we give a complete characterization of the optimal stopping strategy. We show that it achieves a (distribution-dependent) constant-factor approximation to the prophet’s cost for almost all distributions and that this constant is tight. In particular, for distributions for which the integral of the hazard rate is a polynomial $H(x) = \sum_{i=1}^{k} a_i x^{d_i}$, where $d_1 < \cdots < d_k$, the approximation factor is $\lambda(d_1)$, a decreasing function of $d_1$, and is the best possible for $H(x) = x^{d_1}$. Furthermore, when the hazard rate is monotonically increasing (i.e. the distribution is MHR), we show that this constant is at most 2, and this again is the best possible for the MHR distributions.

For the classical prophet inequality for reward maximization, single-threshold strategies have been powerful enough to achieve the best possible approximation factor. Motivated by this, we analyze single-threshold strategies for the cost prophet inequality problem. We design a threshold that achieves a $O(\text{polylog } n)$-factor approximation, where the exponent in the logarithmic factor is a distribution-dependent constant, and we show a matching lower bound.

We believe that our results may be of independent interest for analyzing approximately optimal (posted price-style) mechanisms for procuring items.

Keywords: prophet inequalities, cost minimization, mhr distributions, online algorithms

---

1Dept. of Computer Science, University of Illinois at Urbana-Champaign, IL 61801. {livanos3, rutameht}@illinois.edu

2For example, a (desperate) house buyer trying to decide when to buy in a sellers’ market, where houses are selling fast and they are trying to minimize the price they pay.

3The $n = 1$ case is trivial since both the algorithm and the prophet have to take $X_1$ and thereby both incur cost $E[X_1]$. Thus any algorithm is optimal.

4The hazard rate, also known as failure rate, of a distribution with probability density function $f$ and cumulative distribution function $F$ is $h(x) = \frac{f(x)}{1-F(x)}$. 
1 Introduction

The classical prophet inequality due to Krengel, Sucheston, and Garling (1978) [KS77] concerns the setting where one is presented with \( \text{take-it-or-leave-it} \) rewards \( X_1, \ldots, X_n \) in an online manner, drawn independently from known distributions, and can “stop” at any point and collect the last reward seen. Given that the distributions are known, the inequality ensures the existence of a stopping strategy \( S \) (online algorithm) for any arrival order of the random variables, with expected reward at least half that of a prophet who can see the realizations of all the \( X_i \)’s upfront (offline optimum), i.e., \( \mathbb{E}[S] \geq \frac{1}{2} \mathbb{E} [\max_i X_i] \). This result, and its variations and generalizations, have found extensive applications to online optimization and mechanism design, particularly, in the design of simple yet approximately optimal (sequential posted price) mechanisms, both online and offline, for revenue (rewards) maximization while selling items [HKS07, CHMS10, CHK07, KW12] (see Section 1.2 for a detailed discussion).

However, what if the \( X_i \)’s are costs and the goal is cost minimization, like in the case of procuring items while minimizing the payment? For example, consider a house buyer trying to decide when to buy a house in a sellers’ market – where houses are selling fast. When a house arrives with its price (cost) listed, she may have to decide the same day whether to buy it or not. Given that the buyer may have only distributional knowledge of the house prices, the goal is to devise a buying strategy so that the price paid is minimized.

Towards this, we study the cost counterpart of the prophet inequality, where the \( X_i \)’s represent costs arriving in an online manner, and one must “stop” at some point and select the last cost seen. Note that here the constraint is upwards-closed – if one makes it to \( X_n \), they are forced to pick its realization no matter how high it is. The goal is to design a stopping strategy (online algorithm) \( ALG \) that minimizes the expected cost, and is comparable to the cost of an all-knowing prophet who can always select the minimum realization and thus has expected cost \( \mathbb{E} [\min_i X_i] \). For an \( \alpha \geq 1 \), we say that algorithm \( ALG \) achieves an \( \alpha \)-factor cost prophet inequality, or is \( \alpha \)-competitive/approximate, if

\[
\mathbb{E}[ALG] \leq \alpha \cdot \mathbb{E} \left[ \min_i X_i \right].
\] (1)

One may wonder why the cost setting is not equivalent to the classical prophet inequality with negative \( X_i \)’s. The main reason is that an algorithm for classical prophet inequalities in the rewards setting works with downwards-closed constraints (see Section 1.2) and therefore, if all \( X_i \)’s are negative, the optimal solution is trivial: the algorithm will not select any \( X_i \) and obtain a value of 0. However, this violates the upwards-closed constraint of the cost prophet inequality. In fact, this difference turns out to be a crucial one, as we demonstrate that upwards-closed constraints lead to qualitatively different guarantees.

For the rewards setting, the ratio of 1/2 in the classical prophet inequality is achievable through simple single-threshold algorithms [SC84, KW12] of the form “accept the first \( X_i \geq \tau \) for some threshold \( \tau \)”, and is known to be tight. Furthermore, there exist simple online algorithms that achieve constant-factor approximations even for general multi-dimensional settings with complicated constraints (matroids, matchings, etc) [KW12, Ala14, JMZ22a, GW0, EFGT20]. Motivated by these works, we ask:

For the cost minimization setting, what is the optimal online algorithm, i.e., the algorithm that achieves the best (smallest) possible \( \alpha \) in the cost prophet inequality of (1)?

Is the factor achieved a constant? Is it achievable by simple single-threshold algorithms?
In this paper, we study the above questions, and obtain tight bounds for the optimal and single threshold algorithms. Somewhat surprisingly, the bounds turn out to be qualitatively different from what is known for the rewards maximization setting. In what follows, we give an overview of our results and the techniques used.

1.1 Our Contributions

To our surprise, we first observe that for the case where the $X_i$’s are not identically distributed, no algorithm can achieve any bounded approximation if the arrival order of the $X_i$’s is adversarial or even random! In particular, $\alpha$ in (1) can be unbounded, even in the special case of $n = 2$ and the distributions with support at most two (see Example 2.1). This strong negative result leads us to consider the I.I.D. case where the $X_i$’s are drawn independently from the same known distribution.

**Independent and Identically Distributed (I.I.D.) Costs.** In the I.I.D. case, all the $X_i$’s are drawn from a common non-negative known distribution $D$. In the search of constant-factor competitive algorithms for this case, we study optimal algorithms, i.e., algorithms that achieve the smallest possible $\alpha$ in (1). We show that, much like the classical prophet inequality [CFH+17, LLP+21], it suffices to only consider threshold-based algorithms with oblivious thresholds: set thresholds $\tau_1, \ldots, \tau_n$ upfront, and accept the first $X_i \leq \tau_i$. Intuitively, this is because the process is memoryless, i.e. the decision in the $i$-th round is independent of the past realizations and depends only on the realization of $X_i$ and the distribution of the future costs.

Notice that $\tau_n = +\infty$ because, if for all $i < n$ we have $X_i > \tau_i$, then the algorithm has to accept $X_n$ no matter how high its value is. In that case, it incurs cost of $E_{X \sim D}[X]$.

Using this fact, we show that it suffices to set $\tau_{n-1} = E_{X \sim D}[X]$, which is the expected cost of the algorithm if it decides not to select $X_{n-1}$. Then, by inducting this argument backwards, we show that the optimal $\tau_i$ is equal to the expected cost that the optimal algorithm incurs if there are $n - i$ remaining random variables to be drawn from $D$ (see Section 3.1). This fact verifies the natural intuition that we should skip $X_i$ only if its realization is worse than the expected future cost.

**Hazard-rate.** Our analysis of optimal (threshold-based) algorithms crucially relies on the hazard-rate of the distribution. For a given distribution $D$ with probability density and cumulative distribution functions $f$ and $F$ respectively, the hazard rate of $D$ is defined for all $x$ in the support of $D$ as $h(x) \triangleq \frac{f(x)}{1-F(x)}$. Also referred to by some communities as the failure rate, it is a fundamental quantity within several fields of economics and mathematics and has found a lot of applications in survival analysis [KP02], reliability theory [RH04], pricing [HR09, GPZ21] and even forensic analysis [KAnA11]. For our results, we utilize the integral of the hazard rate, $H(x) = f_0^x h(x) \, dx$, which we call the cumulative hazard rate of $D$. In particular, we study distributions in which $H$ is an arbitrary polynomial, subject to being a valid cumulative hazard rate. Since any given hazard rate function $H$ can be approximated to an arbitrary precision by a polynomial, if the support of $D$ is a union of bounded intervals, our results hold for almost all distributions.

**Optimal (Threshold-Based) Algorithm.** Using the characterization of the optimal threshold-based algorithm, we show that, for almost all distributions, the algorithm’s competitive ratio is at most a constant, and this constant depends on the distribution. In particular, for a distribution with cumulative hazard rate being an (arbitrary) polynomial $H(x) = \sum_{i=1}^{k} a_i x^{d_i}$ where $d_1 < d_2 < \cdots < d_k$, the optimal algorithm is $\lambda(d_1)$-competitive, where $\lambda(d_1)$ is a decreasing
function of $d_1$. Furthermore, we show that for $H(x) = x^d$, the optimal algorithm achieves a competitive ratio that is exactly $\lambda(d_1)$. This, together with the fact that this algorithm achieves the best possible factor, implies a matching lower bound. Interestingly, the approximation depends only on the smallest degree $d_1$ of $H$, and this dependence is inverse. Intuitively, this is because $H$ grows rapidly as $d_1$ increases, and thus $D$ has a less heavy tail which leads to a better approximation. It turns out that $\lambda(d_1)$ depends on the Gamma function, the extension of the factorial function over the reals. For its definition, see Section 2. One might find the constraint of focusing only on distributions with polynomial cumulative hazard rate peculiar; however, to the best of our knowledge, all distributions studied or used in practice have cumulative hazard rate that is either a polynomial or is approximable by polynomials to arbitrary precision (see also Remark 3.8). This includes the uniform, exponential, normal, Weibull, Rayleigh, beta and gamma distributions, among many others.

**Theorem 1.1.** For the I.I.D. setting under any given non-negative distribution $D$ with polynomial cumulative hazard rate, for large enough $n$, there exists a $\lambda(d)$-factor cost prophet inequality, where

$$
\lambda(d) = \frac{(1 + 1/d)^{1/d}}{\Gamma(1 + 1/d)},
$$

d is the smallest degree of the polynomial, and $\Gamma(\cdot)$ is the Gamma function.

Moreover, this constant is tight for the distribution with cumulative hazard rate $H(x) = x^d$.

To understand how $\lambda(d)$ grows with $d$, consider Stirling’s approximation for the Gamma function, $\Gamma(z) \approx \sqrt{2\pi z} \left(\frac{z}{e}\right)^z$. Replacing this in the expression of $\lambda(d)$, we have

$$
\lambda(d) \approx \frac{(1 + 1/d)^{1/d}}{\sqrt{2\pi (1 + 1/d)^{1+1/d}}} = \frac{e}{\sqrt{2\pi}} e^{1/d}.
$$

Thus the dependence of $\lambda(d)$ in $d$ is approximately inversely exponential.

**MHR Distributions.** Distributions with monotonically increasing hazard rate have been extensively studied in the mechanism design literature due to their sought after properties and applications (e.g., see [GPZ21, BBDS17, BGGM10, CD11, DW12, DRY15, GKL17, HR09]). These are known as monotone hazard rate (MHR) distributions. For such distributions, we show that the optimal algorithm is 2-factor competitive. In addition, the factor of 2 is tight for the exponential distribution, which has constant hazard rate.

**Theorem 1.2.** For every MHR distribution with polynomial cumulative hazard rate or with support on a union of bounded intervals, there exists a 2-competitive cost prophet inequality for large enough $n$.

This factor is tight, since there is no $(2 - \varepsilon)$-cost prophet inequality for any $\varepsilon > 0$ for the exponential distribution, which has constant hazard rate.

**Single-Threshold Algorithms.** Given the success of single-threshold algorithms in the classical rewards setting, where they are able to achieve the best possible competitive ratio, we ask if there exists a single-threshold algorithm that achieves a constant-factor competitive ratio for the cost prophet inequality setting as well. The answer turns out to be negative. We show that no single-threshold algorithm can achieve a better than poly-logarithmic competitive ratio. We also obtain a matching upper bound for almost all distributions. In particular, given distribution
with a polynomial cumulative hazard rate $H$, we design a threshold $T$ such that the algorithm that selects the first $X_i \leq T$ for $i < n$ and $X_n$ otherwise, yields a $O(\text{polylog } n)$-factor cost prophet inequality, when $n$ is large enough. Here, the power in the poly-logarithmic factor inversely depends on the smallest degree of $H$.

**Theorem 1.3.** Given $X_1, \ldots, X_n$ drawn independently from a non-negative distribution $D$ with polynomial cumulative hazard rate $H(x) = \sum_{i=1}^{k} a_i x^{d_i}$, where $d_1 < d_2 < \cdots < d_k$, there exists a single-threshold algorithm that is $O(\text{polylog } n)$-competitive, for large enough $n$. Moreover, this factor is tight, i.e. there exist distributions for which no single-threshold algorithm is $o(\text{polylog } n)$-competitive.

**Remark 1.4.** While our proofs show that the results hold for large enough $n$, we explain why we believe this to be a technicality and why we expect them to hold for all $n$. Intuitively, the competitive ratio should be an increasing function of $n$, for all distributions. To see this, first notice that the competitive ratio of any algorithm is exactly 1 for $n = 1$, since both the algorithm and the prophet have to pick the realization of $X_1$ whatever its value. Next, notice that one can look at the per decision cost incurred by an algorithm, compared to the prophet’s cost. Since no algorithm can ever make a better decision than the prophet, the more decisions that need to be made by an algorithm, the worse this extra cost should be. We have verified this intuition empirically for several distributions.

**Application to Mechanism Design.** Finally we note that, given the extensive application of classical prophet inequalities in designing simple yet approximately optimal posted pricing mechanisms for selling items (see Section 1.2), our algorithms and results for the cost prophet inequality can be used for the design and analysis of posted-price-style mechanisms for procuring items.

Specifically, consider a procurement auction (also known as a reverse auction), in which one buyer (auctioneer) wants to procure a single item sold by $n$ different sellers, with an I.I.D. distribution governing the sellers’ valuation for selling the item to the buyer or, in other words, the sellers’ costs. If the buyer observes the sellers’ bids in an online manner, for example as is the case in a seller’s housing market, then the standard reduction of a posted-price mechanism to a prophet inequality due to Hajiaghayi et al. [HKS07] applies directly to the cost setting, if one wants to optimize the social cost.

To minimize the cost (price) paid by the buyer (auctioneer), the equivalent to revenue maximization in the cost setting, one simply needs to use the equivalent of virtual valuations, the virtual cost $\phi(c) = c + \frac{F(c)}{f(c)}$, as Myerson’s optimal auction [Mye81] applies to any single-parameter environment. This holds only if $D$ is a regular distribution (a class of distributions which includes MHR distributions among others). For non-regular $D$, one simply needs to “iron” the social cost function, just as in the classical rewards setting and proceed similarly afterwards.

**1.2 Related Work**

Prophet inequalities for reward maximization have been extensively studied in the mechanism design literature. The works of HajiAghayi et al. [HKS07] and Chawla et al. [CHMS10] pioneered the use of prophet inequalities to analyze (sequential) posted price mechanisms for selling items. Specifically, Chawla et al. [CHMS10] observed that the problem of designing optimal (revenue maximizing) posted price mechanisms can be reduced to an appropriate optimal stopping theory problem. This result led to a significant effort to understand how the expected revenue of an optimal posted price mechanism compares to that of the optimal
auction [HKS07, CHK07, CHMS10, Yan11, BH08, ABF+17, Ala14, DKL0, FGL15, DFKL20]. Recently, in a surprising result, Correa at al [CFPV19] showed that the reverse direction also holds, establishing an equivalence between finding stopping rules in an optimal stopping problem and designing optimal posted price mechanisms – for more information on these applications see [Luc17].

The 1/2-competitive factor guaranteed by the classical prophet inequality for adversarial arrival order has been shown to hold for more general classes of downwards-closed constraints, all the way up to matroids [KW12]. For the special case of k-uniform matroids, where one can select up to k values, Alaei [Ala14] showed a $\left(1 - \frac{1}{\sqrt{k+3}}\right)$-competitive ratio. This was recently improved for small k via the use of a static threshold by Chawla et al. [CDL21] and later made tight for all k by Jiang et al. [JMZZ22a]. Ezra et al. [EFGT20] showed a 0.337-prophet inequality for matching constraints. Rubinstein [Rub16] considered general downwards-closed feasibility constraints and obtained logarithmic approximations. The standard setting has been extended to combinatorial valuation functions [RS17, CL21], where one can obtain a constant competitive ratio when maximizing a submodular function but a logarithmic hardness is known for subadditive functions [RS17]. Recently, Chawla et al. [CGKM20] studied non-adaptive threshold algorithms for matroid constraints and gave the first constant-factor competitive algorithm for graphic matroids.

Esfandiari et al. [EHLM15] introduced prophet secretary, in which the arrival order of the random variables is chosen uniformly at random, instead of by an adversary. They gave an adaptive-threshold algorithm that achieves a $1 - 1/e$-competitive ratio and showed no algorithm can achieve a factor better than 0.75. Ehsani et al. [EHK18] extend this result to matroid constraints. The factor of $1 - 1/e$ was recently beaten, first for the case where the algorithm is allowed to choose the arrival order, called the free order setting, by Abolhasani et al. [AEE+17] and later for random arrival order by Azar et al. [ACK17]. The best currently known ratio is obtained by Correa et al. [CSZ20], where they also improve the upper bound to 0.732. When one can select up to k values, Arnosti et al. [AM21] recently gave a surprising and quite beautiful single-threshold algorithm that achieves the best competitive ratio of $1 - e^{-k}\frac{k^k}{k!}$. More general feasibility constraints have also been studied in the random arrival order case, i.e. for matroids [AW20] and matchings [BGMS21, PRSW22].

Our work is most closely related to the long line of work that considers the case of I.I.D. random variables drawn from a known distribution, which dates back to Hill and Kertz [HK82]. Kertz [Ker86] showed that the competitive ratio in the I.I.D. case approaches $\approx 0.745$ as $n$ goes to infinity, via a recursive approach, and conjectured that this is the best bound possible. A simpler proof of this can be found in [SM02]. The bound of $\approx 0.745$ was shown to be tight by Correa et al. [CFH+21]. The proofs of both the upper and lower bounds were recently simplified, by Jiang et al. [JMZh22] and Liu et al. [LLPz21], respectively. We refer the reader to the surveys by Hill and Kertz [HK92] and Correa et al. [CFH+19] for more results about prophet inequalities.

Several of these results are described in the context of an online contention resolution scheme (OCRS), which is an algorithm used to round a fractional solution of a linear program in an online manner. Originally introduced by Chekuri et al [CVZ11] for the offline case, Feldman et al [FSZ16] showed the existence of constant-factor approximate OCRSs for several classes of interesting constraints and demonstrated that an $\alpha$-approximate OCRS for a constraint implies an $\alpha$-competitive prophet inequality for the same constraint. This connection was proved to be deeper, as Lee et al [LS18] used ex-ante prophet inequalities to design optimal OCRSs for matroids. Whether such a connection exists between cost prophet inequalities and OCRSs for
upwards-closed constraints is an interesting open question.

Organization. Section 2 introduces the cost prophet inequality setting, and contains relevant definitions, as well as important observations. Section 3 characterizes the optimal algorithm and shows that it achieves a tight constant-factor approximation. For the special case of MHR distributions, it shows that this constant is exactly 2. In Section 4, we focus on single-threshold algorithms and design a fixed threshold that yields a tight $O$ (polylog $n$)-competitive cost prophet inequality. Finally, we conclude with some interesting open problems in Section 5.

Due to space constraints and to improve the readability, some of the proofs and technical background about the Gamma function can be found in the Appendices.

2 Preliminaries

In this section we formalize the cost prophet inequality setting, and define several important quantities. Given $n$ distributions $D_1, \ldots, D_n$ supported on $[0, +\infty)$, we sequentially observe the realizations of $n$ random costs $X_1 \sim D_1, \ldots, X_n \sim D_n$. We must “stop” at some point and take the last cost seen. In particular, at any point after observing an $X_i$, we can choose to select or discard it. If we select $X_i$, then the process ends and we receive a cost equal to $X_i$. Otherwise $X_i$ gets discarded forever and the process continues. An all-knowing prophet, who can see the realizations of all $X_i$’s upfront can always select the minimum realized cost and hence their expected cost is,

$$\text{Offline-OPT} = \mathbb{E} \left[ \min_i X_i \right].$$

Given $D_1, \ldots, D_n$, the goal is to design a stopping strategy that minimizes the expected cost. That is, design an (online) algorithm $ALG$ to decide when to “stop” and select the last cost seen, such that expected cost incurred is minimized, and ideally is comparable to the prophet’s cost. Formally, for $\alpha \geq 1$, we say that $ALG$ is $\alpha$-factor approximate/competitive, or achieves an $\alpha$-cost prophet inequality if

$$\mathbb{E} [ALG] \leq \alpha \cdot \mathbb{E} \left[ \min_i X_i \right] = \alpha \cdot \text{Offline-OPT}. \quad (2)$$

We first observe that, if $X_i$’s are not identically distributed, no algorithm can achieve any bounded competitive factor, even in restricted settings with simple distributions, when the arrival order is adversarial or random.

Theorem 2.1. For the cost prophet inequality problem with adversarial or random order arrival, no algorithm is $\alpha$-factor competitive for any bounded $\alpha$, even when restricted to $n = 2$ and distributions with support size at most two.

The theorem follows from the following example.

Example 2.1. Let $n = 2$ and consider the following random variables:

$$X_1 = 1 \text{ w.p. } 1, \quad X_2 = \begin{cases} 0 & \text{w.p. } 1 - 1/L \\ L & \text{w.p. } 1/L \end{cases},$$

for an arbitrarily large number $L > 0$. If the arrival order of $X_1$ and $X_2$ is adversarial, an adversary can force every algorithm to see $X_1$ before $X_2$. In this case, every algorithm receives
an expected value of $E[ALG] = 1$, regardless of whether it stops at $X_1$ or at $X_2$. However, the prophet will select $X_1 = 1$ whenever $X_2 = L$, and $X_2 = 0$ otherwise. Thus the prophet’s expected cost is

$$OPT = 0 \cdot (1 - 1/L) + 1 \cdot 1/L = 1/L,$$

which implies an $L$-competitive factor. For random arrival order, notice that with probability $1/2$, the algorithm sees $X_1$ before $X_2$ and thus our previous analysis holds. Therefore, $E[ALG] \geq 1/2$, which implies a competitive factor at least $L/2$.

Since $L$ is arbitrary large, the competitive factor can be made arbitrarily large.

**Remark 2.2.** Observe that the approximation factor $L$ is also distribution dependent, as is the approximation factors we obtain in the I.I.D. setting in Sections 3 and 4. However, both distributions in the example above are MHR. In Section 3.3, we obtain a 2-approximation for MHR distributions, thus showcasing that the non-I.I.D. setting is qualitatively different than the I.I.D. setting.

**I.I.D. Setting.** The negative results of the above theorem leads us to consider the case where the $X_i$’s are independent and identically distributed (I.I.D.). Formally, our algorithm sees one-by-one the realizations of $n$ I.I.D. random variables $X_1, \ldots, X_n$ drawn from a given distribution $D$ with support on $[0, +\infty)$. $D$ is defined by its Cumulative Distribution Function (CDF) $F : [0, +\infty) \rightarrow [0, 1]$, where $F(x) = \Pr_{X \sim D}[X \leq x]$, and let $f : [0, +\infty) \rightarrow [0, 1]$ denote the Probability Density Function (PDF) of $D$. For brevity, we use $\beta_n$ for the remainder of the paper to denote the expected cost of the prophet who can always select the minimum of the $n$ realizations. The following observation that characterizes $\beta_n$ is crucial in our analysis, and its proof can be found in Appendix A.2.

**Observation 2.3.** For $n \geq 1$,

$$\beta_n = \mathbb{E} \left[ \min_{i=1}^{n} X_i \right] = \int_{0}^{\infty} (1 - F(s))^n ds.$$

Given an algorithm $A$, let $G_A(i)$ denote its expected cost, when it observes $i$ I.I.D. random variables drawn from $D$. Thus, the expected cost of $A$ is denoted by $E[A] = G_A(n)$. We also use $R_A(n)$ to denote the competitive ratio of $A$ against the prophet’s cost $\beta_n$, i.e. $R_A(n) = \frac{G_A(n)}{\beta_n}$. Whenever the algorithm is clear from context, we drop the subscript and just use $G(n)$ and $R(n)$.

**Hazard Rate.** All of our results make heavy use of the hazard (failure) rate of a distribution. The hazard rate has been extensively used in the mechanism design literature [GPZ21, BBDS17, BGGM10, CD11, DW12, DRY15, GKL17, HR09], as well as in the mathematics literature under the name of failure rate. We refer the reader to [BPH96] for an extensive overview. Intuitively, for discrete distributions, the hazard rate at a point $t$ represents the probability that an event occurs at time $t$, given that the event has not occurred up to time $t$. For continuous distributions, the hazard rate instead quantifies the instantaneous rate of the event’s occurrence at time $t$.

**Definition 2.4.** For a distribution $D$ with cumulative distribution function $F$ and probability density function $f$, the hazard rate of $D$ is defined as

$$h(x) \triangleq \frac{f(x)}{1 - F(x)},$$
for all $x$ in the support of $D$. Furthermore, let $H$ denote the integral of $h$, which we call the cumulative hazard rate of $D$,

$$H(x) \triangleq \int_0^x h(u) \, du.$$  

Notice that, $H(x) = \int_0^x h(u) \, du = \int_0^x \frac{f(u)}{1 - F(u)} \, du = - \int_0^x \left( \log (1 - F(u)) \right)' \, du = - \ln (1 - F(x))$, which implies that $1 - F(x) = e^{-H(x)}$, and thus we get that

$$\beta_n = \int_0^\infty e^{-nH(u)} \, du. \quad (3)$$

Since $h$ is a non-negative function, $H$ is a non-negative and monotonically non-decreasing function. Using this we obtain the following observation, the proof of which can be found in Appendix A.2.

**Observation 2.5.** Consider a distribution $D$ supported on $[0, +\infty)$ with cumulative hazard rate $H(x) = \sum_{i=1}^k a_i x^{d_i}$, where $d_1 < \cdots < d_k$. Then, $a_1 > 0$ and $d_1 > 0$.

Distributions with monotonically increasing hazard rate have found a special place within mechanism design literature, originally introduced for the study of revenue maximization. They are known as MHR (or IFR) distributions.

**Definition 2.6.** A distribution $D$ is called a Monotone Hazard Rate (MHR) distribution if and only if the hazard rate function $h$ (Definition 2.4) of $D$ is monotonically increasing.

**Gamma function.** The Gamma ($\Gamma$) function – which is an extension of the factorial function over the reals – and its relatives arise in our analysis of the expected cost of the optimal algorithm. For $x > 0$, it is defined as $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t} \, dt$. Of particular interest to us is the lower incomplete Gamma function $\gamma$, which is defined for $s > 0, x \geq 0$ as $\gamma(s, x) = \int_0^x t^{s-1}e^{-t} \, dt$.

To assist the reader, we include a primer on the Gamma function and its relatives in Appendix A.1, along with a few technical lemmas used in our analysis.

### 3 Optimal Algorithm: Constant Approximation via Multiple Thresholds

In this section, we focus on optimal algorithms for the cost prophet inequality (CPI) setting, i.e., algorithms that achieve the smallest possible $\alpha$ in (1). We show that these algorithms achieve a (distribution-dependent) constant-factor CPI for almost all distributions, and a 2-factor CPI for MHR distributions.

We first observe that, just as in the classical prophet inequality, it suffices to focus on threshold-based algorithms to achieve the optimal competitive ratio. A threshold-based algorithm decides thresholds $\tau_1, \ldots, \tau_n$ upfront using only the knowledge of the underlying distribution $D$, and selects the first $X_i \leq \tau_i$. Since the thresholds do not depend on the realizations of the $X_i$'s, the optimal threshold-based algorithm is an oblivious algorithm. The proof of the following proposition can be found in Appendix A.2.

**Proposition 3.1.** For any instance of the cost prophet inequality setting, one can achieve the optimal competitive ratio with a threshold-based oblivious algorithm.
Intuitively, this is because the algorithm’s decision in round \( i \) is independent of past realizations and only depends on the realization of \( X_i \) and the number of remaining random variables, i.e. it is a memoryless process.

Given Proposition 3.1, we focus on threshold-based algorithms. Let \( \tau_1, \ldots, \tau_n \) denote the optimal thresholds of an optimal algorithm. As it turns out, the optimal thresholds have a very natural interpretation; the optimal threshold for the next random variable when we have \( k \) realizations left to see is exactly the expected cost incurred by an optimal algorithm when its input is \( k \) I.I.D. random variables. This implies that the optimal-threshold algorithm will select the next random variable \( X_i \) if and only if the value it sees is smaller than the value it expects to receive by ignoring \( X_i \) and continuing the process.

We then analyze the performance of the optimal-threshold algorithm and show it obtains a constant-factor competitive ratio for (almost) every distribution. We identify each distribution by its cumulative hazard rate \( H(x) \) and show that the constant factor depends on the growth rate of \( H \). Interestingly, when \( H(x) \) is a polynomial, the constant-factor is dominated by the smallest-degree in \( H \). In particular, for a distribution with \( H(x) = \sum_{i=1}^{k} a_i x^{d_i} \), where \( d_1 < \cdots < d_k \), the precise constant factor we obtain is

\[
\lambda(d_1) = \frac{(1 + 1/d_1)^{1/d_1}}{\Gamma(1 + 1/d_1)}.
\]

Perhaps surprisingly, we show that this constant is tight for the distribution with \( H(x) = x^{d_1} \).

We view this as both a positive and a negative result; while we can achieve a constant-factor competitive ratio for every fixed distribution, the constant can be arbitrarily large, as \( \lim_{d \to 0} \lambda(d) = +\infty \). Since all interesting distributions, with the exception of a few pathogenic cases, have a cumulative hazard rate \( H \) that can be approximated to arbitrary precision by a polynomial, our results imply a (distribution-dependent) constant-factor cost prophet inequality for almost all distributions.

Finally, we focus on the special case of MHR distributions and show that if \( D \) is an MHR distribution with \( H(x) = \sum_{i=1}^{k} a_i x^{d_i} \), then \( d_1 \geq 1 \). Since \( \lambda \) is decreasing in \( d_1 \) and \( \lambda(1) = 2 \), this directly implies a tight 2-competitive ratio for almost all MHR distributions.

### 3.1 Characterizing the Optimal Thresholds

In this section we obtain an exact formulation for the optimal thresholds and, using these, design an optimal threshold-based algorithm. In what follows, we use \( G(i) \) to denote \( G_{OPTALG}(i) \) for brevity, where \( OPTALG \) is an optimal algorithm.

**Lemma 3.2.** For the cost prophet inequality problem with random variables \( X_1, X_2, \ldots, X_n \), \( \tau_n = +\infty \) for every algorithm. For \( 1 \leq i \leq n-1 \), the optimal threshold for the random variable \( X_i \) is

\[
\tau_i = G(n - i).
\]

**Proof.** The lemma follows by backwards induction on \( n \).

**Base case.** Since we are forced to select a single value, if the algorithm ever observes \( X_n \), it must select its realization. This is equivalent to \( \tau_n = +\infty \). It then follows that \( G(1) = \mathbb{E}_{X \sim D}[X] \).

**Induction.** Consider the \( i \)-th step, where \( i < n \). For our induction hypothesis, assume that \( \tau_j = G(n - j) \) for all \( i < j < n \). Conditioned to the fact that the algorithm has reached the \( i \)-th step, the expected cost of the optimal algorithm is \( G(n - i + 1) \); i.e. the cost that the optimal

\footnote{Recall that \( H \) is non-decreasing, since its derivative, the hazard rate function \( h \), is non-negative.}
where the second equality follows by the definition of $E$. This is because with probability $\tau$ for any $i$, we will show that $\tau_i = G(n-i)$ minimizes $G(n-i+1)$.

We rearrange (4) and obtain

$$G(n-i+1) = F(\tau_i) \mathbb{E}[X | X \leq \tau_i] + (1 - F(\tau_i)) G(n-i)$$

where the second equality follows by the definition of $\mathbb{E}[X | X \leq \tau_i]$ and the second-to-last equality follows via integration by parts.

We will show that the optimal threshold at the $i$-th step is

$$\tau_i = G(n-i).$$

In other words, we will show that

$$G(n-i)F(G(n-i)) - \int_{0}^{G(n-i)} F(u) du + (1 - F(G(n-i))) G(n-i)$$

$$\leq \tau_i F(\tau_i) - \int_{0}^{\tau_i} F(u) du + (1 - F(\tau_i)) G(n-i),$$

for any $\tau_i \neq G(n-i)$. Rearranging (5), we get

$$G(n-i)F(G(n-i)) - \int_{0}^{G(n-i)} F(u) du + (1 - F(G(n-i))) G(n-i)$$

$$\leq \tau_i F(\tau_i) - \int_{0}^{\tau_i} F(u) du + (1 - F(\tau_i)) G(n-i) \iff$$

$$G(n-i) - \int_{0}^{G(n-i)} F(u) du \leq \tau_i F(\tau_i) - \int_{0}^{\tau_i} F(u) du + G(n-i) - F(\tau_i) G(n-i) \iff$$

$$F(\tau_i) (G(n-i) - \tau_i) \leq \int_{0}^{G(n-i)} F(u) du - \int_{\tau_i}^{G(n-i)} F(u) du \iff$$

$$10$$
\[ F(\tau_i) (G(n - i) - \tau_i) \leq \int_{\tau_i}^{G(n-i)} F(u) \, du. \]  

(6)

We distinguish between two cases: \( \tau_i < G(n - i) \) and \( \tau_i > G(n - i) \). In the case where \( \tau_i < G(n - i) \), (6) becomes

\[ F(\tau_i) \leq \int_{\tau_i}^{G(n-i)} \frac{F(u)}{G(n - i) - \tau_i} \, du, \]

which is true by the mean value theorem, since \( F \) is increasing and \( \tau_i < G(n - i) \). Similarly, in the case where \( \tau_i > G(n - i) \), (6) becomes

\[ F(\tau_i) \geq \int_{\tau_i}^{G(n-i)} \frac{F(u)}{G(n - i) - \tau_i} \, du = \int_{G(n-i)}^{\tau_i} \frac{F(u)}{\tau_i - G(n - i)} \, du, \]

which is again true by the mean value theorem, since \( F \) is increasing and \( \tau_i > G(n - i) \).

We conclude that the optimal threshold for \( X_i \) is

\[ \tau_i = G(n - i). \]

Lemma 3.2 implies that the following threshold-based algorithm is an optimal algorithm; it achieves the best possible competitive ratio for the cost prophet inequality (CPI) problem.

Algorithm 1: Optimal Threshold Algorithm (D)

1. \( \text{Set } \tau_n \leftarrow +\infty. \)
2. \( \text{for } i \leftarrow n - 1 \text{ to } 1 \text{ do} \)
3. \( \quad \tau_i \leftarrow F(\tau_{i+1}) \mathbb{E}[X \mid X \leq \tau_{i+1}] + (1 - F(\tau_{i+1})) \tau_{i+1}. \)
4. \( \text{end} \)
5. \( \text{for } i \leftarrow 1 \text{ to } n - 1 \text{ do} \)
6. \( \quad \text{Let } z_i \text{ be the realization of } X_i. \)
7. \( \quad \text{if } z_1, \ldots, z_{i-1} \text{ were not selected and } z_i \leq \tau_i \text{ then} \)
8. \( \quad \quad \text{Select } z_i. \)
9. \( \quad \text{end} \)
10. \( \text{end} \)

3.2 Constant Factor Competitive Ratio

In this section, we show Theorem 1.1 which states that, for almost all distributions, the optimal-threshold algorithm yields a distribution-dependent constant-factor competitive ratio for the cost prophet inequality problem and this constant is tight.

3.2.1 Upper Bound

For the upper bound, the idea is to show the result for all distributions for which \( H(x) \) is a polynomial. Interestingly, in this case, we show that the constant factor depends only on the smallest degree of \( H(x) \). Since polynomials can approximate almost all functions within an arbitrary degree of precision, this shows that the optimal-threshold algorithm yields a constant-factor cost prophet inequality for almost all distributions.
**Theorem 3.3.** Let $D$ be a distribution on $[0, +\infty)$ for which $H(x) = \sum_{i=1}^{k} a_i x^{d_i}$, where $d_1 < \cdots < d_k$, and let

$$\lambda(d_1) = \frac{(1 + 1/d_1)^{1/d_1}}{\Gamma (1 + 1/d_1)}.$$ 

Then, Algorithm 1 achieves a $\lambda(d_1)$-competitive ratio with respect to $\beta_n$.

**Proof.** By Observation 2.5 we have that $d_1 > 0$ and $a_1 > 0$. For the competitive ratio of Algorithm 1, we start by analyzing its expected cost with respect to the cumulative hazard rate $H(x)$.

**Lemma 3.4.** The expected cost incurred by Algorithm 1 is

$$G(n) = \int_{0}^{G(n-1)} e^{-H(u)} du.$$ 

**Proof.** Recall that $G(n)$ satisfies the recurrence relation in (4)

$$G(n) = F(\tau_1) \mathbb{E}[X \mid X \leq \tau_1] + (1 - F(\tau_1)) G(n - 1).$$

Substituting the optimal thresholds from Lemma 3.2 into the recurrence above, we obtain

$$G(n) = F(G(n - 1)) \mathbb{E}[X \mid X \leq G(n - 1)] + (1 - F(G(n - 1))) G(n - 1)$$

$$= F(G(n - 1)) \int_{0}^{G(n-1)} u f(u) du + (1 - F(G(n - 1))) G(n - 1)$$

$$= \int_{0}^{G(n-1)} u f(u) du + (1 - F(G(n - 1))) G(n - 1)$$

$$= [uf(u) du]_{0}^{G(n-1)} - \int_{0}^{G(n-1)} F(u) du + (1 - F(G(n - 1))) G(n - 1)$$

$$= G(n - 1)F(G(n - 1)) - \int_{0}^{G(n-1)} F(u) du + G(n - 1) - G(n - 1)F(G(n - 1))$$

$$= \int_{0}^{G(n-1)} (1 - F(u)) du.$$ 

Next, recall that $H(x) = -\log (1 - F(x))$, and thus we obtain

$$G(n) = \int_{0}^{G(n-1)} e^{-H(u)} du.$$ 

\[\square\]

Recall that $R(n)$ denotes the competitive ratio of Algorithm 1 for $n$ random variables, and that our algorithm compares against the prophet who always selects the minimum value out of all realizations, i.e. $\beta_n$ on expectation. We want to show that $R(n)$ is upper bounded by a constant for all $n \geq 1$. By Lemma 3.4, we have

$$R(n) = \frac{G(n)}{\beta_n} = \frac{1}{\beta_n} \int_{0}^{G(n-1)} e^{-H(u)} du = \frac{1}{\beta_n} \int_{0}^{G(n-1)} e^{-\sum_{i=1}^{k} a_i u^{d_i}} du.$$ 

Before we proceed, we analyze $\beta_n$. 

12
Lemma 3.5. For every \( n \geq 1 \),
\[
\beta_n = \frac{\Gamma(1 + 1/d_1)}{(a_1 n)^{1/d_1}} + o\left(\frac{1}{n^{1/d_1}}\right).
\]

Proof.
\[
\beta_n = \int_0^\infty e^{-n H(u)}\, du = \int_0^\infty e^{-n \sum_{i=1}^k a_i u^{d_i}}\, du = \int_0^\infty e^{-n a_1 u^{d_1}} \cdot e^{-n \sum_{i=2}^k a_i u^{d_i}}\, du \\
= \int_0^\infty e^{-n a_1 u^{d_1}} \cdot \prod_{i=2}^k e^{-n a_i u^{d_i}}\, du = \int_0^\infty e^{-n a_1 u^{d_1}} \cdot \sum_{\ell_1 \geq 0} \frac{(-n a_1 u^{d_1})^{\ell_1}}{\ell_1!} \, du.
\]
(7)

Let \( x = n a_1 u^{d_1} \leftrightarrow u = \left(\frac{x}{n a_1}\right)^{1/d_1} \). Then,
\[
dx = n a_1 d_1 u^{d_1-1}\, du \leftrightarrow \, du = \frac{u^{1/d_1}}{n a_1 d_1}\, dx = \frac{x^{1/d_1-1}}{n a_1^{1/d_1} d_1}\, dx,
\]
and (7) becomes
\[
\beta_n = \frac{1}{n a_1^{1/d_1} d_1} \int_0^\infty e^{-x^{1/d_1-1}} \cdot \prod_{i=2}^k \sum_{\ell_i \geq 0} \frac{(-n a_i \left(\frac{x}{n a_1}\right)^{d_i/d_1})^{\ell_i}}{\ell_i!} \, dx
\]
\[
= \frac{1}{n a_1^{1/d_1} d_1} \sum_{\ell_2, \ldots, \ell_k \geq 0} \int_0^\infty e^{-x^{1/d_1-1}} \cdot \prod_{i=2}^k \frac{(-n a_i \left(\frac{x}{n a_1}\right)^{d_i/d_1})^{\ell_i}}{\ell_i!} \, dx
\]
\[
= \frac{1}{n a_1^{1/d_1} d_1} \sum_{\ell_2, \ldots, \ell_k \geq 0} \int_0^\infty e^{-x^{1/d_1+1/d_1}} \cdot \prod_{i=2}^k \frac{(-n a_i \left(\frac{x}{n a_1}\right)^{-d_i/d_1})^{\ell_i}}{\ell_i!} \, dx
\]
\[
= \frac{1}{n a_1^{1/d_1} d_1} \sum_{\ell_2, \ldots, \ell_k \geq 0} \prod_{i=2}^k \frac{(-n a_i \left(\frac{x}{n a_1}\right)^{-d_i/d_1})^{\ell_i}}{\ell_i!} \cdot \Gamma\left(1/d_1 + 1/d_1 \sum_{j=2}^k \ell_j\right)
\]
(8)
\[
= \frac{\Gamma(1/d_1)}{d_1 (n a_1)^{1/d_1}} + \frac{1}{d_1 (n a_1)^{1/d_1}} \sum_{\ell_2, \ldots, \ell_k \geq 0, \ell_2, \ldots, \ell_k \neq (0, \ldots, 0)} \prod_{i=2}^k \frac{(-n a_i \left(\frac{x}{n a_1}\right)^{-d_i/d_1})^{\ell_i}}{\ell_i!} \cdot \Gamma\left(1/d_1 + 1/d_1 \sum_{j=2}^k \ell_j\right)
\]
\[
= \frac{\Gamma(1 + 1/d_1)}{(n a_1)^{1/d_1}} + \frac{1}{d_1 (n a_1)^{1/d_1}} \sum_{\ell_2, \ldots, \ell_k \geq 0, \ell_2, \ldots, \ell_k \neq (0, \ldots, 0)} n^{\sum_{j=2}^k \ell_j(1-d_j/d_1)} \prod_{i=2}^k \frac{(-n a_i \left(\frac{x}{n a_1}\right)^{-d_i/d_1})^{\ell_i}}{\ell_i!} \cdot \Gamma\left(1/d_1 + 1/d_1 \sum_{j=2}^k \ell_j\right)
\]
(9)
Thus, (10) becomes
\[ \beta_n = \frac{\Gamma(1 + 1/d_1)}{(a_1 n)^{1/d_1}} + o\left(\frac{1}{n^{1/d_1}}\right). \]

We are now ready to upper bound \( R(n) \).

**Lemma 3.6.** For every \( n \geq 1 \), we have
\[ R(n) \leq \frac{(1 + 1/d_1)^{1/d_1}}{\Gamma(1 + 1/d_1)}. \]

**Proof.** We show that \( R(n) \leq \frac{(1+1/d_1)^{1/d_1}}{\Gamma(1+1/d_1)} \) via induction on \( n \). For \( n = 1 \), \( R(1) = 1 \) and \( \frac{(1+1/d_1)^{1/d_1}}{\Gamma(1+1/d_1)} \geq 1 \) for all \( d_1 > 0 \). For the induction hypothesis, assume \( R(k) \leq \frac{(1+1/d_1)^{1/d_1}}{\Gamma(1+1/d_1)} \) for all \( k \leq n \), and let \( \lambda(d_1) = \frac{(1+1/d_1)^{1/d_1}}{\Gamma(1+1/d_1)} \) for brevity. For \( n + 1 \) we have
\[
R(n + 1) = \frac{1}{\beta_{n+1}} \int_0^{\lambda(1)} e^{\sum_{i=1}^k a_i u_i} du \\
\leq \frac{1}{\beta_{n+1}} \int_0^{\lambda(d_1)\beta_n} e^{\sum_{i=1}^k a_i u_i} du \\
= \frac{1}{\beta_{n+1}} \int_0^{\lambda(d_1)\beta_n} e^{-a_1 u_1^d_1} \prod_{i=2}^k e^{-a_i u_i^d_i} du \\
= \frac{1}{\beta_{n+1}} \int_0^{\lambda(d_1)\beta_n} e^{-a_1 u_1^d_1} \prod_{i=2}^k \sum_{\ell_i \geq 0} \frac{(-a_i u_i^d_i)^{\ell_i}}{\ell_i!} du, \tag{10}
\]
where the second inequality follows by our induction hypothesis, since \( G(n) \leq \lambda(d_1)\beta_n \). Let \( x = a_1 u_1^{d_1} \iff u = \left(\frac{x}{a_1}\right)^{1/d_1} \). Also,
\[
dx = a_1 d_1 u_1^{d_1-1} du \iff du = \frac{a_1^{1-d_1}}{d_1} \cdot \frac{x^{1/d_1-1}}{a_1^{1/d_1} d_1} dx.
\]

Thus, (10) becomes
\[
R(n + 1) \leq \frac{1}{d_1 a_1^{1/d_1} \beta_{n+1}} \int_0^{\alpha_1(\lambda(d_1)\beta_n)^{d_1}} e^{-x} x^{1/d_1-1} \sum_{\ell_2, \ldots, \ell_k \geq 0} \frac{(-a_i \left(\frac{x}{a_1}\right)^{d_i/d_1})^{\ell_i}}{\ell_i!} dx \\
= \frac{1}{d_1 a_1^{1/d_1} \beta_{n+1}} \int_0^{\alpha_1(\lambda(d_1)\beta_n)^{d_1}} e^{-x} x^{1/d_1-1} \sum_{\ell_2, \ldots, \ell_k \geq 0} \prod_{i=2}^k \frac{(-a_i \left(\frac{x}{a_1}\right)^{d_i/d_1})^{\ell_i}}{\ell_i!} dx.
\]
where the third inequality follows by multiplying together the terms of each sum, the fourth inequality follows by exchanging the order of summation and integration, the fifth inequality follows because the product does not depend on \(x\), and the last inequality follows by the definition of \(\gamma(s, x)\).

Now, using Fact A.4, (11) becomes

\[
R(n + 1) \leq \frac{1}{d_1 a_1^{1/d_1} \beta_n^{1 + 1/d_1} \sum_{\ell_1, \ldots, \ell_k \geq 0} \prod_{i=2}^k \frac{(-a_i a_1^{-d_i/d_1})^{\ell_i}}{\ell_i!} \left( a_1 \left( \frac{\lambda}{\beta_n} \right)^{d_1} \right)^{1/d_1 + 1/d_1 \sum_{j=2}^k d_j \ell_j} \left[ \frac{-a_1 (\lambda(d_1) \beta_n)^{d_1} \lambda}{a_1^{1/d_1} \beta_n^{1 + 1/d_1} \sum_{\ell_1, \ldots, \ell_k \geq 0} \prod_{i=2}^k \frac{(-a_i (\lambda(d_1) \beta_n)^{d_1})^{\ell_i}}{\ell_i!} \left( 1 + \sum_{j=1}^k d_j \ell_j \right) \right] \left[ \frac{\lambda(d_1) \beta_n}{\beta_n^{1 + 1/d_1}} \sum_{\ell_1, \ldots, \ell_k \geq 0} \prod_{i=2}^k \frac{(-a_i (\lambda(d_1) \beta_n)^{d_1})^{\ell_i}}{\ell_i!} \left( 1 + \sum_{j=1}^k d_j \ell_j \right) \right]^{-1} \left( 1 + \sum_{j=1}^k d_j \ell_j \right) \right]^{-\ell_1} \left( 1 + \sum_{j=1}^k d_j \ell_j \right) \right) \right]^{-\ell_1} \left( 1 + \sum_{j=1}^k d_j \ell_j \right) \right)}

\]

Claim 3.7. For large enough \(n\),

\[
\frac{\beta_n}{\beta_n^{1 + 1/d_1}} \sum_{\ell_1, \ldots, \ell_k \geq 0} \prod_{i=1}^k \frac{(-a_i (\lambda(d_1) \beta_n)^{d_1})^{\ell_i}}{\ell_i!} \left( 1 + \sum_{j=1}^k d_j \ell_j \right) \leq 1.
\]

Proof. By Lemma 3.5, we know that, for large enough \(n\), there exist a constant \(c \geq 0\) such that

\[
\frac{\beta_n}{(a_1 n)^{1/d_1}} + o \left( \frac{1}{n^{1/d_1}} \right) \leq \frac{\Gamma \left( 1 + 1/d_1 \right)}{(a_1 n)^{1/d_1}} \left( 1 + o \left( 1 \right) \right).
\]
Therefore, we have
\[
\frac{\beta_n}{\beta_{n+1}} = \left( \frac{n+1}{n} \right)^{1/d_1} (1 + o(1)).
\]

Thus
\[
\frac{\beta_n}{\beta_{n+1}} \sum_{\ell_1, \ell_2, \ldots, \ell_k \geq 0} \prod_{i=1}^k \left( \frac{a_i (\lambda(d_1)\beta_n)^{d_i}}{\ell_i! (1 + \sum_{j=1}^k d_j \ell_j)} \right)^{\ell_i} = \left( 1 + \frac{1}{n} \right)^{1/d_1} (1 + o(1)) \sum_{\ell_1, \ell_2, \ldots, \ell_k \geq 0} \prod_{i=1}^k \left( \frac{-a_i (\lambda(d_1)\beta_n)^{d_i}}{\ell_i! (1 + \sum_{j=1}^k d_j \ell_j)} \right)^{\ell_i}
\]
\[
\leq \left( 1 + \frac{1}{n} \right)^{1/d_1} (1 + o(1)) \left( 1 - \sum_{i=1}^k \frac{a_i}{1 + d_i} (\lambda(d_1)\beta_n)^{d_i} + o\left( \frac{1}{n^{1/d_i}} \right) \right). \tag{13}
\]

Notice that
\[
\sum_{i=1}^k \frac{a_i}{1 + d_i} (\lambda(d_1)\beta_n)^{d_i} = \frac{a_1}{1 + d_1} (\lambda(d_1))^{d_1} \beta_n^{d_1} + \sum_{i=2}^k \frac{a_i}{1 + d_i} (\lambda(d_1)\beta_n)^{d_i}.
\]

Also, \(\beta_n^{d_i} = O\left( \frac{1}{n^{d_i/d_1}} \right)\), and for \(i \geq 2\), we have \(d_i > d_1\), which implies that \(\beta_n^{d_i} = o\left( \frac{1}{n^{1/d_i}} \right)\). Thus, (13) becomes
\[
\frac{\beta_n}{\beta_{n+1}} \sum_{\ell_1, \ell_2, \ldots, \ell_k \geq 0} \prod_{i=1}^k \left( \frac{a_i (\lambda(d_1)\beta_n)^{d_i}}{\ell_i! (1 + \sum_{j=1}^k d_j \ell_j)} \right)^{\ell_i} \leq \left( 1 + \frac{1}{n} \right)^{1/d_1} (1 + o(1)) \left( 1 - \frac{a_1}{1 + d_1} (\lambda(d_1)\beta_n)^{d_1} + o\left( \frac{1}{n^{1/d_1}} \right) \right)
\]
\[
\leq \left( 1 + \frac{1}{n} \right)^{1/d_1} + o\left( \frac{1}{n^{1/d_1}} \right) \left( 1 - \frac{a_1}{1 + d_1} (\lambda(d_1)\beta_n)^{d_1} + o\left( \frac{1}{n^{1/d_1}} \right) \right)
\]
\[
= 1 + \frac{1}{d_1 n} \frac{a_1}{1 + d_1} (\lambda(d_1))^{d_1} \frac{\Gamma(1 + 1/d_1)}{a_1 n} + o\left( \frac{1}{n} \right) \tag{14}
\]

For (14) to be bounded above by 1, we need
\[
\frac{1}{d_1 n} \leq \frac{a_1}{1 + d_1} (\lambda(d_1))^{d_1} \frac{\Gamma(1 + 1/d_1)}{a_1 n} \iff (\lambda(d_1))^{d_1} \geq \frac{1 + 1/d_1}{\Gamma(1 + 1/d_1)} \iff \lambda(d_1) \geq \frac{(1 + 1/d_1)^{1/d_1}}{\Gamma(1 + 1/d_1)},
\]
which holds, since \(\lambda(d_1) = \frac{(1 + 1/d_1)^{d_1}}{\Gamma(1 + 1/d_1)}\).

Combining (12) with Claim 3.7, we get that \(R(n+1) \leq \lambda(d_1)\).

Thus, it follows that Algorithm 1 achieves a \(\frac{(1 + 1/d_1)^{1/d_1}}{\Gamma(1 + 1/d_1)}\)-competitive ratio with respect to \(\beta_n\).

**Remark 3.8.** While our proofs show that the results hold for distributions with polynomial cumulative hazard rate, we note that, by the Stone–Weierstrass theorem, any continuous function in a union of closed intervals can be approximated with a sequence of polynomials within arbitrary
precision. Furthermore, if the given distribution $D$ has non-compact domain, one can truncate the domain at a high value, after which the density of the distribution is negligible, and proceed to analyze the truncated distribution.

Furthermore, we note that the approximation factor changes smoothly with respect to $H$ and therefore is similar for similar $H$. Specifically, consider a given $H^*$ and let $\{H_n\}_{n=1}^\infty$ be a sequence of polynomials approximating and converging to $H^*$. Denote the smallest degree of the $n$-th polynomial by $d_{1,n}$. There exists a converging subsequence of $d_{1,n}$, say $d_{1,n_k}$, and the upper bound on the approximation factor depends on $d_{1}^* = \lim_{k \to \infty} d_{1,n_k}$.

Hence, we claim our results hold for all distributions that either have polynomial cumulative hazard rate, or have support that is a compact set. For distributions with non-compact support, our results hold with a small loss in the approximation factors.

Remark 3.9. The astute reader might observe that throughout the paper we’ve assumed that the support of $D$ begins at 0, which implies that $H(x) = \int_0^x h(u) du$, and thus $H(0) = 0$, which in turn implies that $d_{1} > 0$. This is without loss of generality. Specifically, if the support of $D$ begins at $a > 0$, one can “shift” it to the origin to find the approximation factor. Formally, we have $H(x) = \int_0^x h(u) du = \int_{x-a}^a h(u) du$, and thus $H(a) = 0$. Define $H'(x) = H(x + a)$. We have $H'(0) = 0$ and the approximation factor of the original distribution depends on $d_{1}^* > 0$. Thus, this dependence is a technicality that does not affect the approximation factor.

3.2.2 Lower Bound

In this section, we show that there exist distributions for which the upper bounds given by $\lambda$ of the previous section is tight.

Theorem 3.10. Consider the distribution $D$ for which $H(x) = x^d$ for $d \geq 0$. For any $\varepsilon > 0$, there is no $(\frac{(1+1/d)^{1/d}}{\Gamma(1+1/d)} - \varepsilon)$-competitive cost prophet inequality for the single-item setting and I.I.D. random variables drawn from $D$.

Proof. Let $\lambda(d) = \frac{(1+1/d)^{1/d}}{\Gamma(1+1/d)}$.

Lemma 3.11. For every $n \geq 1$,

$$\beta_n = \frac{\Gamma(1 + 1/d)}{n^{1/d}}.$$  

Proof. The proof follows immediately from the proof of Lemma 3.5. In particular, we have

$$\beta_n = \int_0^\infty e^{-nH(u)} du = \int_0^\infty e^{-nu^d} du,$$  

and, by (9) of Lemma 3.5, since $a_1 = 1$ and $a_2 = \cdots = a_k = 0$, we get that

$$\beta_n = \frac{\Gamma(1 + 1/d)}{n^{1/d}}.$$  

Using Lemma 3.11, we have that

$$R(n) = \frac{G(n)}{\beta_n} = \frac{n^{1/d}}{\Gamma(1+1/d)} \int_0^{G(n-1)} e^{-H(u)} du = \frac{n^{1/d}}{\Gamma(1+1/d)} \int_0^{G(n-1)} e^{-u^d} du.$$  

Let \( x = u^d \iff u = x^{1/d} \). Also,

\[
\begin{align*}
  dx &= u^{1/d-1} du \iff du = \frac{u^{1-d}}{d} dx = \frac{x^{1/d-1}}{d} dx,
\end{align*}
\]

and thus (16) becomes

\[
R(n) = \frac{n^{1/d}}{d \Gamma(1 + 1/d)} \int_0^{(G(n-1))^d} e^{-x x^{1/d-1}} dx = \frac{n^{1/d}}{d \Gamma(1 + 1/d)} \frac{1}{d} \gamma \left( 1/d, (G(n-1))^d \right).
\]  

(17)

where the second equality follows from the definition of the lower incomplete Gamma function.

**Lemma 3.12.** \( R(n) \) is increasing in \( n \).

**Proof.** Recall that, by (17), we have

\[
R(n) = \frac{n^{1/d}}{d \Gamma(1 + 1/d)} \gamma \left( 1/d, (G(n-1))^d \right)
= \frac{1}{d \beta_n} \frac{n^{1/d}}{d \beta_n} \gamma \left( 1/d, (G(n-1))^d \right)
= \frac{1}{d \beta_n} \gamma \left( 1/d, (G(n-1))^d \right).
\]

However, by Fact A.4, we have

\[
\gamma \left( 1/d, (G(n-1))^d \right) = G(n-1) \sum_{k=0}^{\infty} \frac{(- (G(n-1))^d)^k}{k! (1/d + k)}.
\]

Thus

\[
R(n) = \frac{G(n-1)}{d \beta_n} \sum_{k=0}^{\infty} \frac{(- (G(n-1))^d)^k}{k! (1/d + k)}
= \frac{G(n-1)}{\beta_{n-1}} \frac{\beta_{n-1} \beta_n}{\beta_n} \sum_{k=0}^{\infty} \frac{(- (G(n-1))^d)^k}{k! (1 + d k)}
= R(n-1) \frac{\beta_{n-1}}{\beta_n} \sum_{k=0}^{\infty} \frac{(- (G(n-1))^d)^k}{k! (1 + d k)}.
\]

It suffices to show that

\[
\frac{\beta_{n-1}}{\beta_n} \sum_{k=0}^{\infty} \frac{(- (G(n-1))^d)^k}{k! (1 + d k)} \geq 1.
\]

Notice that

\[
\frac{\beta_{n-1}}{\beta_n} = \left( \frac{n}{n-1} \right)^{1/d} = \left( 1 + \frac{1}{n-1} \right)^{1/d} = \sum_{\ell=0}^{1/d} \frac{1}{(n-1)^\ell} \left( \frac{1}{\ell} \right).
\]
Thus, it suffices to show that

\[
\sum_{\ell=0}^{1/d} \frac{1}{(n-1)^{\ell}} \left( \frac{1}{d} \right) \cdot \sum_{k=0}^{\infty} \frac{(- (G(n-1))^d)^k}{k! \, (1 + d k)} \geq 1.
\]

We use the fact that \(G(n-1) \leq \lambda(d) \beta_{n-1} = (\frac{1+1/d}{n-1})^{1/d}\) and get

\[
\sum_{\ell=0}^{1/d} \frac{1}{(n-1)^{\ell}} \left( \frac{1}{d} \right) \cdot \sum_{k=0}^{\infty} \frac{(- (G(n-1))^d)^k}{k! \, (1 + d k)} = \sum_{k=0}^{\infty} \sum_{\ell=0}^{1/d} \frac{1}{(n-1)^{\ell}} \left( \frac{1}{d} \right) \cdot \frac{(- (G(n-1))^d)^k}{k! \, (1 + d k)}
\]

\[
\geq \sum_{k=0}^{\infty} \frac{1}{(n-1)^{\ell}} \left( \frac{1}{d} \right) \cdot \frac{(- (1 + 1/d)^k}{k! \, (n-1)^k \,(1 + d k)}
\]

\[
= \sum_{k=0}^{\infty} \frac{1}{(n-1)^{\ell}} \left( \frac{1}{d} \right) \cdot \frac{(- (1 + 1/d)^k}{k! \, (1 + d k)} \cdot \frac{1}{(n-1)^{\ell+k}}
\]

\[
= 1 + \frac{1}{d(n-1)} - \frac{1 + 1/d}{(d+1)(n-1)} + O\left( \frac{1}{n^2} \right)
\]

Thus, for this quantity to be greater than 1, it suffices to have

\[
\frac{1}{d(n-1)} \geq \frac{1 + 1/d}{(d+1)(n-1)} \iff \frac{d + 1}{d} \geq 1 + 1/d,
\]

which is true.

Assume, towards contradiction, that \(\lim_{n \to \infty} R(n) = \lambda^* < \lambda(d_1) = \frac{(1+1/d)^{1/d}}{\Gamma(1+1/d)}\).

We know that \(G(n-1) = R(n-1) \beta_{n-1} = \frac{\Gamma(1+1/d)}{(n-1)^{\gamma}} R(n-1)\). Thus we get

\[
R(n) = \frac{n^{1/d}}{\Gamma(1+1/d)} \left( \frac{1}{d}, \frac{(1 + 1/d)^d}{(n-1)} \right) \left( R(n-1) \right)^d.
\] (18)

Recall that, by Fact A.4, \(\gamma(s, x) = x^s \sum_{k=0}^{\infty} \frac{(-x)^k}{k! \, (s+k)}\), and thus (18) becomes

\[
R(n) = \left( \frac{n}{n-1} \right)^{1/d} \left( \frac{1}{d}, \frac{\Gamma(1+1/d)^d}{(n-1)} \right) \left( R(n-1) \right)^d
\]

\[
R(n) = \left( \frac{n}{n-1} \right)^{1/d} \left( \frac{1}{d}, \frac{\Gamma(1+1/d)^d}{(n-1)} \right) \left( R(n-1) \right)^d
\]

\[
= R(n-1) \left( 1 + \frac{1}{n-1} \right)^{1/d} \sum_{k=0}^{\infty} \frac{\Gamma(1+1/d)^d}{(n-1)} \left( R(n-1) \right)^d
\]

Notice that

\[
\left( 1 + \frac{1}{n-1} \right)^{1/d} = \sum_{\ell=0}^{1/d} \frac{1}{(n-1)^{\ell}} \left( \frac{1}{d} \right).
\]
Thus,\[
\left(1 + \frac{1}{n-1}\right)^{1/d} \sum_{k=0}^{\infty} \frac{\left(-\frac{(\Gamma(1+1/d))^{d}}{(n-1)^d}\right)^k}{k! (1 + d k)} = \sum_{\ell=0}^{1/d} \frac{1}{(n-1)^\ell} \frac{1}{\ell!} \sum_{k=0}^{\infty} \frac{\left(-\frac{(\Gamma(1+1/d))^{d}}{(n-1)^d}\right)^k}{k! (1 + d k)} \\
= \sum_{k=0}^{\infty} \sum_{\ell=0}^{1/d} \frac{1}{(n-1)^{\ell+k}} \frac{1}{\ell!(\ell+k)!} \frac{\left(-\frac{(\Gamma(1+1/d))^{d}}{(n-1)^d}\right)^k}{k! (1 + d k)} \\
= \sum_{k=0}^{\infty} \frac{1}{d (n-1)^n} \frac{1}{d! (n-1)!} \left(1 + \frac{1}{n-1}\right)^{1/d} \sum_{k=0}^{\infty} \frac{\left(-\frac{(\Gamma(1+1/d))^{d}}{(n-1)^d}\right)^k}{k! (1 + d k)} \\
\approx 1 + \frac{1}{d (n-1)} - \frac{(\Gamma(1+1/d) \cdot R(n-1))^d}{(d+1) (n-1)},
\]
where, for large enough \(n\), we can ignore higher order terms and we also have \(R(n-1) \approx \lambda^*\). Thus, for \(R(n) \leq \lambda^*\), it must be that
\[
\frac{1}{d} - \frac{(\Gamma(1+1/d) \cdot \lambda^*)^d}{(d+1)} \leq 0 \iff \frac{\Gamma(1+1/d) \cdot \lambda^*)^d}{(d+1)} \geq 1 + 1/d \iff \lambda^* \geq \frac{(1 + 1/d)^{1/d}}{\Gamma(1+1/d)},
\]
and we arrive at a contradiction.

Therefore, for any \(\varepsilon > 0\), there is no \((1+1/d)^{1/d} - \varepsilon\)-competitive cost prophet inequality for the single-item setting and i.i.d. random variables drawn from \(\mathcal{D}\).

One can see Theorem 1.1 as both a positive and a negative result, since even though for almost all distributions there exists an algorithm that achieves a constant-factor competitive ratio, this constant can be arbitrarily large.

Now, Theorem 1.1 follows by Theorems 3.3 and 3.10.

### 3.3 Special Case: MHR Distributions

Even though the constant-factor competitive ratio obtained by Algorithm 1 is distribution-dependent, it turns out that we can show a uniform factor of 2 when the distributions are MHR. This factor is also tight, and it provides a nice parallel to the standard \(1/2\)-competitive prophet inequality in the rewards setting [KS77, KS78, SC84, KW12].

**Theorem 3.13.** For every MHR distribution with polynomial cumulative hazard rate or with support on a union of bounded intervals, there exists a 2-competitive cost prophet inequality for large enough \(n\).

**Proof.** Let \(\mathcal{D}\) be an MHR distribution with cumulative hazard rate \(H(x) = \sum_{i=1}^{k} a_i x^{d_i}\), where \(d_1 < \cdots < d_k\). Notice that since \(\mathcal{D}\) has a monotonically increasing hazard rate, we have \(h'(x) = H''(x) \geq 0\) everywhere in \([0, +\infty)\). Thus,
\[
\left(\sum_{i=1}^{k} a_i x^{d_i}\right)^{''} \geq 0 \iff \left(\sum_{i=1}^{k} a_i d_i x^{d_i-1}\right)' \geq 0 \iff \sum_{i=1}^{k} a_i d_i (d_i - 1) x^{d_i-2} \geq 0,
\]
for all \( x \geq 0 \). Recall that, by Observation 2.5, for \( H \) to be the cumulative hazard rate of a distribution \( D \), it must be an increasing function in \( x \), and thus \( a_1 > 0 \).

Assume towards contradiction that \( d_1 < 1 \), which implies that the first term of \( H \) is negative. We use this to contradict the fact that \( H''(x) \geq 0 \) everywhere. In particular, consider a point \( y \) where

\[
\begin{align*}
    a_1 d_1 (1 - d_1) y^{d_1} > & \sum_{i=2}^{k} a_i d_i (d_i - 1) y^{d_i} \iff \\
    y^2 \left( a_1 d_1 (1 - d_1) y^{d_1 - 2} - \sum_{i=2}^{k} a_i d_i (d_i - 1) y^{d_i - 2} \right) > & 0 \iff \\
    -y^2 H''(y) > & 0 \implies H''(y) < 0.
\end{align*}
\]

Such a point can always be found because, for any choice of \( a_1, \ldots, a_k \) and \( d_1 < \cdots < d_k \), one can pick a small enough \( y \) that ensures \( a_1 d_1 (1 - d_1) y^{d_1} \) dominates the term \( \sum_{i=2}^{k} a_i d_i (d_i - 1) y^{d_i} \).

Therefore, for all MHR distributions with cumulative hazard rate \( H(x) = \sum_{i=1}^{k} a_i x^{d_i} \), it must be the case that \( d_1 \geq 1 \). This implies that for every MHR distribution \( D \) with polynomial cumulative hazard rate, \( \lambda(d_1) \leq \lambda(1) = 2 \), and thus Algorithm 1 obtains a 2-factor approximation to the prophet’s cost.

Furthermore, notice that if we consider the distribution with \( H(x) = x \), i.e. the exponential distribution, then, as a corollary of Theorem 1.1 for \( d = 1 \), we get that the factor of 2 is tight. The exponential distribution is MHR as it has a constant hazard rate, and hence we obtain the following result.

**Theorem 3.14.** For any \( \varepsilon > 0 \), there exists no \((2 - \varepsilon)\)-factor cost prophet inequality for the exponential distribution.

Now, Theorem 1.2 follows by Theorems 3.13 and 3.14.

### 4 Single Threshold Algorithm

This section is dedicated to proving Theorem 1.3. We design an algorithm which sets a fixed threshold \( T \) and selects the first realization that is below \( T \). If our algorithm ever reaches \( X_n \) and has not selected any value, it is forced to pick the realization of \( X_n \) regardless of its cost. Our choice of \( T \) is

\[
T = O \left( \left( \frac{\log n}{n} \right)^k \right),
\]

for an appropriate value of \( k \) that depends on the given distribution.

As in Section 3, we analyze our algorithm’s performance for a distribution with \( H(x) = \sum_{i=1}^{k} a_i x^{d_i} \), where \( d_1 < \cdots < d_k \), and obtain a \( O \left( (\log n)^{1/d_1} \right) \)-competitive ratio. We then proceed to show that this ratio is asymptotically tight, as we show that no single threshold algorithm can achieve a competitive ratio better than \( \Omega \left( (\log n)^{1/d} \right) \) for the distribution with \( H(x) = x^d \). As before, our results imply a \( O(\text{polylog } n) \)-factor single-threshold cost prophet inequality for the single-item setting, for almost all distributions.
4.1 Upper Bound

**Theorem 4.1.** Let $D$ be a distribution on $[0, +\infty)$ for which $H(x) = \sum_{i=1}^{k} a_i x^{d_i}$, where $d_1 < \cdots < d_k$. Then, there exists a single threshold $T = T(n, d_1, a_1)$ such that the algorithm that selects the first value $X_i \leq T$ for $i < n$ and $X_n$ otherwise, achieves a $O\left((\log n)^{1/d_1}\right)$-competitive ratio compared to $\beta_n$, for large enough $n$.

**Proof.** We start by analyzing the algorithm’s performance for an arbitrary choice of $T$. We have

$$
E[ALG] = \left(1 - (1 - F(T))^{n-1}\right) E[X | X \leq T] + (1 - F(T))^{n-1} E[X] 
$$

Using the above, (20) becomes

$$
R(n) = \frac{1}{\beta_n} \left( \frac{1 - e^{-(n-1)H(T)}}{1 - e^{-H(T)}} \right) \left( \int_{0}^{T} e^{-H(x)} dx - Te^{-H(T)} \right) + e^{-(n-1)H(T)} \beta_1)
$$

Notice that,

$$
\int_{0}^{T} e^{-H(x)} dx - Te^{-H(T)} \leq T \left(1 - e^{-H(T)}\right).
$$

Using the above, (20) becomes

$$
R(n) \leq \frac{1}{\beta_n} \left( \frac{1 - e^{-(n-1)H(T)}}{1 - e^{-H(T)}} \right) \left(1 - e^{-H(T)}\right) + e^{-(n-1)H(T)} \beta_1)
$$

By Lemma 3.5, we know that there exist constants $c_1, c_2 > 0$ such that for large enough $n$, we have

$$
c_1 \frac{\Gamma(1 + 1/d_1)}{n^{1/d_1}} \leq \beta_n \leq c_2 \frac{\Gamma(1 + 1/d_1)}{n^{1/d_1}}.
$$

Thus, (21) becomes

$$
R(n) \leq \frac{n^{1/d_1}}{c_1 \Gamma(1 + 1/d_1)} \left(1 - e^{-(n-1)H(T)}\right) T + e^{-(n-1)H(T)} c_2 \Gamma(1 + 1/d_1).
$$

Let

$$
T = \left(\log \left(\frac{n}{\log n}\right)\right)^{1/d_1}.
$$
Since \( H(T) = \sum_{i=1}^{k} a_i T^{d_i} \), we have
\[
H(T) = a_1 \frac{\log \left( \frac{n}{\log n} \right)}{d_1 a_1 (n - 1)} + \sum_{i=2}^{k} a_i \left( \frac{\log \left( \frac{n}{\log n} \right)}{d_1 a_1 (n - 1)} \right)^{d_i/d_i} = \frac{\log \left( \frac{n}{\log n} \right)}{d_1 a_1 (n - 1)} + \sum_{i=2}^{k} a_i \left( \frac{\log \left( \frac{n}{\log n} \right)}{d_1 a_1 (n - 1)} \right)^{d_i/d_i}.
\]

Since \( d_i > d_1 \) for all \( i \geq 2 \), we have that, for large enough \( n \),
\[
H(T) \approx a_1 T^{d_1},
\]
as \( \sum_{i=2}^{k} a_i T^{d_i} = o \left( T^{d_1} \right) \). Thus, (23) becomes
\[
R(n) \leq \frac{n^{1/d_1}}{c_1 \Gamma \left( 1 + 1/d_1 \right)} \left( 1 - e^{-\left( n-1 \right) a_1 T^{d_1}} \right) T + e^{-\left( n-1 \right) a_1 T^{d_1}} \frac{c_2}{c_1} n^{1/d_1}
\]
\[
= \frac{n^{1/d_1}}{c_1 \Gamma \left( 1 + 1/d_1 \right)} \left( 1 - \left( \frac{\log n}{n} \right)^{1/d_1} \right) \cdot \left( \frac{\log \left( \frac{n}{\log n} \right)}{d_1 a_1 (n - 1)} \right)^{1/d_1} + e^{-\left( n-1 \right) a_1 \frac{\log \left( \frac{n}{\log n} \right)}{d_1 a_1 (n - 1)}} \frac{c_2}{c_1} \left( \frac{\log n}{n} \right)^{1/d_1} n^{1/d_1}
\]
\[
= \frac{1}{c_1 \Gamma \left( 1 + 1/d_1 \right) \left( d_1 a_1 \right)^{1/d_1}} \left( \frac{n}{n-1} \right)^{1/d_1} \left( 1 - \left( \frac{\log n}{n} \right)^{1/d_1} \right) \cdot \left( \frac{\log \left( \frac{n}{\log n} \right)}{n} \right)^{1/d_1} \frac{c_2}{c_1} \left( \log n \right)^{1/d_1} n^{1/d_1}
\]

However, there exists a constant \( c_3 > 0 \) such that for large enough \( n \),
\[
\left( 1 + \frac{1}{n-1} \right)^{1+1/d_1} \left( 1 - \left( \frac{\log n}{n} \right)^{1/d_1} \right) \leq c_3,
\]
and also \( \left( \log \left( \frac{n}{\log n} \right) \right)^{1/d_1} \leq \left( \log n \right)^{1/d_1} \). Thus,
\[
R(n) \leq \frac{c_3}{c_1 \Gamma \left( 1 + 1/d_1 \right) \left( d_1 a_1 \right)^{1/d_1}} \cdot \left( \log n \right)^{1/d_1} + \frac{c_2}{c_1} \left( \log n \right)^{1/d_1} = O \left( \left( \log n \right)^{1/d_1} \right).
\]

\[ \square \]

4.2 Lower Bound

**Theorem 4.2.** Consider the distribution \( D \) for which \( H(x) = x^d \) for \( d \geq 0 \). There is no \( o \left( \left( \log n \right)^{1/d} \right) \)-competitive single-threshold cost prophet inequality for the single-item setting and I.I.D. random variables drawn from \( D \).

**Proof.** Recall by (20) that
\[
R(n) = \frac{1}{\beta_n} \frac{1}{1 - e^{-H(T)}} \left( \int_{0}^{T} e^{-H(x)} \, dx - T e^{-H(T)} \right) + e^{-\left( n-1 \right) H(T)} \beta_1.
\]
Assume, towards contradiction, that \( R(n) = o \left( (\log n)^{1/d} \right) \). For this to be the case, it must be that
\[
e^{-(n-1)H(T)} \frac{\beta_1}{\beta_n} = o \left( (\log n)^{1/d} \right),
\] (24)
and also that
\[
\frac{1}{\beta_n} \left( 1 - e^{-(n-1)H(T)} \right) \left( \int_0^T e^{-H(x)} \, dx - T e^{-H(T)} \right) = o \left( (\log n)^{1/d} \right). \quad (25)
\]

By (24) and the definition of \( o(\cdot) \), we have that for every \( \varepsilon > 0 \), there must exist a \( n_0 \geq 1 \) such that for all \( n \geq n_0 \), we have
\[
e^{-(n-1)H(T)} \frac{\beta_1}{\beta_n} \leq \varepsilon (\log n)^{1/d} \iff e^{-(n-1)H(T)} \leq \varepsilon \frac{\beta_n}{\beta_1} (\log n)^{1/d} \iff
\]
\[-(n-1)H(T) \leq \log \left( \frac{\varepsilon \beta_n}{\beta_1} (\log n)^{1/d} \right) \iff H(T) \geq \frac{\log \left( \frac{\beta_1}{\varepsilon \beta_n (\log n)^{1/d}} \right)}{n-1} \iff T^d \geq \frac{\log \left( \frac{\beta_1}{\varepsilon \beta_n (\log n)^{1/d}} \right)}{n-1} \iff T \geq \left( \frac{\log \left( \frac{\beta_1}{\varepsilon \beta_n (\log n)^{1/d}} \right)}{n-1} \right)^{1/d}. \quad (26)
\]

However, by (25), we have that for every \( \varepsilon' > 0 \), there must exist a \( n_1 \geq 1 \) such that for all \( n \geq n_1 \), we have
\[
\frac{1}{\beta_n} \left( 1 - e^{-(n-1)H(T)} \right) \left( \int_0^T e^{-H(x)} \, dx - T e^{-H(T)} \right) \leq \varepsilon' (\log n)^{1/d}
\]
\[
\frac{1}{\beta_n} \left( 1 - e^{-(n-1)H(T)} \right) \left( \int_0^T e^{-x^d} \, dx - T e^{-H(T)} \right) \leq \varepsilon' \beta_n (\log n)^{1/d}
\]
\[
\frac{1}{\beta_n} \left( 1 - e^{-(n-1)H(T)} \right) \left( \frac{1}{d} \gamma \left( 1/d, T^d \right) - T e^{-H(T)} \right) \leq \varepsilon' \beta_n (\log n)^{1/d}. \quad (27)
\]
where the last equality follows by substituting \( t = x^d \) in the integral, as seen several other times in the paper.

Notice that \( T \) has to be decreasing in \( n \), since, if not, one can easily see from (19) that the algorithm is too eager to select a value and its performance degrades rapidly as \( n \) increases. Therefore, we know that \( \lim_{n \to \infty} T = 0 \). Furthermore, by Fact A.5, we know that for small \( T \), i.e. large enough \( n \), we have
\[
\gamma \left( 1/d, T^d \right) \approx d T,
\]
and thus (27) becomes
\[
\frac{1}{\beta_n} \left( 1 - e^{-(n-1)H(T)} \right) \left( T - T e^{-H(T)} \right) \leq \varepsilon' \beta_n (\log n)^{1/d} \iff
\]
\[
\frac{1}{\beta_n} \left( 1 - e^{-(n-1)H(T)} \right) - T \left( 1 - e^{-H(T)} \right) \leq \varepsilon' \beta_n (\log n)^{1/d} \iff
\]
\[
T \left( 1 - e^{-(n-1)H(T)} \right) \leq \varepsilon' \beta_n (\log n)^{1/d}.
\]
However, by (26) we know that we must have
\[ T \geq \left( \log \left( \frac{\beta_1 \beta_n (\log n)^{1/d}}{n - 1} \right) \right)^{1/d}, \]
and if
\[ T \left( 1 - e^{-(n-1)H(T)} \right) \leq \varepsilon' \beta_n (\log n)^{1/d}, \]
then it also must be the case that
\[ T \left( 1 - \frac{\varepsilon \beta_n (\log n)^{1/d}}{\beta_1} \right) \leq \varepsilon' \beta_n (\log n)^{1/d}. \]
Notice that by Lemma 3.5
\[ \beta_n = \frac{\Gamma (1 + 1/d)}{n^{1/d}} \quad \text{and} \quad \beta_1 = \Gamma (1 + 1/d), \]
and thus
\[ T \left( 1 - \varepsilon \left( \frac{\log n}{n} \right)^{1/d} \right) \leq \varepsilon' \beta_1 \left( \frac{\log n}{n} \right)^{1/d}. \]
For every \( \varepsilon > 0 \), for \( n \) large enough, we have \( 1 - \varepsilon \left( \frac{\log n}{n} \right)^{1/d} > 0 \), and thus
\[ T \leq \varepsilon' \beta_1 \left( \frac{\log n}{n} \right)^{1/d}. \quad \text{(28)} \]
To arrive at a contradiction, we use (26) and (28) to show that it suffices to find, for every \( \varepsilon > 0 \), a constant \( \varepsilon' > 0 \) such that
\[ \varepsilon' \frac{\beta_1 \left( \frac{\log n}{n} \right)^{1/d}}{1 - \varepsilon \left( \frac{\log n}{n} \right)^{1/d}} < \left( \log \left( \frac{\beta_1 \beta_n (\log n)^{1/d}}{n - 1} \right) \right)^{1/d} \cdot \left( \log \left( \frac{1 - \frac{\log n}{n}}{n - 1} \right) \right)^{1/d}. \]
Indeed, rearranging the terms above, we get
\[
\begin{align*}
\varepsilon' \frac{\beta_1 \left( \frac{\log n}{n} \right)^{1/d}}{1 - \varepsilon \left( \frac{\log n}{n} \right)^{1/d}} &< \left( \log \left( \frac{\beta_1 \beta_n (\log n)^{1/d}}{n - 1} \right) \right)^{1/d} \cdot \left( \log \left( \frac{1 - \frac{\log n}{n}}{n - 1} \right) \right)^{1/d} \\
&= \frac{1}{\beta_1} \cdot \left( 1 - \varepsilon \left( \frac{\log n}{n} \right)^{1/d} \right)^{1/d} \cdot \left( \frac{1}{d} \cdot \log \left( \frac{\log n}{n} \right) \right)^{1/d} \\
&= \frac{1}{\beta_1} \cdot \left( 1 - \varepsilon \left( \frac{\log n}{n} \right)^{1/d} \right)^{1/d} \cdot \left( \frac{1}{d} \cdot \log n \cdot \log \left( \frac{\log n}{n} \right) \right)^{1/d}.
\end{align*}
\]
\[
\begin{align*}
&= \frac{1}{\beta_1} \cdot \left( 1 - \varepsilon \left( \frac{\log n}{n} \right)^{1/d} \right) \cdot \left( \frac{1}{d} \cdot \frac{n}{n - 1} \cdot \frac{\log \left( \frac{1}{\varepsilon d \log n} \right)}{\log n} \right)^{1/d} \\
&= \frac{1}{\beta_1 d^{1/d}} \cdot \left( 1 - \varepsilon \left( \frac{\log n}{n} \right)^{1/d} \right) \cdot \left( \frac{n}{n - 1} \cdot \frac{\log n - \log \left( \varepsilon^d \log n \right)}{\log n} \right)^{1/d}.
\end{align*}
\]

Notice, however, that for any fixed \( \varepsilon > 0 \), we have

\[
\lim_{n \to \infty} \left( 1 - \varepsilon \left( \frac{\log n}{n} \right)^{1/d} \right) \cdot \left( \frac{n}{n - 1} \cdot \frac{\log n - \log \left( \varepsilon^d \log n \right)}{\log n} \right)^{1/d} = 1,
\]

and thus, for every \( \varepsilon > 0 \) there exists a large enough \( n \) and a constant \( 0 < \varepsilon' \frac{1}{\beta_1 d^{1/d}} \) such that (26) and (28) cannot simultaneously hold, and we arrive at a contradiction. \( \square \)

Theorem 1.3 now follows by Theorems 4.1 and 4.2.

5 Conclusion

In this paper, we studied the cost minimization counterpart of the classical prophet inequality due to Krengel, Sucheston and Garling [KS77]. The upwards-closed constraint in our setting makes it fundamentally different and more complex. We show a strong negative result for the non-I.I.D. case when the arrival order is adversarial or random. For the I.I.D. case, we show that the best approximation possible is a distribution-dependent constant, which depends on the growth rate of the cumulative hazard rate of the distribution. This constant is at most 2 for MHR distributions. Furthermore, when restricted to single-threshold algorithms, the best possible approximation is poly-logarithmic, where the power of the logarithm is again a distribution-dependent constant. In all three cases, the results are tight.

Our work opens up a number of interesting questions.

- Our optimal algorithm has \( n \) distinct thresholds, one for each \( X_i \), which is at the other extreme compared to the single-threshold algorithms. What if we are allowed to use at most \( k \)-thresholds for \( k > 1 \)? How does the competitive ratio improve with \( k \), starting with the poly-logarithmic factor we show for \( k = 1 \)?

- Apart from MHR, are there other interesting classes of distributions for which we can get constant-factor approximation for a fixed (distribution independent) constant? A commonly studied class of distributions in the mechanism design literature that could be a reasonable candidate is the class of regular distributions.

- If one has only sample access to \( D \), how does the competitive ratio of the optimal algorithm change with the number of samples?

- An interesting non-I.I.D. setting for which our impossibility results do not apply is the free order setting in which the distributions can differ but the algorithm can select the order in which it sees the realizations. One cannot hope to do better than in the I.I.D. setting, but is a (distribution-dependent) constant-factor competitive ratio possible?
Acknowledgements: The authors would like to thank Matt Weinberg for sharing helpful references on prophet inequalities and VL would like to thank Chandra Chekuri, Efstathios Konstantinos Chrortsios Garitsis, Kesav Krishnan, Shubhang Kulkarni, Rhea Jain and Robert Andrews for their insightful comments and helpful discussions.

References


[DKL0] Paul Dütting, Thomas Kesselheim, and Brendan Lucier. An $o(\log \log m)$ prophet inequality for subadditive combinatorial auctions. *SIAM Journal on Computing*, 0(0):FOCS20–239–FOCS20–275, 0.


Jiashuo Jiang, Will Ma, and Jiawei Zhang. Tightness without counterexamples: A new approach and new results for prophet inequalities, 2022.


Euiwoong Lee and Sahil Singla. Optimal online contention resolution schemes via ex-ante prophet inequalities. In Hannah Bast, Grzegorz Herman, and Yossi Azar, editors, *26th European Symposium on Algorithms, ESA 2018*, Leibniz International


\section*{A Appendix}

\section*{A.1 Background on the Gamma Function}

The Gamma function $\Gamma(x)$ extends the factorial function to complex numbers. In particular,

$$\Gamma(n + 1) = n!$$

for every $n \in \mathbb{N}$.

Here we give a brief and incomplete primer on the Gamma function, to assist the reader. However, for a more extensive treatment along with many folklore results about the function, see [Gau98].
**Definition A.1** (Gamma ($\Gamma$) Function). For every $x > 0$, the *Gamma function* is defined as

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t} \, dt.$$  

Like the factorial function, the Gamma function also satisfies the following recurrence

$$\Gamma(x + 1) = x\Gamma(x).$$  

The following fact is closely related to Stirling’s approximation for the Gamma function and is due to [DLMF, Eq. 5.11.E7].

**Fact A.2.** For $a > 0$ and $b \in \mathbb{R}$, we have

$$\Gamma(a + b) \leq \sqrt{2\pi} \left(\frac{a}{e}\right)^a \cdot a^b.$$  

Of particular use to us are the following special functions that are related to the Gamma function.

**Definition A.3** (Upper ($\Gamma(\cdot, \cdot)$ and Lower $\gamma(\cdot, \cdot)$ Incomplete Gamma Functions). For every $s > 0, x \geq 0$, the *Upper Incomplete Gamma function* is defined as

$$\Gamma(s, x) = \int_x^\infty t^{s-1}e^{-t} \, dt,$$  

whereas the *Lower Incomplete Gamma function* is defined as

$$\gamma(s, x) = \int_0^x t^{s-1}e^{-t} \, dt.$$  

For every $s > 0, x \geq 0$, we have

$$\Gamma(s, x) + \gamma(s, x) = \Gamma(s).$$  

Next, we describe a few known results about the lower incomplete Gamma function that we use throughout the paper.

**Fact A.4.** For the lower incomplete Gamma function $\gamma(s, x)$ with $s, x > 0$, we have

$$\gamma(s, x) = x^s \sum_{k=0}^\infty \frac{(-x)^k}{k! \cdot (s + k)}.$$  

*Proof.* By the definition of the lower incomplete Gamma function, we have

$$\gamma(s, x) = \int_0^x t^{s-1}e^{-t} \, dt = \int_0^x \sum_{k=0}^\infty (-1)^k \frac{t^{s+k-1}}{k!} \, dt = \sum_{k=0}^\infty (-1)^k \frac{x^{s+k}}{k! \cdot (s + k)} = x^s \sum_{k=0}^\infty \frac{(-x)^k}{k! \cdot (s + k)}.$$  

The following fact follows easily via Fact A.4.

**Fact A.5.** We have that, as $x \to 0$,

$$\frac{\gamma(s, x)}{x^s} \to s^{-1}.$$  

The following claim is due to Qi and Mei [QM99].

**Claim A.6.** [See 3.1 in [QM99]] For small enough $x$, we have

$$\gamma(s, x) \leq s^{-1} x^{s-1} e^{-x}.$$  

32
A.2 Missing Proofs and Useful Lemmas

A.2.1 Proof of Observation 2.3

Observation 2.3. For \( n \geq 1 \),

\[
\beta_n = \mathbb{E} \left[ \min_{i=1}^{n} X_i \right] = \int_{0}^{\infty} (1 - F(s))^n \, ds.
\]

Proof. Let \( Y_n = \min_{i=1}^{n} X_i \), then the CDF \( F_{Y_n} \) of \( Y_n \) is

\[
F_{Y_n}(x) = \Pr [Y_n \leq x] = 1 - \Pr [Y_n > x] = 1 - \prod_{i=1}^{n} \Pr [X_i > x] = 1 - (1 - F(x))^n, \forall x \in [0, +\infty).
\]

Recall that for a random variable \( X \), we have

\[
\mathbb{E}[X] = \int_{0}^{\infty} x f_X(x) \, dx = \int_{0}^{\infty} f_X(x) \int_{0}^{x} dt \, dx.
\]

By changing the order of integration, we obtain

\[
\mathbb{E}[X] = \int_{0}^{\infty} \int_{0}^{x} f_X(x) \, dx \, dt = \int_{0}^{\infty} \Pr[X \geq t] \, dt = \int_{0}^{\infty} (1 - F_X(t)) \, dt.
\]

Using this, we get that the expected cost of the prophet (offline optimum), denoted by \( \beta_n \) is,

\[
\beta_n = \mathbb{E}[Y_n] = \int_{0}^{\infty} (1 - F_{Y_n}(s)) \, ds = \int_{0}^{\infty} (1 - (1 - F(s))^n) \, ds = \int_{0}^{\infty} (1 - F(s))^n \, ds.
\]

A.2.2 Proof of Observation 2.5

Observation 2.5. Consider a distribution \( D \) supported on \([0, +\infty)\) with cumulative hazard rate \( H(x) = \sum_{i=1}^{k} a_i x^{d_i} \), where \( d_1 < \cdots < d_k \). Then, \( a_1 > 0 \) and \( d_1 > 0 \).

Proof. Once can easily see that \( a_1 > 0 \) since \( H \) is non-negative. Note that, for any choice of \( a_1, \ldots, a_k \) and \( d_1 < d_2 < \cdots < d_k \), there exists a small enough \( x_* \in [0, 1) \) such that,

\[
\left| a_1 x_*^{d_1} \right| > \sum_{i=2}^{k} \left| a_i x_*^{d_i} \right|.
\]

Thus, if \( a_1 < 0 \), we have \( H(x_*) < 0 \), a contradiction.

Next we show that \( d_1 \geq 0 \). Consider the derivative of \( H \), namely \( h(x) = \sum_{i=1}^{k} a_i d_i x^{d_i-1} \). Again, given that fact that \( d_1 < d_i \) for all \( i \geq 2 \), there exists \( y_* \) such that

\[
\left| a_1 d_1 y_*^{d_1-1} \right| > \sum_{i=2}^{k} \left| a_i d_i y_*^{d_i-1} \right|.
\]

Thus since \( a_1 > 0 \), we have \( a_1 \cdot d_1 < 0 \) which implies \( h(y_*) < 0 \), a contradiction to \( h \) being non-negative.

\[33\]
Proof of Proposition 3.1

Proposition 3.1. For any instance of the cost prophet inequality setting, one can achieve the optimal competitive ratio with a threshold-based oblivious algorithm.

Proof Sketch. Since every algorithm has to select a value, if an algorithm observes the realization of $X_n$, it is forced to select it. When an algorithm sees $X_{n-1}$, it has to decide whether to select it or not. Whatever the decision process of the algorithm, let $p^A(r | X_{n-1} = z)$ be the probability that algorithm $A$ selects the realization of $X_i$, given $X_i$. Then, the expected cost of $A$ is

$$\sum_{z \geq 0} z p^A(r | X_{n-1} = z) + \left( 1 - \sum_{z \geq 0} p^A(r | X_{n-1} = z) \right) \mathbb{E}[X_n].$$

For a fixed choice of $L = \sum_{z \geq 0} p^A(r | X_{n-1} = z)$, to maximize this quantity, $A$ will greedily assign all the probability mass of $L$ to the lowest values $z$. Thus, the only choice $A$ has to make is $L$ itself, which is equal to $\Pr \left[ X_{n-1} \leq F^{-1}(L) \right]$. Therefore, every choice of $L$ implies a threshold, namely $F^{-1}(L)$.

Finally, for the remaining random variables, the observation holds via induction, since the random variables are I.I.D. \qed

34