Columbia B9136 January 28, 2025

Lecture 1 — 4.5 Proofs of the basic Prophet Inequality

Lecturer: Will Ma *Scribe: Will Ma*

1 Four (and a half) Proofs of the basic Prophet Inequality

In 1978, Krengel, Sucheston, and Garling (Krengel and Sucheston, 1978) came upon a beautiful and fundamental fact about the performance of a gambler relative to a prophet in optimal stopping. In this game, the gambler observes a sequence of non-negative rewards $V_1, \dots, V_T$, and can stop at any time to collect the most recent observation. The gambler cannot recoup past rewards, and cannot foretell future rewards, only being told at the start that $V_1, \dots, V_T$ are drawn from *independent* distributions $F_1, \dots, F_T$. The gambler, surprisingly, can always collect at least half as much reward in expectation as a prophet, who observes the realizations $V_1, \dots, V_T$ at the start and can stop on the biggest number.

This result has since found application in Bayesian mechanism design, revenue management, and general online resource allocation, and its proof has been refined and re-derived many times by different communities. We now present four (and a half) different proofs of this fact, that show off some of the "greatest hits" in analytical techniques commonly used for online resource allocation.

Terminology. We imagine the rewards $V_1, \dots, V_T$ as being the *valuations* of T *agents*, with $F_1, \dots, F_T$ being valuation distributions that could be heterogeneous across agents. The agents arrive one-by-one, and the valuation of each agent is observed upon arrival, at which point an *online policy* must irrevocably decide whether they should be *accepted* or *rejected*. At most one agent can be accepted. We sometimes refer to this constraint as starting with a single *item* or *resource*, and only being able to give it to a single agent.

Notation. For a positive integer n, let $[n]$ denote the set $\{1, \dots, n\}$. Let $[\cdot]^+$ denote the operator $\max\{\cdot, 0\}$. We write $:=$ to emphasize equality by definition. We use ALG ("algorithm") to denote the reward of the online policy currently being analyzed, i.e. the valuation that it accepts. ALG is a random variable and we call $\mathbb{E}[\mathsf{ALG}]$ the *performance* of the online policy. Meanwhile, we let $\mathsf{OFF} := \max_{t \in [T]} V_t$ denote the reward of the prophet ("offline"), and call $\mathbb{E}[\mathsf{OFF}]$ its performance.

Theorem 1. *For any positive integer T and independent distributions $F_1, \dots, F_T$ over real numbers, an online policy can guarantee $\mathbb{E}[\mathsf{ALG}] \geq \mathbb{E}[\mathsf{OFF}]/2$.*

1.1 (Two) Proofs of Theorem 1 based on Thresholding

Consider a policy that fixes a threshold $\tau \geq 0$ at the start, and accepts the first agent (if any) whose valuation is at least τ. We analyze such a policy, which will inform how τ should be set.

For any $t \in [T]$, let $X_t \in \{0, 1\}$ be the indicator random variable that agent t is accepted. Let $N_t := \sum_{t' \leq t} X_t$ denote the number of accepted agents up to and including t. On any realization of $V_1, \dots, V_T$, the reward collected by the threshold policy can be decomposed as

$$\begin{aligned}\mathsf{ALG} = \sum_{t=1}^{T} V_t X_t &= \sum_t (V_t - \tau) X_t + \tau \sum_t X_t \\ &= \sum_t [V_t - \tau]^+ (1 - N_{t-1}) + \tau N_T\end{aligned}$$

where the second equality holds because the threshold policy accepts agent t if and only if $V_t \geq \tau$ and $N_{t-1} = 0$.

Taking expectations on both sides, the random variables $[V_t - \tau]^+$ and $(1 - N_{t-1})$ crucially are *independent*, since N_{t-1} depends only on valuations $V_1, \ldots, V_{t-1}$ which are independent from V_t. Therefore, $\mathbb{E}\big[[V_t - \tau]^+(1 - N_{t-1})\big] = \mathbb{E}\big[[V_t - \tau]^+\big]\mathbb{E}[1 - N_{t-1}]$ for all t and we get:

$$\begin{aligned} \mathbb{E}[\mathsf{ALG}] &= \sum_t \mathbb{E}\big[[V_t - \tau]^+\big](1 - \mathbb{E}[N_{t-1}]) + \tau\mathbb{E}[N_T] \\ &\geq (1 - \mathbb{E}[N_T]) \sum_t \mathbb{E}\big[[V_t - \tau]^+\big] + \tau\mathbb{E}[N_T]. \end{aligned} \tag{1}$$

The inequality holds because $N_0 \leq \cdots \leq N_T$ and $[V_t - \tau]^+ \geq 0$ (which is why we needed to take the $[\cdot]^+$ operator).

Lower bound (1) is useful because we can control the threshold τ, which affects $\mathbb{E}[N_T]$ (recall that $N_T \in \{0, 1\}$ and equals 1 if and only if an agent has valuation at least τ). Our goal is to set τ in a way that roughly balances the terms, noting that (1) could be close to 0 if τ is too low or too high (because $\mathbb{E}[N_T]$ is decreasing in τ). The following theorem implies that the coefficients $\sum_t \mathbb{E}\big[[V_t - \tau]^+\big]$ and τ combine to upper bound $\mathbb{E}[\mathsf{OFF}]$ for any τ, which will lead to two ways for setting the threshold that prove Theorem 1.

Lemma 2. *On any realization of* $V_1, \ldots, V_T$*, we have* $\sum_t [V_t - \tau]^+ + \tau \geq \mathsf{OFF}$.

Proof of Theorem 2. Assuming valuations are non-negative, OFF always equals V_t for some t, in which case the LHS quantity is clearly at least $[V_t - \tau]^+ + \tau \geq (V_t - \tau) + \tau = V_t$. □

Lemma 3. *(i) Setting a threshold of* $\tau = \mathbb{E}[\mathsf{OFF}]/2$ *yields* $\mathbb{E}[\mathsf{ALG}] \geq \mathbb{E}[\mathsf{OFF}]/2$.

(ii) Setting a threshold τ *such that* $\mathbb{E}[N_T] = 1/2$ *also yields* $\mathbb{E}[\mathsf{ALG}] \geq \mathbb{E}[\mathsf{OFF}]/2$.

Proof of Theorem 3. (i) Substituting $\tau = \mathbb{E}[\mathsf{OFF}]/2$ into Theorem 2 and taking expectations on both sides, we get $\mathbb{E}\big[\sum_t [V_t - \tau]^+\big] \geq \mathbb{E}[\mathsf{OFF}]/2$. Combining with (1), we get

$$\mathbb{E}[\mathsf{ALG}] \geq (1 - \mathbb{E}[N_T])\frac{\mathbb{E}[\mathsf{OFF}]}{2} + \frac{\mathbb{E}[\mathsf{OFF}]}{2}\mathbb{E}[N_T] = \frac{\mathbb{E}[\mathsf{OFF}]}{2}.$$

(ii) Theorem 2 (after taking expectations on both sides) says that $\mathbb{E}\big[\sum_t [V_t - \tau]^+\big] + \tau \geq \mathbb{E}[\mathsf{OFF}]$ at this value of τ where $\mathbb{E}[N_T] = 1/2$. Substituting into (1), we get

$$\mathbb{E}[\mathsf{ALG}] \geq \frac{1}{2}\left(\sum_t \mathbb{E}\big[[V_t - \tau]^+\big] + \tau\right) \geq \frac{\mathbb{E}[\mathsf{OFF}]}{2}.$$

□

Remark 4. In Theorem 3 (ii), when is it possible to set a threshold such that $\mathbb{E}[N_T] = 1/2$?

Recall that $\mathbb{E}[N_T] = \Pr[N_T = 1]$ is the probability of at least one agent realizing a valuation equal to or above the threshold τ. For valuation distributions with point masses, $\mathbb{E}[N_T]$ can "jump" past 1/2 (e.g. if all agents have a deterministic valuation, then $\mathbb{E}[N_T]$ is either 0 or 1). However, if we expand the definition of a threshold policy to allow for randomization when the valuation equals τ, then we can ensure the possibility of $\mathbb{E}[N_T] = 1/2$.

Definition 5 (Static Threshold Policies)**.** Formally, a static threshold policy is defined by two parameters, $\tau \geq 0$ and a tiebreak probability $\rho \in [0,1]$. An agent t *clears the threshold* if either $V_t > \tau$, or $V_t = \tau$ and an independent event ("coin toss") of probability ρ occurs. The policy accepts any agent who clears the threshold, while it is feasible to do so.

Unless otherwise stated, a static threshold policy is assumed to be allowed to use a randomized tiebreak probability $\rho \in (0,1)$. If ρ is unspecified, then it is understood to be 1, which corresponds to the threshold policies defined earlier (where ties are broken in favor of accepting the agent). We sometimes use the phrase *deterministic* to refer to static threshold policies with $\rho \in \{0,1\}$.

The adjective "static" is important in emphasizing that the threshold parameters τ, ρ cannot change over time. It is easy to confirm that the above proofs still hold for any value of $\rho \in [0,1]$, which ensures the existence of a static threshold policy for which $\mathbb{E}[N_T] = 1/2$. This is called the "median" policy, because if no distributions have point masses, then

$$\mathbb{E}[N_T] = 1/2 \quad \iff \quad \Pr[\max_t V_t > \tau] = 1/2 \quad \iff \quad \Pr[\mathsf{OFF} > \tau] = 1/2$$

(τ is the median of the distribution for OFF). By contrast, setting $\tau = \mathbb{E}[\mathsf{OFF}]/2$ is called the "half the mean" policy.

Remark 6. A randomized tiebreak probability $\rho \in (0,1)$ is actually not necessary for Theorem 3 (ii) to go through. Letting τ denote the threshold at which $\mathbb{E}[N_T] = 1/2$ for some $\rho \in [0,1]$, if we shift ρ to 0, then $\mathbb{E}[N_T] \leq 1/2$, while if we shift ρ to 1, then $\mathbb{E}[N_T] \geq 1/2$. Depending on which of the coefficients $\sum_t \mathbb{E}\big[[V_t - \tau]^+\big]$ or τ is bigger, it is ensured that shifting ρ to either 0 or 1 does not decrease the RHS of (1), and hence there exists a deterministic static threshold policy (which also has "median" interpretations; see Samuel-Cahn (1984)) that completes the proof of Theorem 3 (ii).

Although we have now seen two different ways for a deterministic static threshold to have guaranteed performance at least $\mathbb{E}[\mathsf{OFF}]/2$, randomized tie-breaking can still improve performance.

Exercise 1. Let there be two agents, each with valuation that is equally likely to be 2 or 1. Show that a randomized static threshold policy can perform strictly better than any deterministic one.

1.2 Proof of Theorem 1 based on Dynamic Programming

Lemma 7. *The performance an optimal online policy satisfies* $\mathbb{E}[\mathsf{ALG}] \geq \mathbb{E}[\mathsf{OFF}]/2$.

Although Theorem 7 is already implied by Theorem 3 (why?), there is an arguably simpler argument that directly analyzes the optimal online policy using dynamic programming.

Proof of Theorem 7. Let J_t denote the value-to-go (expected reward to be collected) if the game has not ended upon the arrival of agent t. Bellman's equations from dynamic programming tell us that

$$J_t = \mathbb{E}[\max\{V_t, J_{t+1}\}] \qquad \forall t \in [T] \tag{2}$$

where an optimal policy accepts agent t if and only if their realized valuation V_t is at least J_{t+1}. The values of J_t are computed using backward induction over $t = T, \ldots, 1$, with J_{T+1} understood to be 0. From (2), we can derive that

$$J_t - J_{t+1} = \mathbb{E}[\max\{V_t - J_{t+1}, 0\}] = \mathbb{E}\big[[V_t - J_{t+1}]^+\big] \qquad \forall t \in [T]. \tag{3}$$

$\mathbb{E}[\mathsf{ALG}]$ equals J_1, which can be written as

$$\begin{aligned} J_1 = \sum_{t=1}^{T}(J_t - J_{t+1}) &= \mathbb{E}\left[\sum_t [V_t - J_{t+1}]^+\right] \\ &\geq \mathbb{E}\left[\sum_t [V_t - J_1]^+\right] \\ &\geq \mathbb{E}[\mathsf{OFF} - J_1] \end{aligned}$$

where the first equality follows from (3), the first inequality holds because $J_1 \geq \cdots \geq J_{T+1}$, and the second inequality applies Theorem 2 with $\tau = J_1$.

This shows that $J_1 \geq \mathbb{E}[\mathsf{OFF}]/2$, and since $\mathbb{E}[\mathsf{ALG}] = J_1$, the proof is complete. □

A virtue of the dynamic programming proof is that it elegantly extends to the *reusability* setting.

Definition 8 (Reusability Setting)**.** An accepted agent t, instead of consuming our single resource forever, only uses it for a *duration* d_t which is a positive integer. Specifically, if agent t is accepted, then the resource becomes unavailable for agents $t+1, \ldots, t+d_t-1$, but becomes available again starting with agent $t+d_t$. The non-reusable problem is the special case where $d_t = \infty$ for all t.

Exercise 2. Suppose that the usage durations $d_1, \ldots, d_T$ are all deterministic and known at the start. Our goal is to show that an optimal online policy found through dynamic programming has guaranteed performance $\mathbb{E}[\mathsf{ALG}] \geq \mathbb{E}[\mathsf{OFF}]/2$. (Valuations $V_1, \ldots, V_T$ are still initially unknown and drawn independently from known distributions. OFF is still defined by a prophet who knows $V_1, \ldots, V_T$ at the start, and follows the same rules for the reusable resource.)

1. Write Bellman's equations for this problem with reusability, where you only need a variable J_t for the value-to-go when agent t arrives with the resource *available*. Argue that $J_1 \geq \cdots \geq J_T \geq J_{T+1} = 0$.
2. Let $X_t^* \in \{0,1\}$ be the indicator random variable for agent t being accepted by the *prophet*, with $\mathbb{E}[\mathsf{OFF}] = \mathbb{E}[\sum_t V_t X_t^*]$. Writing J_1 as a telescoping sum similar to above, argue that

   $$J_1 \geq \mathbb{E}[\sum_t X_t^*[V_t - J_{t+1} + J_{t+d_t}]^+].$$

3. Complete the proof that $J_1 \geq \mathbb{E}[\mathsf{OFF}]/2$.

   Hint: You will need to use the facts that J_t is decreasing, and that if $X_t^* = X_{t'}^* = 1$ in the prophet's solution and $t' > t$, then $t' \geq t + d_t$.

This proof can be extended to more general settings where the usage durations d_t are random and initially unknown.

1.3 Magician's Proof of Theorem 1

The next proof presented seems unnecessarily abstract, but will turn out to be the most generalizable. It was introduced in Alaei (2014), who referenced a "Magician's problem". The same proof idea is commonly referred to as an "Online Contention Resolution Scheme". A key concept in these proofs is to think in the *quantile* space, instead of the valuation space.

Definition 9 (Assigning a Quantile for each Valuation)**.** Conditional on the realized valuation V_t of an agent t, they are assigned a quantile $Q_t \in [0,1]$. If the measure corresponding to F_t has no mass on the realized value of V_t, then $Q_t = F_t(V_t)$. Otherwise, if it has mass $\delta > 0$ on V_t, then Q_t is assigned uniformly at random from $(F_t(V_t) - \delta, F_t(V_t)]$ (independent of everything else).

Through this construction, the unconditional distribution of Q_t is uniform over [0,1][1], and independent of everything else (under the assumption that V_t is independent from everything else). Higher quantiles correspond to higher valuations.

The construction in Theorem 9 is useful for two reasons. First, we say that agent t draws a *top-x quantile* if $Q_t \in (1-x, 1]$, allowing us to define an event that occurs with probability exactly x (since Q_t is drawn uniformly from [0,1]) whereas in the valuation space we had to worry about tiebreaking. Second and more importantly, the notion of a top-x quantile allows us to upper-bound the expected reward collected by *any* policy (or prophet) from agent t. In particular, let $X_t \in \{0,1\}$ be the indicator for agent t being accepted, a random variable that could be arbitrarily correlated with V_t. Let $x_t = \mathbb{E}[X_t]$. We can reason that:

$$\mathbb{E}[V_t X_t] \le \mathbb{E}[V_t \cdot \mathbb{1}(Q_t > 1 - x_t)] \tag{4}$$

because the LHS is maximized when X_t equals 1 on the largest realizations of V_t. Since X_t can only equal 1 w.p. x_t, and V_t is non-decreasing in Q_t, the LHS is maximized exactly by setting $X_t = \mathbb{1}(Q_t > 1 - x_t)$, completing the proof of (4).

Exercise 3. The goal of this exercise is formulate the problem of "correlating" X_t with V_t as an LP, and shows that the optimal objective value is indeed $\mathbb{E}[V_t \cdot \mathbb{1}(Q_t > 1 - x_t)]$, under the assumption that the distribution is discrete. Let F_t be a valuation distribution whose support is assumed to lie within the finite set $\{r_1, \dots, r_n\}$, where $r_1 > \cdots > r_n \ge 0$. Note that $\sum_{j=1}^{n} \Pr[V_t = r_j] = 1$.

1. For any $x \in [0,1]$, show that the optimal objective value of the following LP:

   $$\begin{aligned} \max \quad & \sum_{j=1}^{n} r_j z_j \\ \text{s.t.} \quad & \sum_{j=1}^{n} z_j \le x \\ & 0 \le z_j \le \Pr[V_t = r_j] \qquad\qquad \forall j \in [n], \end{aligned}$$

   which has decision variables $z_1, \dots, z_n$, is $\mathbb{E}[V_t \cdot \mathbb{1}(Q_t > 1 - x)]$.

2. Recall that $x_t = \mathbb{E}[X_t]$. Complete the proof that $\mathbb{E}[V_t X_t] \le \mathbb{E}[V_t \cdot \mathbb{1}(Q_t > 1 - x_t)]$.

   Hint: Setting $z_j = \Pr[V_t = r_j \cap X_t = 1]$ for all $j \in [n]$ forms a feasible solution to the LP with $x = x_t$.

Definition 10 (Gain Function, Inverse CDF)**.** For each agent t, define $g_t(x)$ to be the function $\mathbb{E}[V_t \cdot \mathbb{1}(Q_t > 1 - x)]$ over $x \in [0,1]$. g_t, the "gain", is an increasing function with $g_t(0) = 0$ and $g_t(1) = \mathbb{E}[V_t]$.

$g_t(x)$ can also be expressed in the following form. Define the inverse CDF for agent t as $F_t^{-1}(q) := \inf\{v \in \mathbb{R} : q \le F_t(v)\}$, which maps a quantile $q \in [0,1]$ to a valuation for distribution

---

[1] Technically, for discrete distributions Q_t is drawn uniformly from $(0,1]$, but the measure-0 difference can be ignored.

F_t. Our construction of Q_t in Theorem 9 satisfies $V_t = F_t^{-1}(Q_t)$ except possibly on a set of measure[2] 0. Therefore, we can express

$$g_t(x) = \mathbb{E}[V_t \cdot \mathbb{1}(Q_t > 1 - x)] = \int_{1-x}^{1} F_t^{-1}(q)dq. \tag{5}$$

From the integral in (5), we can see that $g_t'(x) = F_t^{-1}(1-x)$, which is decreasing in x, and hence g_t is a concave function for any distribution F_t.

Remark 11. We pedantically illustrate the quantile assignment (Theorem 9) and inverse CDF function (Theorem 10) for a discrete distribution that is 0 or 1 each w.p. 1/2. The CDF is $F(v) = 0$ for $v < 0$, $F(v) = 1/2$ for $0 \le v < 1$, and $F(v) = 1$ for $1 \le v$. If the agent draws valuation $V = 1$, then they are assigned a quantile uniformly at random from $(1/2, 1]$ (the "top-half quantile"), while if the agent draws valuation $V = 0$, then they are assigned a quantile uniformly at random from $(0, 1/2]$. Quantile 0 is never assigned. Meanwhile, the inverse CDF is $F^{-1}(q) = 1$ if $q \in (1/2, 1]$, $F^{-1}(q) = 0$ if $q \in (0, 1/2]$, and $F^{-1}(0) = -\infty$. If the agent has quantile Q, then their valuation is $V = F^{-1}(Q)$ unless $Q = 0$.

Returning to (4) and using the notation in Theorem 10, we can upper-bound $\mathbb{E}[\mathsf{OFF}]$ as follows. For each agent t, let X_t^* be the indicator that they are accepted by the prophet, and let $x_t := \mathbb{E}[X_t^*]$. Then

$$\mathbb{E}[\mathsf{OFF}] = \sum_{t=1}^{T} \mathbb{E}[V_t X_t^*] \le \sum_t \mathbb{E}[V_t \cdot \mathbb{1}(Q_t > 1 - x_t)] = \sum_t g_t(x_t). \tag{6}$$

Note that on every sample path, $\sum_{t=1}^{T} X_t^* \le 1$, and hence taking expectations on both sides, we always have $\sum_t x_t \le 1$.

Assume for now that we are able to compute the prophet's acceptance probabilities $x_1, \ldots, x_T$, which motivates the following online policy. Upon the arrival of an agent t, based on the observed V_t, assign them a quantile Q_t as in Theorem 9. If this is a top-x_t quantile (and the item is still available), then accept agent t. The intuition is that we are "imitating" the prophet, and conditional on the item being available for agent t, the expected reward collected is $\mathbb{E}[V_t \cdot \mathbb{1}(Q_t > 1 - x_t)]$ which upper-bounds what the prophet collects from agent t. However, in order to preserve the probability that the item is available for later agents, another independent coin needs to be tossed to determine whether to accept each agent t. Formally, let $\mathrm{Ber}(\rho_t)$ denote an independent Bernoulli random variable of mean $\rho_t \in [0, 1]$, and an arriving agent is accepted if all three of the following conditions are met: (i) the item is available; (ii) $Q_t > 1 - x_t$; and (iii) $\mathrm{Ber}(\rho_t) = 1$.

We now analyze such a policy, which will motivate why we need the Bernoulli probabilities ρ_t to be less than 1 and how to set them. As before, let X_t be the indicator for the online policy accepting an agent t, and let N_{t-1} be the indicator that an agent before t has been accepted (and

---

[2] We will not formally prove this. The formal definition of the inverse CDF looks scary, but intuitively, complications only arise when F_t is "flat" in a region or "jumps" (both of which can occur with discrete distributions). Jumps are easy—the inverse CDF simply maps different quantiles q to the same value, which has a point mass and caused F_t to jump. When the CDF is flat, the corresponding quantile could map ambiguously to valuations—but the CDF can only be flat for a countable number of quantiles, which has measure 0.

hence the item is unavailable for agent t). We can derive

$$\begin{aligned}\sum_{t=1}^{T}\mathbb{E}[V_tX_t] &= \sum_t \mathbb{E}[V_t(1-N_{t-1})\mathbb{1}(Q_t>1-x_t)\mathrm{Ber}(\rho_t)]\\ &= \sum_t g_t(x_t)\rho_t\Pr[N_{t-1}=0]\\ &= \sum_t g_t(x_t)\rho_t\prod_{t'<t}(1-x_{t'}\rho_{t'})\end{aligned}$$

where the first equality holds because $X_t=(1-N_{t-1})\mathbb{1}(Q_t>1-x_t)\mathrm{Ber}(\rho_t)$ by the three conditions of the online policy, the second equality holds because N_{t-1} and $\mathrm{Ber}(\rho_t)$ are independent from $V_t\mathbb{1}(Q_t>1-x_t)$, the third equality holds by definition of N_{t-1}—the item is available for agent t if no agent $t'<t$ both drew a top-$x_{t'}$ quantile and had $\mathrm{Ber}(\rho_{t'})=1$, and all of these are independent events.

Recall from (6) that $\mathbb{E}[\mathsf{OFF}]=\sum_t g_t(x_t)$, and now we have a lower bound of $\sum_t g_t(x_t)\rho_t\prod_{t'<t}(1-x_{t'}\rho_{t'})$ on $\mathbb{E}[\mathsf{ALG}]$. Therefore, if we can set $\rho_1,\dots,\rho_T$ so that the multiplier

$$\rho_t\prod_{t'<t}(1-x_{t'}\rho_{t'}) \tag{7}$$

is at least some constant $c\in[0,1]$ for all $t\in[T]$, then we have proven that $\mathbb{E}[\mathsf{ALG}]\geq c\cdot\mathbb{E}[\mathsf{OFF}]$. Generally, setting $\rho_t=1$ for all t will lead to $c=0$, which is why we needed the probabilities ρ_t in the first place. Indeed, in this case the multiplier (7) would equal $\prod_{t'<t}(1-x_{t'})$, so if $x_1=1-\varepsilon$ for some small ε then the multiplier for agent $t=2$ can already be no larger than ε.

Proposition 12. *Setting $\rho_t=1/2$ for all $t\in[T]$ implies that the multipliers satisfy*

$$\rho_1\prod_{t'<1}(1-x_{t'}\rho_{t'})\geq\cdots\geq\rho_T\prod_{t'<T}(1-x_{t'}\rho_{t'})\geq\frac{1}{4}.$$

Proof of Theorem 12. Substituting in $\rho_t=1/2$ for all t, the inequalities except the final one are immediate. The final one holds because

$$\frac{1}{2}\prod_{t'<T}(1-\frac{1}{2}x_{t'})\geq\frac{1}{2}(1-\sum_{t'<T}\frac{1}{2}x_{t'})\geq\frac{1}{2}(1-\frac{1}{2}(1))=\frac{1}{4}$$

where we have applied both a union bound and the fact that $\sum_{t=1}^{T}x_t\leq 1$. ☐

Theorem 12 in conjunction with the preceding analysis implies that $\mathbb{E}[\mathsf{ALG}]\geq\mathbb{E}[\mathsf{OFF}]/4$. The constant of $c=1/4$ is not ideal because the multipliers for the early-arriving agents are higher than necessary, making the bound for the last agent T as low as $1/4$. To improve Theorem 12, we can increase ρ_t over time, and set them in a way so that the multiplier for every agent t is *exactly* c. That is, we set $\rho_1=c$, set ρ_2 so that $\rho_2(1-\rho_1x_1)=c$, and so forth, which inductively leads to

$$\rho_t=\frac{c}{1-c\sum_{t'<t}x_{t'}}\qquad\text{and}\quad\rho_t\prod_{t'<t}(1-x_{t'}\rho_{t'})=c\qquad\forall t\in[T]. \tag{8}$$

Importantly, we need to ensure that ρ_t is a probability, i.e. ensure that $\frac{c}{1-c\sum_{t'<t}x_{t'}}\leq 1$ for all t, which prevents us from simply setting the constant c to be as big as possible. Since $\sum_{t'<t}x_{t'}$ can be as large as 1, the biggest c that can ensure $\rho_t\leq 1$ for all $t\in[T]$ is $c=1/2$.

Lemma 13. *Setting* $\rho_t = 1/(2 - \sum_{t'<t} x_{t'})$ *for all* $t \in [T]$ *defines a policy with* $\mathbb{E}[\mathsf{ALG}] \geq \mathbb{E}[\mathsf{OFF}]/2$.

Proof of Theorem 13. Given the instructions in (8), this proof writes itself and is omitted. Everything holds as equality, so one needs to plug in the equations and complete the induction. □

Theorem 13 gives us a third proof of Theorem 1. A virtue of this analysis is that it reduces the random state of each agent t to binary—either they draw a top-x_t quantile, in which case we want to accept them (without discriminating on what their exact valuation is); or they do not, in which case we always reject them. And although it seems restrictive to ensure that the item is available with sufficient probability to *every* agent, having more constrains often helps us pins down an exact policy, as we saw in the derivation of Theorem 13.

Exercise 4. Given constants $\rho_1, \ldots, \rho_T \in [0, 1]$, consider an alternate online policy: accept an agent t if $Q_t > 1 - x_t\rho_t$ (and the item is still available). Prove that this policy also satisfies

$$\mathbb{E}[\mathsf{ALG}] \geq \mathbb{E}[\mathsf{OFF}] \cdot \min_{t\in[T]} \rho_t \prod_{t'<t} (1 - x_{t'}\rho_{t'}).$$

Hint: You will have to use the fact that the functions g_t are concave, established in Theorem 10.

1.4 Proof of Theorem 1 based on Sampling

The final proof presented analyzes a fixed threshold policy that is determined based on one sample from each distribution. The policy can be run even if the distributions themselves are unknown, as long as these samples are given. (If the distributions are known, then the policy can also be run by artificially generating these samples.)

Lemma 14. *Consider the algorithm that draws independent samples* $\tilde{V}_1, \ldots, \tilde{V}_T$ *from respective distributions* $F_1, \ldots, F_T$ *at the start, and then accepts the first agent* t *whose valuation* V_t *exceeds* $\max_t \tilde{V}_t$. *The resulting online policy has performance* $\mathbb{E}[\mathsf{ALG}] \geq \mathbb{E}[\mathsf{OFF}]/2$, *where the expectation is taken over the randomness in both the samples and the valuations.*

Remark 15. For this threshold policy to work, we need a form of randomized tiebreaking that differs from Theorem 5—the $2T$ values $\tilde{V}_1, \ldots, \tilde{V}_T, V_1, \ldots, V_T$ need to be treated as distinct and have a strict ordering. This can be achieved e.g. by the policy independently drawing a tiebreak seed uniformly from [0,1] for each of these $2T$ values, and favoring larger tiebreak seeds over smaller ones.

Proof of Theorem 14. The key idea is to re-imagine the random process as follows. Instead of first drawing the samples $\tilde{V}_t$ and then drawing the valuations V_t, we will draw *two* independent samples U_t, U'_t from each distribution F_t, and then randomly choose one of them to be the sample and the other to be the actual valuation, with equal likelihood. We show that

$$\mathbb{E}[\mathsf{ALG} - \mathsf{OFF}/2 | (U_t, U'_t)_{t=1}^T] \geq 0 \tag{9}$$

for any values of $U_1, U'_1, \ldots, U_T, U'_T$, which would complete the proof due to the tower property of conditional expectation.

Intuitively, (9) is not asking for too much because given any values of U_t, U'_t, each of them is equally likely to be the actual valuation. When the biggest value is the actual valuation, OFF may be bigger, but ALG also tends to collect more. Note that the random threshold τ also depends on which of U_t, U'_t is the sample, across the different $t \in [T]$.

Fix $U_1, U'_1, \ldots, U_T, U'_T$ and the tiebreak seed associated with each value, and relabel these $2T$ elements as $W_1, \ldots, W_{2T}$ where $W_1 \geq \cdots \geq W_{2T}$ and identical values are sorted by decreasing tiebreak seed. Note that every element in sequence $W_1, \ldots, W_{2T}$ is paired with exactly one other element in the sequence—the one that was drawn from the same distribution F_t for some $t \in [T]$. Based on sequence $W_1, \ldots, W_{2T}$, let J denote the *smallest* index in $1, \ldots, 2T$ for which W_J is paired with an earlier element W_t, with $t < J$.

Observe that OFF can be no smaller than W_J; its conditional expectation exactly equals

$$\mathbb{E}[\mathsf{OFF} | (U_t, U'_t)_{t=1}^T] = \sum_{j=1}^{J-1} \frac{1}{2^j} W_j + \frac{1}{2^{J-1}} W_J. \tag{10}$$

To see this, let j denote the smallest index in $1, \ldots, J$ for which W_j is an actual valuation (as opposed to a sample, which does not affect OFF). The probability of this index taking any value in $1, \ldots, J-1$ is $1/2^j$, since none of $W_1, \ldots, W_j$ are paired with one another (by definition of the index J). Meanwhile, the probability of this index taking value J is $1/2^{J-1}$, since W_J is paired with one of $W_1, \ldots, W_{J-1}$. Given this index j, we have $\mathsf{OFF} = W_j$, leading to the expectation in (10).

On the other hand, the algorithm's conditional expectation can be lower-bounded as

$$\mathbb{E}[\mathsf{ALG} | (U_t, U'_t)_{t=1}^T] \geq \sum_{j=2}^{J-1} \frac{1}{2^j} W_{j-1} + \frac{1}{2^{J-1}} W_{J-1}. \tag{11}$$

To see this, let j denote the smallest index in $1, \ldots, J$ for which W_j is a *sample*. The algorithm collects at least W_{j-1} as long as $j > 1$, because an actual valuation that exceeds the threshold $\tau = W_j$ (considering tiebreaking) exists, and moreover the policy cannot stop on valuations $W_{j+1}, \ldots, W_{2T}$ since they are either strictly smaller than τ, or equal to τ but with a smaller tiebreak seed. Using the same probabilities for each value of j as in the analysis of (10), we get the lower bound in (11) (if $j = 1$, then the algorithm collects nothing).

Finally, the RHS of 11 can be manipulated as

$$\frac{W_1}{2^2} + \cdots + \frac{W_{J-2}}{2^{J-1}} + \frac{1}{2^{J-1}} W_{J-1} \geq \frac{W_1}{2^2} + \cdots + \frac{W_{J-2}}{2^{J-1}} + \frac{1}{2^J} W_{J-1} + \frac{1}{2^J} W_J$$

with the lower bound being exactly half of (10), completing the proof. □

1.5 Tightness of Theorem 1

Example 16 (Tightness of 1/2 Guarantee)**.** There are $T = 2$ agents, with $V_1 = 1$ w.p. 1, and $V_2 = 1/\varepsilon$ w.p. ε and $V_2 = 0$ w.p. $1 - \varepsilon$. An online policy is indifferent between accepting vs. rejecting agent 1, as both lead to $\mathbb{E}[\mathsf{ALG}] = 1$. On the other hand, $\mathbb{E}[\mathsf{OFF}] = \varepsilon \cdot \frac{1}{\varepsilon} + (1-\varepsilon)1 = 2 - \varepsilon$. Taking the limit as $\varepsilon \to 0$ shows that for any online policy, $\mathbb{E}[\mathsf{ALG}]/\mathbb{E}[\mathsf{OFF}]$ can be as low as 1/2.

Bibliographical notes. The presentation found in these notes is mostly original, synthesizing proofs from the literature. Subsection 1.1 is based on proofs from Samuel-Cahn (1984); Kleinberg and Weinberg (2012). Subsection 1.2 is based on proofs from Rusmevichientong et al. (2020); Baek and Ma (2022). Subsection 1.3 is based on the proof from Alaei (2014). Subsection 1.4 reproduces exactly the proof from Rubinstein et al. (2020).

References

S. Alaei. Bayesian combinatorial auctions: Expanding single buyer mechanisms to many buyers. *SIAM Journal on Computing*, 43(2):930–972, 2014.

J. Baek and W. Ma. Bifurcating constraints to improve approximation ratios for network revenue management with reusable resources. *Operations Research*, 70(4):2226–2236, 2022.

R. Kleinberg and S. M. Weinberg. Matroid prophet inequalities. In *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*, pages 123–136, 2012.

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. *Innovations in Theoretical Computer Science*, page ., 2020.

P. Rusmevichientong, M. Sumida, and H. Topaloglu. Dynamic assortment optimization for reusable products with random usage durations. *Management Science*, 66(7):2820–2844, 2020.

E. Samuel-Cahn. Comparison of threshold stop rules and maximum for independent nonnegative random variables. *the Annals of Probability*, 12(4):1213–1216, 1984.