

The Power of Two Choices in Dynamic Pricing: Near-Instance-Optimal Regret and Query Complexity

Yan Dai^{1,3(ex-intern)}

Tao Lin^{2,3(ex-postdoc)}

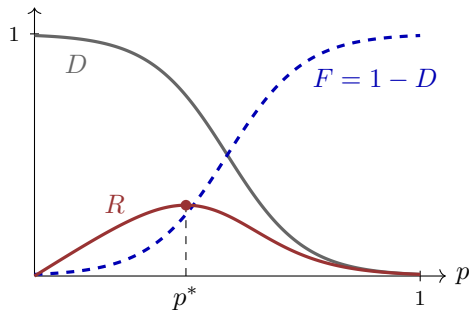
Aleksandrs Slivkins³

¹MIT ORC

²CUHK-Shenzhen

³Microsoft Research

Demand Function = Value Distribution



- Demand $D(p)$: probability of buying at price p , decreasing in p (normalized to $[0, 1]$)
- $F := 1 - D$ increases from 0 to 1
- Equivalently: each customer draws a private **value** $v \sim F$, and buys iff $v \geq p$

$$D(p) = \Pr_{v \sim F}\{v \geq p\}$$

- Revenue $R(p) := p \cdot \Pr_{v \sim F}\{v \geq p\}$
- Monopoly price $p^* := \operatorname{argmax}_p R(p)$

Dynamic Pricing with Unknown Demand

In each round $t = 1, 2, \dots, T$

- A customer arrives with value $v_t \stackrel{\text{i.i.d.}}{\sim} F$, where F is **unknown** to the firm
- The firm posts a price p_t , based on the information from past rounds
- The customer buys iff $v_t \geq p_t$: the firm only sees $\mathbb{1}[v_t \geq p_t]$, and earns $p_t \mathbb{1}[v_t \geq p_t]$

Regret: revenue gap compared with knowing F

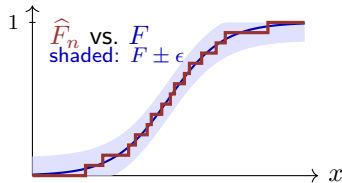
$$\mathfrak{R}_T := T \cdot R(p^*) - \mathbb{E} \left[\sum_{t=1}^T p_t \mathbb{1}[v_t \geq p_t] \right]$$

Goal: $\mathfrak{R}_T = o(T)$ (i.e., as good as p^* when $T \rightarrow \infty$) for any possible distribution F
Or at least... We need to learn p^* up to ϵ $R(\hat{p}) \geq R(p^*) - \epsilon$ in $\text{poly}(\epsilon^{-1})$ rounds

Warm-Up: What If We Saw Each v_t after Each Round t ?

If we have n i.i.d. samples $v_1, v_2, \dots, v_n \sim F$

- **Empirical CDF** $\hat{F}_n(x) := \frac{1}{n} \sum_{i=1}^n \mathbb{1}[v_i \leq x]$
- $\Pr\{\sup_x |\hat{F}_n(x) - F(x)| > \epsilon\} \leq 2e^{-2n\epsilon^2}$ [DKW56]
- $\hat{R}_n(p) := p(1 - \hat{F}_n(p))$ is $1/\sqrt{n}$ -close to R

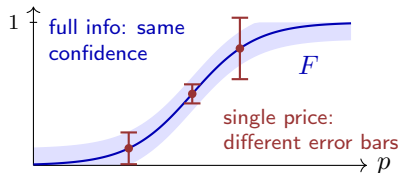
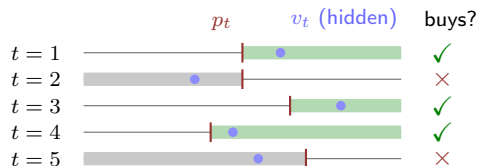


Sample Complexity $\text{Sample}(\epsilon)$

Number of samples $v_1, \dots, v_n \stackrel{\text{i.i.d.}}{\sim} F$ needed to find a $\hat{p}$ with $\mathbb{E}[R(p^*) - R(\hat{p})] \leq \epsilon$

$\hat{R}_n$ and R ϵ -close $\implies \hat{p}_n := \operatorname{argmax}_p \hat{R}_n(p)$ is 2ϵ -optimal $\implies \text{Sample}(\epsilon) = \mathcal{O}(\epsilon^{-2})$

Single Prices: One Bit per Customer



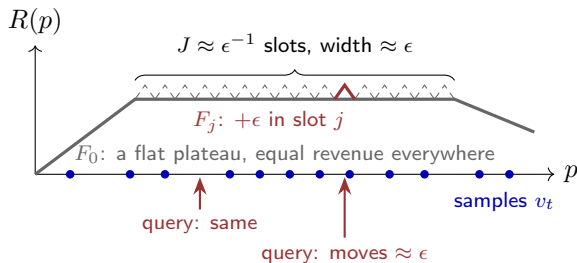
- Round t : only see $\mathbb{1}[v_t \geq p_t]$, which we call a **threshold query** at p_t
- Using the same p for n rounds: can est $F(p)$ within $\pm \frac{1}{\sqrt{n}}$, but **only at this p**
- Previously: one sample v_t answers the threshold queries at **all** prices at once

Threshold Complexity Threshold(ϵ)

Number of threshold queries needed to find a $\hat{p}$ with $\mathbb{E}[R(p^*) - R(\hat{p})] \leq \epsilon$

Threshold(ϵ) $\geq$ Sample(ϵ). How much larger? A hard instance $\rightarrow$

A Hard Instance: Where is the Needle?



Single prices

$\mathbb{1}[v \geq p]$: same under F_0, F_j
 unless $p \in \text{slot } j$, which only off by ϵ
 $\implies \epsilon^{-2}$ queries $\times \epsilon^{-1}$ slots
 $\implies \text{Threshold}(\epsilon) = \Omega(\epsilon^{-3})$

Full information

One v_t answers $\mathbb{1}[v_t \geq p]$ for all p
 $\implies \text{Sample}(\epsilon) = \Theta(\epsilon^{-2})$

Single prices $\text{Threshold}(\epsilon) = \tilde{\Theta}(\epsilon^{-3})$ [LSTW23] vs **Full info** $\text{Sample}(\epsilon) = \Theta(\epsilon^{-2})$

Menus

Video streaming

Basic (HD, with ads)	\$8
Standard (Full HD)	\$16
Premium (4K)	\$24

Flight

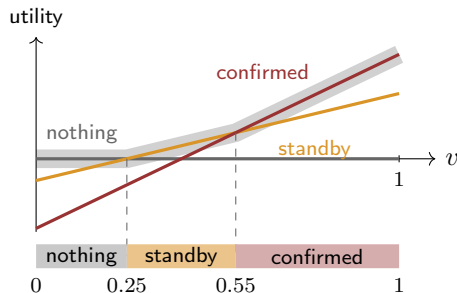
Confirmed business class	\$400
Bid-for-upgrade (50% chance)	\$250 if seated

$$(q, p) = (0.4, \$8), (0.7, \$16), (1, \$24)$$

$$(q, p) = (1, \$400), (0.5, \$125)$$

- **Option** (q, p) : get the item w.p. q (or a q -fraction), pay p in expectation
- **Menu**: a set of options. Customer picks the best (q, p) maximizing $vq - p$ (or nothing if they are all bad). Firm sees the chosen option (q, p) and earns revenue p
- Single price = $\{(1, p)\}$ (which we call a **size-1** menu)

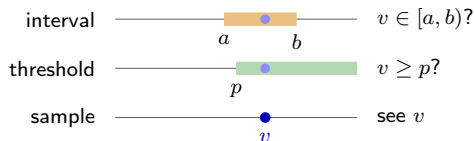
A Menu Splits Values into Intervals



- Option (q, p) : line $v \mapsto qv - p$ (slope q)
- Flight example: standby $\frac{1}{2}v - 0.125$, confirmed $v - 0.4$
- Value v picks the **highest line** at v
- Each option optimal on an **interval**
- With m options: firm now observes which interval v is located in
 - Single price: $[0, p)$ vs $[p, 1]$
 - 2 options: $[0, p_1)$ vs $[p_1, p_2)$ vs $[p_2, 1]$
 - Full information: “infinite intervals”

We Can Use Menus to Ask Interval Queries

- **Interval query** $[a, b) = \text{see } \mathbb{1}[v \in [a, b)]$
- **Threshold query** $p = \text{interval query } [0, p)$
- **Sample** = interval query at every $[v, v + \epsilon)$



Interval Complexity $\text{Interval}(\epsilon)$

Number of interval queries needed to find a $\hat{p}$ with $\mathbb{E}[R(p^*) - R(\hat{p})] \leq \epsilon$

Sample $\Theta(\epsilon^{-2}) \leq \text{Interval } ? \leq \text{Threshold } \tilde{\Theta}(\epsilon^{-3})$
 Interval = $\tilde{\mathcal{O}}(\epsilon^{-2})$ – as good as having full info!

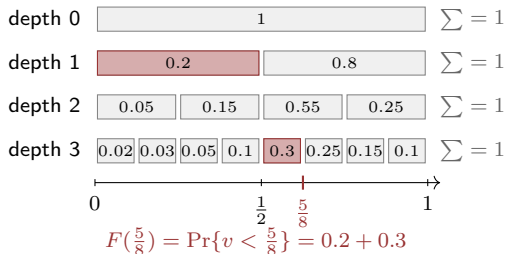
Results: Interval Queries Are as Good as Samples

Queries to learn p^* within ϵ	samples	$\asymp$	interval	threshold
General	ϵ^{-2}	$\asymp$	ϵ^{-2}	ϵ^{-3}
Regular	ϵ^{-2}	$\asymp$	ϵ^{-2}	ϵ^{-2}
MHR	$\epsilon^{-1.5}$	$\asymp$	$\epsilon^{-1.5}$	ϵ^{-2}
Curvature κ	$\kappa^{-1/2}\epsilon^{-1.5}$	$\asymp$	$\kappa^{-1/2}\epsilon^{-1.5}$	$\epsilon^{-2.5}$

- All tight up to logs. Prior work: [DKW56; HMR15; KL03; LSTW23; KSU19]
- Curvature $\kappa > 0$: $R(p^*) - R(p) \geq \kappa(p - p^*)^2$ for all $p \in [0, 1]$

Still one-bit feedback per customer, yet as good as revealing the full v !

Learning the CDF: Segment Tree + Empirical Bernstein



- Let $N = \epsilon^{-1}$. For all $k = 0, 1, \dots, N$, estimated $\hat{F}(k\epsilon)$ must be $F(k\epsilon) \pm \epsilon$ (threshold queries: $\Omega(\epsilon^{-2})$ for each k)
- Segment tree: $[0, \frac{k}{N}) = \leq \log_2 N$ tree intervals; estimate each within $\pm \frac{\epsilon}{\log_2 N}$
- Query I : coin with mean $\mu_I = \Pr\{v \in I\}$
- $\text{Var} \leq \text{mean} \Rightarrow$ **small mass, small noise**

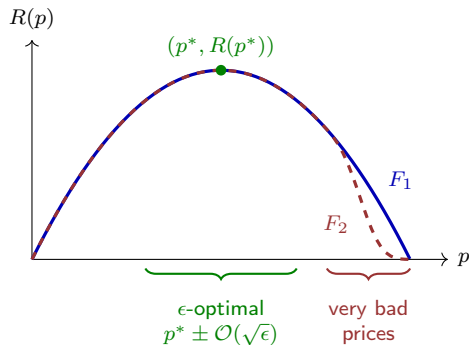
one tree interval I	one depth (disjoint: $\sum_I \mu_I \leq 1$)	all $\log_2 N$ depths
$\tilde{\mathcal{O}}(\frac{\mu_I}{\epsilon^2} + \frac{1}{\epsilon})$ [MP09]	$\tilde{\mathcal{O}}(\frac{1}{\epsilon^2} + \frac{N}{\epsilon}) = \tilde{\mathcal{O}}(\epsilon^{-2})$	$\tilde{\mathcal{O}}(\epsilon^{-2})$

$\text{Interval}(\epsilon) = \tilde{\mathcal{O}}(\epsilon^{-2})$: **as good as samples** (threshold queries: $\tilde{\Theta}(\epsilon^{-3})$ [LSTW23])

Learning the Whole CDF Is Wasteful

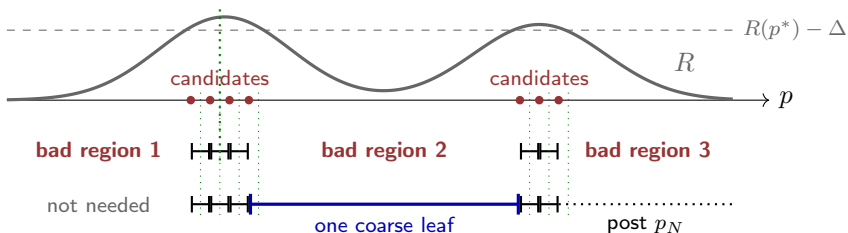
- Two similar unimodal F_1, F_2 (**MHR**, a nice class of distributions)
- Revenue: the same near p^* , different only at very bad prices
- To price: no need to tell F_1 from F_2
- To learn CDF: must tell them apart

Estimating the rest of F precisely is wasteful



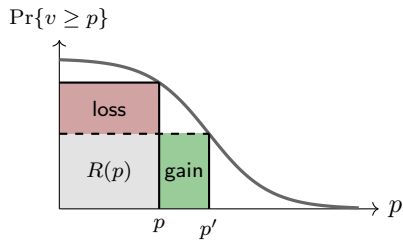
MHR: $\text{Sample}(\epsilon) = \tilde{\Theta}(\epsilon^{-1.5})$ [HMR15]. Can we match it using interval queries? $\rightarrow$

Zoom In without Cutting the Tree



- Layer $\Delta = \frac{1}{2}, \frac{1}{4}, \dots$: **candidates** $p_0 < \dots < p_N$ near-optimal ($R(p_i) \geq R(p^*) - 4\Delta$)
- Estimating $R(p) = p \Pr\{v \geq p\}$ needs the mass of **everything** above p
 $\implies$ cannot simply “cut away” those bad regions
- Instead: tree over candidates, **fine** near good prices, **coarse** in bad region
- **Estimate** candidates' R (up to $\pm\Delta$). **Refine** to half. **Eliminate** sub-optimal.

Key Property: Little Mass between Near-Optimal Prices



- $R(p)$ = area of the rectangle.
- Raise the price from p to p' : **gain**
 $(p' - p) \Pr\{v \geq p'\}$, **lose** $p \Pr\{p \leq v < p'\}$
- Both Δ -optimal $\implies$ **loss** $\leq$ **gain** $+\Delta$
 $\implies p \Pr\{p \leq v < p'\} \leq p' - p + \Delta$
 $\implies$ **little mass between near-opt prices**

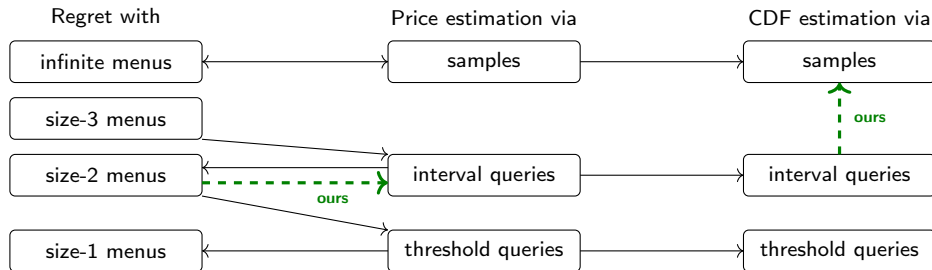
one interval $[p_a, p_b)$	one depth (telescoping)	layer Δ
$\tilde{O}\left(\frac{p_a \Pr\{p_a \leq v < p_b\}}{\Delta^2} + \frac{1}{\Delta}\right)$	$\tilde{O}\left(\frac{p_N - p_0 + \Delta}{\Delta^2}\right)$	$\tilde{O}\left(\frac{p_N - p_0 + \Delta}{\Delta^2}\right)$

General: $p_N - p_0 \leq 1 \implies \text{Interval}(\epsilon) = \tilde{O}(\epsilon^{-2})$

MHR: $p_N - p_0 = \mathcal{O}(\sqrt{\Delta}) \implies \text{Interval}(\epsilon) = \tilde{O}(\epsilon^{-1.5})$, **as good as samples**

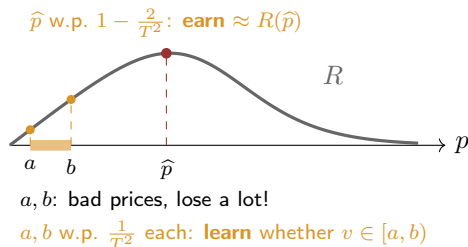
Back to Regret: Chains of Reductions

- $\mathfrak{R}_T^{(m)}$: best worst-case regret with $\leq m$ options (single price: $m = 1$)
- Arrows: black-box reductions; **dashed**: our algorithms. How? $\longrightarrow$



One Price to Earn, Two to Learn

- Segment tree asks $[a, b) =$ a menu with boundaries a, b
- A menu also **earns** the same as posting each boundary w.p. whatever we want (flight: \$250 or \$550, $\frac{1}{2}$ each)
- But a, b can both be bad (e.g., $[0, \frac{1}{N})$)
- Add best guess $\hat{p}$ as a 3rd option: $\hat{p}$ w.p. $1 - \frac{2}{T^2}$, a, b w.p. $\frac{1}{T^2}$ each



Repeat with $\epsilon = \frac{1}{2}, \frac{1}{4}, \dots$: $\tilde{O}(\epsilon^{-2})$ rounds, each losing $\mathcal{O}(\epsilon) \implies$ regret $\sum_{\epsilon} \tilde{O}(\epsilon^{-1}) = \tilde{O}(\sqrt{T})$

Size-3 menus: $\mathfrak{R}_T^{(3)} = \tilde{O}(\sqrt{T})$ for every F (same as full information)

Takeaway: A Slightly Worse Second Option

- Zooming only asks $[p_i, p_j)$ between candidates, so p_i is near-optimal (unlike the segment tree: $[0, \frac{1}{N})$): post p_i w.p. $1 - T^{-2}$, p_j w.p. $T^{-2} \implies$ **2 options**

Second option

get it w.p. $1 - T^{-2}$
pay p_i only if you get it

Main option

get it for sure
pay $(1 - T^{-2})p_i + T^{-2}p_j \approx p_i$



- $\implies$ Customers with $v \in [p_i, p_j)$ take second option: an **interval query** for free

Size 2 Regret: $\mathfrak{R}_T^{(2)} = \tilde{O}(\sqrt{T})$, $\tilde{O}(T^{1/3})$ for MHR. **Matching full-info** (up to logs)

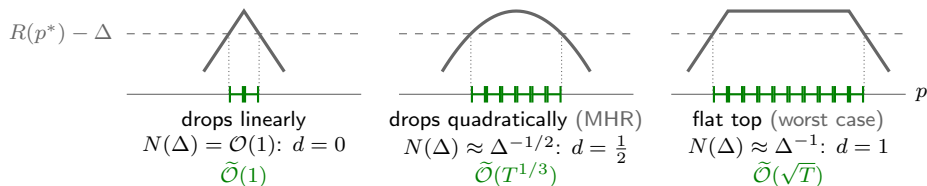
Regret: Two Options Suffice

Regret	size ∞ (full info)	$\asymp$	size 2	size 1 (single price)
General	$\sqrt{T}$	$\asymp$	$\sqrt{T}$	$T^{2/3}$
Regular	$\sqrt{T}$	$\asymp$	$\sqrt{T}$	$\sqrt{T}$
MHR	$T^{1/3}$	$\asymp$	$T^{1/3}$	$\sqrt{T}$
Curvature κ	$\kappa^{-1/3}T^{1/3}$	$\asymp$	$\kappa^{-1/3}T^{1/3}$	$T^{3/5}$
ZoomDim d	$T^{d/(d+1)}$	$\asymp$	$T^{d/(d+1)}$	$T^{(d+1)/(d+2)}$

- All tight up to logs. Prior work: [DKW56; HMR15; KL03; SW24; LSTW23; KSU19]
- ZoomDim d : Δ^{-d} intervals of length Δ cover all Δ -optimal prices. **Every single** F : size 2 attains $\tilde{O}(V_T)$, $V_T := \sup_{\Delta} \min\{T\Delta, N(\Delta)\}$, and even full info cannot beat $\Omega(V_T)$ locally near F (local minimax)

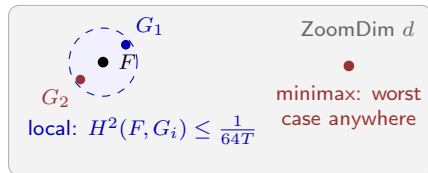
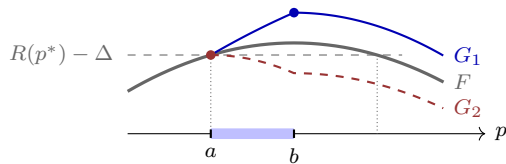
Extension: Instance-Optimality

- $N(\Delta)$: number of length- Δ intervals covering the Δ -optimal region
- **Zooming dimension** d : $N(\Delta) \approx \Delta^{-d}$ (as in Lipschitz bandits [KSU19])
- Candidates span $p_N - p_0 \lesssim \Delta N(\Delta) \implies \tilde{\mathcal{O}}(N(\Delta)/\Delta)$ rounds, each losing $\mathcal{O}(\Delta)$
- Halve Δ until $N(\Delta)/\Delta \approx T \implies$ regret $\tilde{\mathcal{O}}(T^{d/(d+1)})$



Every single F : size 2 gets $\tilde{\mathcal{O}}(T^{d/(d+1)})$, without knowing d

Lower Bound: Hard Instances Locally Around F

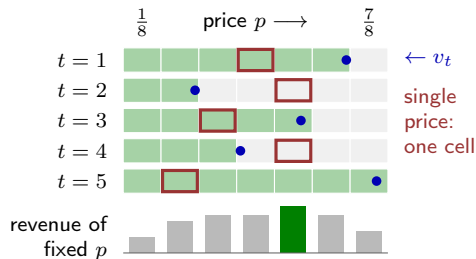


- On F alone, “always post p^* ” has 0 regret: must compare F with **neighbors**
- Half $[a, b)$ of the Δ -optimal interval: mass $\approx \Delta N(\Delta)$. Move masses $\Rightarrow G_1, G_2$
- Density on $[a, b)$: $\times(1 \pm \frac{1}{N(\Delta)}) \Rightarrow H^2(F, G_i) \approx \frac{\Delta}{N(\Delta)} \approx \frac{1}{T}$ if $T\Delta \approx N(\Delta)$
- T rounds can't tell G_1 from G_2 , yet every price is Δ -bad under one

Regret $\Omega(T\Delta) = \Omega(T^{d/(d+1)})$ on F, G_1 , or G_2 , even with full info

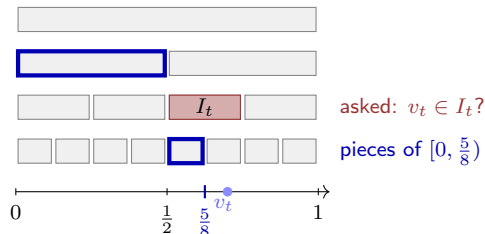
Extension: Adversarial Dynamic Pricing

- $v_1, \dots, v_T$ picked by an **adversary**
- Row t : prices customer t would buy at
- Benchmark: best **fixed** price p^* , defined as $\max_p \sum_t p \mathbb{1}[v_t \geq p]$
- With full info (**Hedge** [FS97]): post p w.p. $\propto \exp(\eta \cdot \text{total revenue of } p) \implies \tilde{O}(\sqrt{T})$
- Single price: $\tilde{\Theta}(T^{2/3})$ [KL03]



Size 3: Earn with Hedge, Learn Another Interval

- Menu: Hedge's p_t w.p. $1 - \frac{2}{T}$ (**earn**);
the two ends of a tree interval I_t w.p. $\frac{1}{T}$ each (**learn** whether $v_t \in I_t$)
- Row t : $\mathbb{1}[v_t \geq p] = 1 - \sum_I \mathbb{1}[v_t \in I]$
over the tree pieces I of $[0, p]$
- Guess $\mathbb{1}[v_t \in I]$ by $\frac{\mathbb{1}[v_t \in I]}{\Pr\{I_t = I\}}$ if $I_t = I$,
else 0 (importance-weighting)



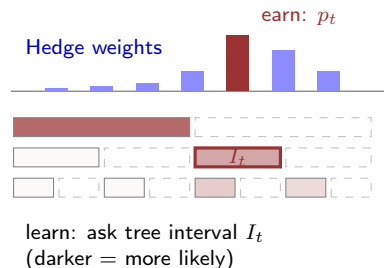
Unbiased revenue of **every** price: Hedge runs as if seeing v_t , but with noise
How much noise? Depends on **which** I_t we ask $\rightarrow$

Ask Where Hedge Looks: Small Noise

- Hedge's regret $\approx \sqrt{T \cdot \text{noise}}$, noise = mean square of the guesses under Hedge's weights
- Uniform I_t (e.g., [MN21]): noise up to #prices $\implies$ only $\tilde{O}(T^{2/3})$
- Ours: $\Pr\{I_t = I\} \propto$ Hedge's weight on p 's passing through I : cancels $\frac{1}{\Pr\{I_t = I\}}$
- v_t is in one tree interval per depth $\implies$ noise $\mathcal{O}(\log^2 T) \implies \tilde{O}(\sqrt{T})$ regret!

Size 3: $\tilde{O}(\sqrt{T})$ against adaptive adversaries, as good as seeing v_t

Also: $\tilde{O}(\sqrt{(S+1)T})$ against prices switching S times; limited supply [BDKS15]



Offering a slightly worse second option is (almost) as good as having full information

size 2 = size ∞ = full information, for every class and every single F

Thank you!

✉ yandai20@mit.edu 🌐 <https://yandaichn.github.io/>

References I

- [BDKS15] Moshe Babaioff, Shaddin Dughmi, Robert Kleinberg, and Aleksandrs Slivkins. “Dynamic Pricing with Limited Supply”. In: *ACM Transactions on Economics and Computation* 3.1 (2015).
- [DKW56] Aryeh Dvoretzky, Jack Kiefer, and Jacob Wolfowitz. “Asymptotic Minimax Character of the Sample Distribution Function and of the Classical Multinomial Estimator”. In: *The Annals of Mathematical Statistics* (1956), pp. 642–669.
- [FS97] Yoav Freund and Robert E Schapire. “A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting”. In: *Journal of Computer and System Sciences* 55.1 (1997), pp. 119–139.
- [HMR15] Zhiyi Huang, Yishay Mansour, and Tim Roughgarden. “Making the Most of Your Samples”. In: *Proceedings of the Sixteenth ACM Conference on Economics and Computation*. 2015, pp. 45–60.
- [KL03] Robert Kleinberg and Tom Leighton. “The Value of Knowing a Demand Curve: Bounds on Regret for Online Posted-Price Auctions”. In: *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science*. 2003, p. 594.
- [KSU19] Robert Kleinberg, Aleksandrs Slivkins, and Eli Upfal. “Bandits and Experts in Metric Spaces”. In: *Journal of the ACM (JACM)* 66.4 (2019), pp. 1–77.
- [LSTW23] Renato Paes Leme, Balasubramanian Sivan, Yifeng Teng, and Pratik Worah. “Pricing Query Complexity of Revenue Maximization”. In: *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. SIAM. 2023, pp. 399–415.
- [MN21] Michela Meister and Sloan Nietert. “Learning with Comparison Feedback: Online Estimation of Sample Statistics”. In: *Algorithmic Learning Theory*. PMLR. 2021, pp. 983–1001.

References II

- [MP09] Andreas Maurer and Massimiliano Pontil. "Empirical Bernstein Bounds and Sample-Variance Penalization". In: *Conference on Learning Theory*. 2009.
- [SW24] Sahil Singla and Yifan Wang. "Bandit Sequential Posted Pricing via Half-Concavity". In: *Proceedings of the 25th ACM Conference on Economics and Computation*. 2024, pp. 922–939.