Prophet Inequalities

Part 4: Data-driven prophet inequalities

Paul Dütting, Google Research
duetting@google.com

ADFOCS 2024 Summer School

August 2024

The Prophet Inequality Problem

- Given known distributions $\mathcal{D}_1, \mathcal{D}_2, \dots, \mathcal{D}_n$ over (non-negative) values:
  - A gambler gets to see realizations $v_i \sim \mathcal{D}_i$ one-by-one, and needs to immediately and irrevocable decide whether to accept v_i
  - The prophet sees the entire sequence of values $v_1, v_2, \dots, v_n$ at once, and can simply choose the maximum value
- **Question:** What's the worst-case gap between $\mathbb{E}[\text{value accepted by gambler}]$ and $\mathbb{E}[\text{value accepted by prophet}]$?

$\mathbb{E}[\text{value accepted by prophet}] = \mathbb{E}[\max_i v_i]$

$\mathbb{E}[\text{value accepted by gambler}] =: \mathbb{E}[ALG]$

A Data-Driven Approach

(Pioneered in [Azar Kleinberg Weinberg 2014])

Working in the same prophet setting, what can we achieve when the underlying distributions are unknown?

In particular, what can we do with limited access to the underlying distributions through samples?

Why Cool?

- Unlike in the setting with known distributions, it is not clear what optimal algorithms for this setting would look like
- Can we do any learning? What should we learn if we can learn something?
- How do the answers to these questions change with different amounts of information available?

Plan for Part 4

- A closer look at the results and techniques for the data-driven single-choice prophet inequality problem
  - Non-identical distributions
  - Identical distributions
- A brief discussion of state-of-the art for data-driven combinatorial prophet inequality problems

Outline Other Parts

Part 1: Introduction

Part 2: Online matching and contention resolution

Part 3: Online combinatorial auctions and balanced prices

Part 4: Data-driven prophet inequalities

Non-Identical Distributions

[Rubinstein Wang Weinberg 2020]

The Problem

(Single-Sample Prophet Inequality (SSPI) Problem)

samples known to gambler

- Given unknown distributions $\mathcal{D}_1, \mathcal{D}_2, \dots, \mathcal{D}_n$ over (non-negative) values, and samples $s_1 \sim \mathcal{D}_1, s_2 \sim \mathcal{D}_2, \dots, s_n \sim \mathcal{D}_n$:
  - A gambler gets to see realizations $v_i \sim \mathcal{D}_i$ one-by-one, and needs to immediately and irrevocable decide whether to accept v_i
  - The prophet sees the entire sequence of values $v_1, v_2, \dots, v_n$ at once, and can simply choose the maximum value
- **Question:** What's the worst-case gap between $\mathbb{E}[\text{value accepted by gambler}]$ and $\mathbb{E}[\text{value accepted by prophet}]$?

$\mathbb{E}[\text{value accepted by prophet}] = \mathbb{E}[\max_i v_i]$

$\mathbb{E}[\text{value accepted by gambler}] =: \mathbb{E}[ALG]$

Game of Googol

- An adversary determines n pairs of (non-negative) numbers $x_{i1}; x_{i2}$ for $i \in [n]$
- For each $i \in [n]$, we toss a fair coin to decide between
  - $V_i = x_{i1}; H_i = x_{i2}$ or
  - $V_i = x_{i2}; H_i = x_{i1}$

V: "visible", H: "hidden"

Game of Googol

- An adversary determines n pairs of (non-negative) numbers $x_{i1}; x_{i2}$ for $i \in [n]$
- For each $i \in [n]$, we toss a fair coin to decide between
  - $V_i = x_{i1}; H_i = x_{i2}$ or
  - $V_i = x_{i2}; H_i = x_{i1}$

V: “visible”, H: “hidden”

- A gambler gets to see $V_1, \ldots, V_n$, then observes the H_i one-by-one, and needs to immediately and irrevocably decide whether to accept H_i
- The prophet gets to see $H_1, \ldots, H_n$ at once, and chooses the maximum value

Game of Googol

- An adversary determines n pairs of (non-negative) numbers $x_{i1}; x_{i2}$ for $i \in [n]$
- For each $i \in [n]$, we toss a fair coin to decide between
  - $V_i = x_{i1}; H_i = x_{i2}$ or
  - $V_i = x_{i2}; H_i = x_{i1}$

V: "visible", H: "hidden"

- A gambler gets to see $V_1, \dots, V_n$, then observes the H_i one-by-one, and needs to immediately and irrevocably decide whether to accept H_i
- The prophet gets to see $H_1, \dots, H_n$ at once, and chooses the maximum value
- **Compare:** $\mathbb{E}[\text{value accepted by gambler}]$ to $\mathbb{E}[\text{value accepted by prophet}]$

  (where the expectation is over the random coin tosses)

Example

randomly assign x_{i1}, x_{i2} to V_i, H_i

$x_{11} = 2;\ x_{12} = 3$

$x_{21} = 9;\ x_{22} = 1$

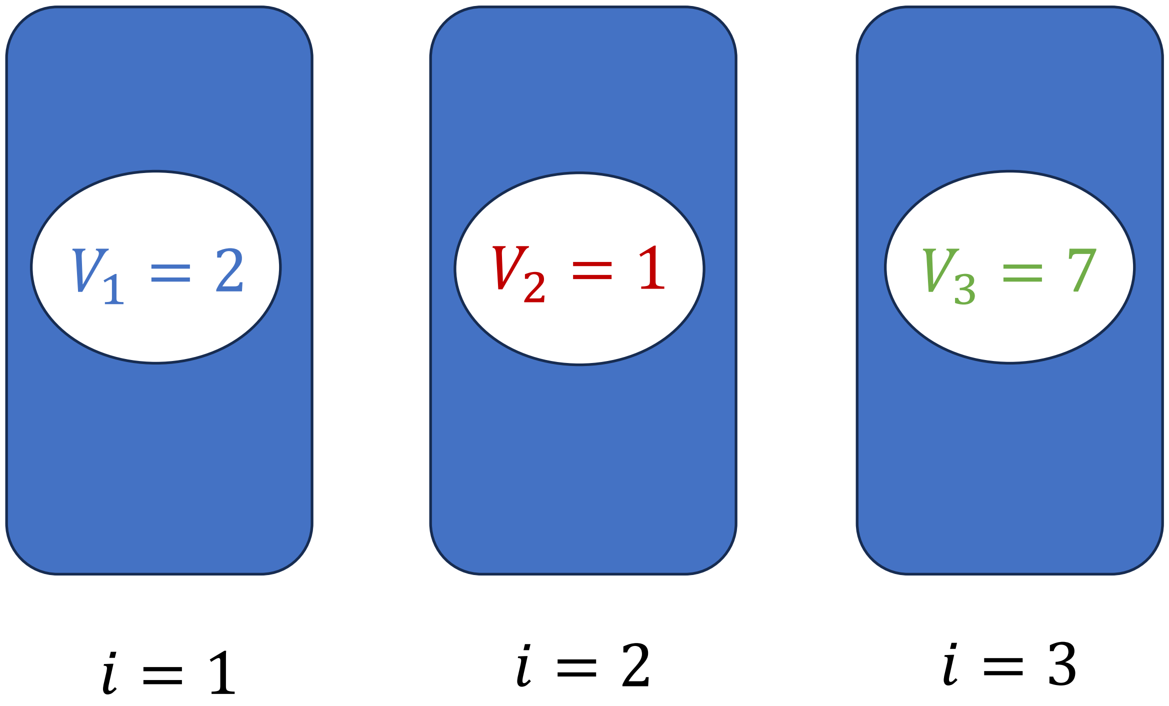
$x_{31} = 6;\ x_{32} = 7$

$V_1 = 2;\ H_1 = 3$

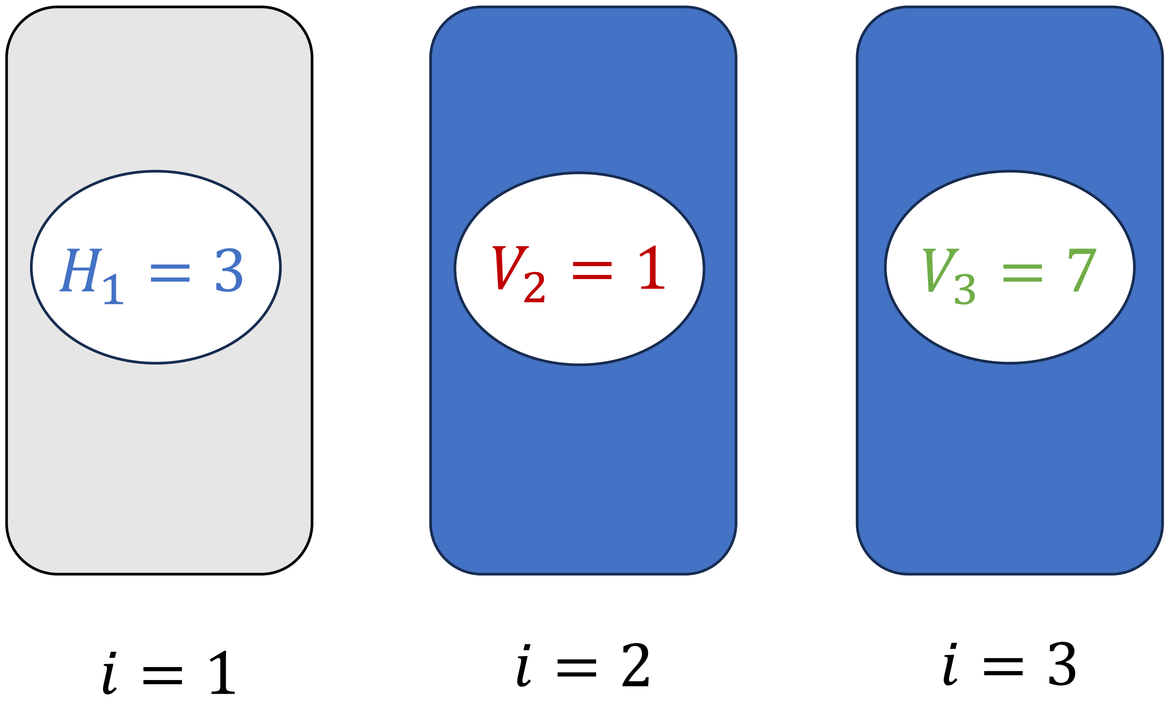
$H_2 = 9;\ V_2 = 1$

$H_2 = 6;\ V_2 = 7$

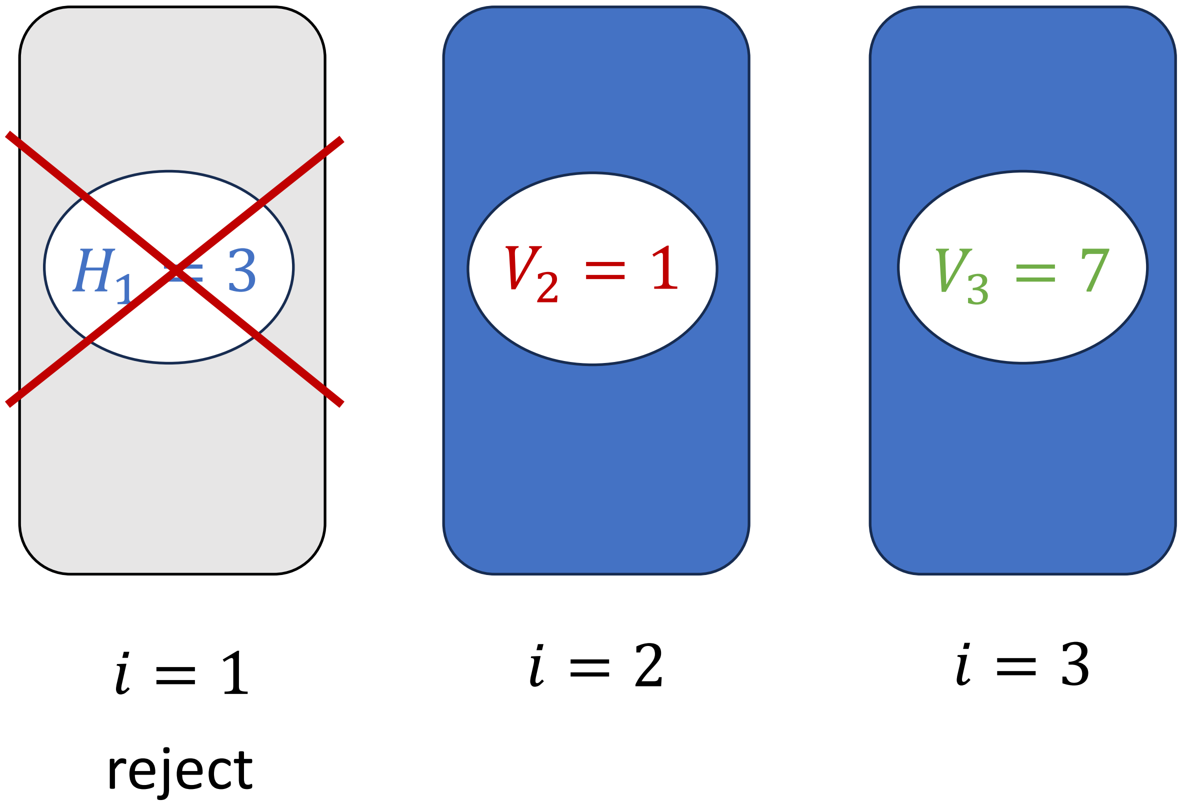
Example



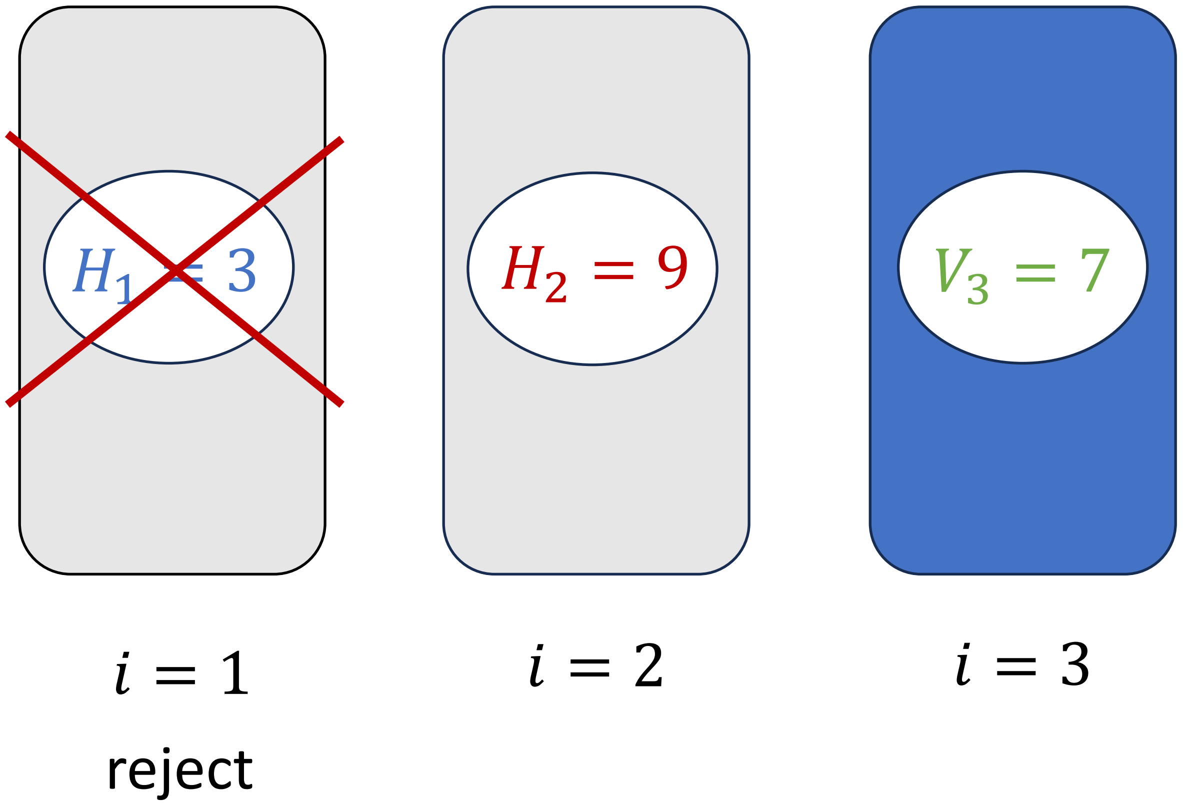
Example



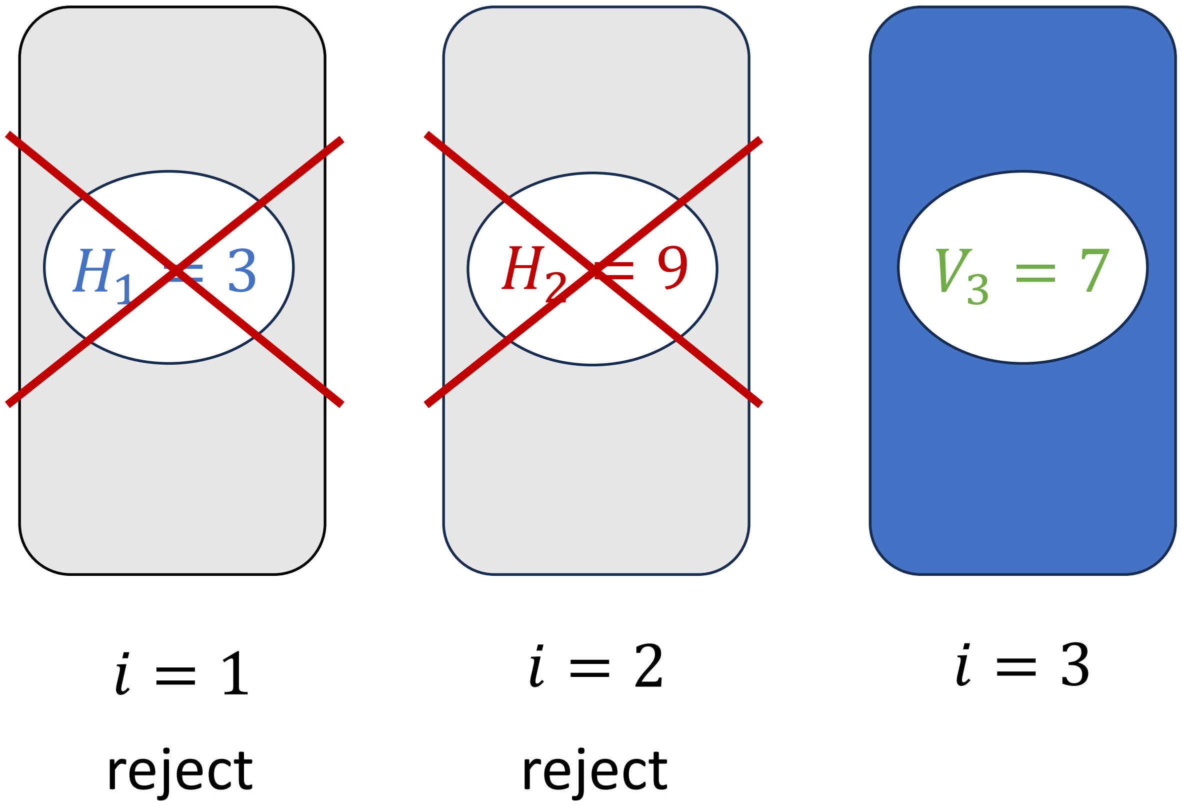
Example



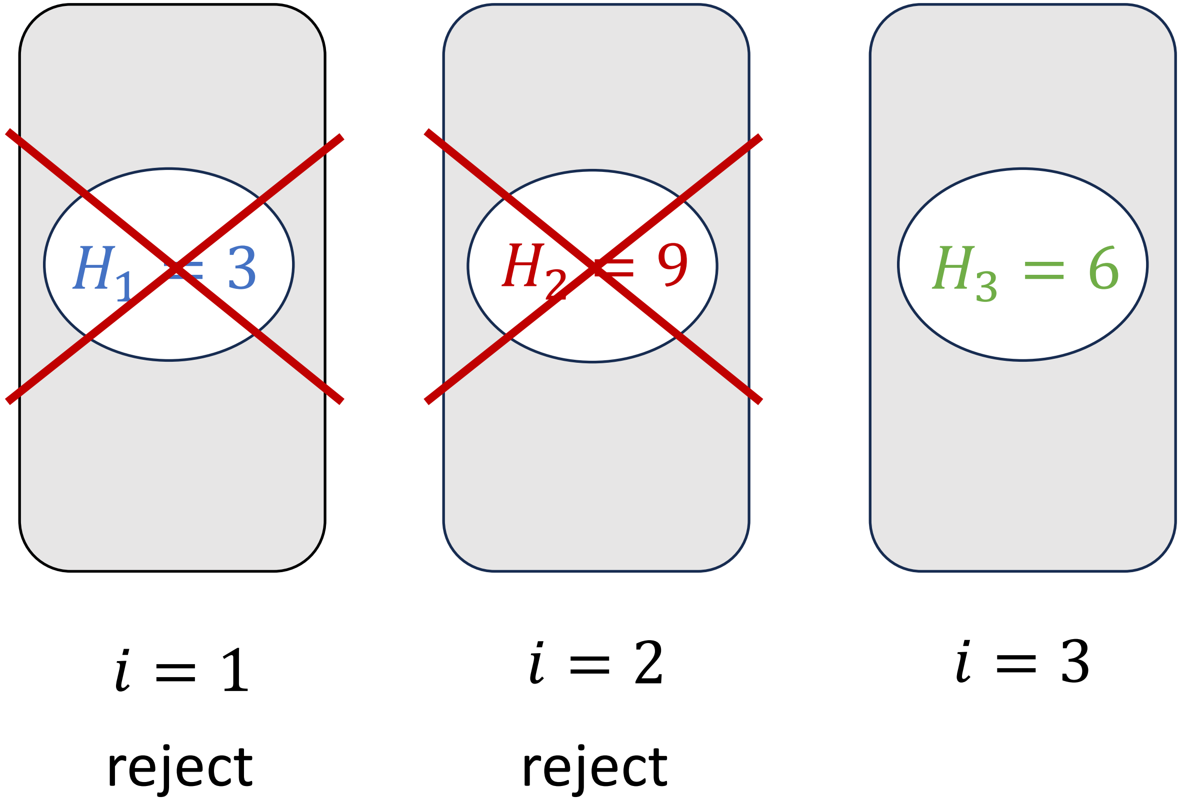
Example



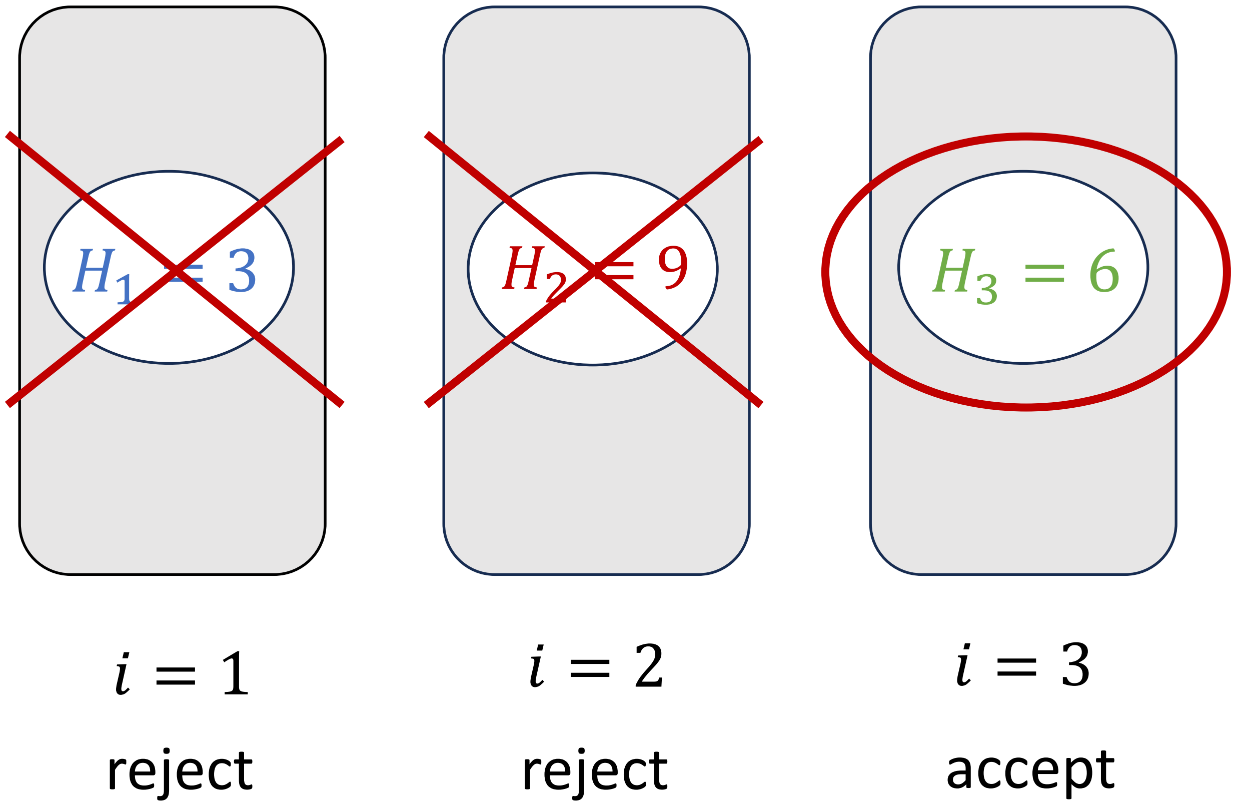
Example



Example

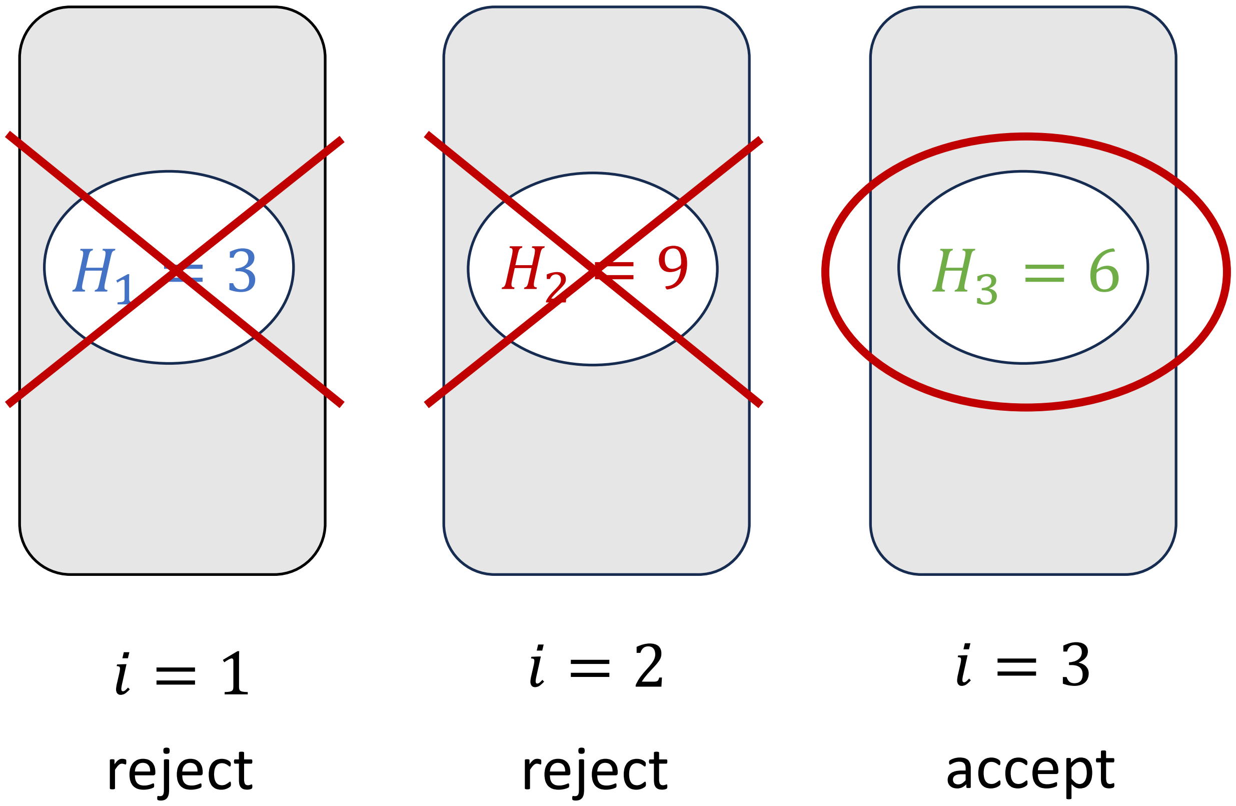


Example



Example

ALG = 6 vs. OPT = 9



Reduction

Observation: If we have a guarantee for the Game of Googol, that is for any (adversarially) chosen set of n pairs of numbers we get:

$$\mathbb{E}[ALG_\tau] \geq \frac{1}{\alpha} \cdot \mathbb{E}[\max_i H_i]$$

then we also get an α-competitive Single-Sample Prophet Inequality (SSPI).

The reduction:

- The sequence $v_1, \dots, v_n$ that is revealed online plays the role of $H_1, .., H_n$
- The independent samples $s_1, \dots, s_n$ play the role of $V_1, .., V_n$

The Result

Theorem [Rubinstein Wang Weinberg 2020]

In the Game of Googol, setting a threshold of $\tau = \max_i V_i$ and accepting the first H_i such that $H_i \geq \tau$ ensures that

$$\mathbb{E}[ALG_\tau] \geq \frac{1}{2}\,\mathbb{E}[\max_i H_i]\,.$$

(i.e., we can achieve the optimal factor 2 of the original prophet inequality problem with a single sample from each distribution!)

Analysis: Notation

- **Recall:** We have fixed $x_{i,j}$ for $i \in n$, $j \in \{1,2\}$ (numbers chosen by adversary)

  w.l.o.g. assume that the $x_{i,j}$ are all distinct

Analysis: Notation

- **Recall:** We have fixed $x_{i,j}$ for $i \in n$, $j \in \{1,2\}$ (numbers chosen by adversary)

  w.l.o.g. assume that the $x_{i,j}$ are all distinct

- Let's sort the $x_{i,j}$ so that: $W_1 > W_2 > \ \dots \ > W_{2n}$
- Define: pivotal index $j^\star \in \{1, \dots, n+1\}$:
  - Going left to right in the W_i sequence, this is the first time we see the second number of a pair

Note: Irrespective of the coin tosses,

$OPT = W_j$ for some $j \leq j^\star$

Analysis: Notation

- **Recall:** We have fixed $x_{i,j}$ for $i \in n, j \in \{1,2\}$ (numbers chosen by adversary)

  w.l.o.g. assume that the $x_{i,j}$ are all distinct

- Let's sort the $x_{i,j}$ so that: $W_1 > W_2 > \ldots > W_{2n}$
- Define: pivotal index $j^\star \in \{1, \ldots, n+1\}$:
  - Going left to right in the W_i sequence, this is the first time we see the second number of a pair

W_1	W_2	W_3	W_4	W_5	W_6
9	7	6	3	2	1

$j^\star = 3$

Note: Irrespective of the coin tosses,

$OPT = W_j$ for some $j \leq j^\star$

Analysis: Formula for OPT

Lemma [Rubinstein, Wang, Weinberg 2020]

It holds that

$$\underbrace{\mathbb{E}[OPT]}_{= \max_i H_i} = \left(\sum_{j=1}^{j^\star - 1} \frac{W_j}{2^j} \right) + \frac{W_{j^\star}}{2^{j^\star - 1}} \ .$$

Analysis: Formula for OPT

Lemma [Rubinstein, Wang, Weinberg 2020]

It holds that

$$\underbrace{\mathbb{E}[OPT]}_{=\max_i H_i} = \left(\sum_{j=1}^{j^\star-1} \frac{W_j}{2^j}\right) + \frac{W_{j^\star}}{2^{j^\star-1}} .$$

Note: It suffices to show that

(1) For $j \leq j^\star - 1$: $\Pr[OPT = W_j] = 1/2^j$

(2) For $j = j^\star$: $\Pr[OPT = W_j] = 1/2^{j-1}$

Proof: Formula for OPT

$j^\star = 3$

W_1	W_2	W_3	W_4	W_5	W_6
9	7	6	3	2	1

$= W_1$							
$= W_2$							
$= W_3$							
OPT	sequences for which this is the case						$\Pr[\text{these sequences}]$

Proof: Formula for OPT

$j^\star = 3$

W_1	W_2	W_3	W_4	W_5	W_6
9	7	6	3	2	1

$= W_1$	hidden	*	*	*	*	*	½
$= W_2$							
$= W_3$							

OPT — sequences for which this is the case — $\Pr[\text{these sequences}]$

Proof: Formula for OPT

$j^\star = 3$

W_1	W_2	W_3	W_4	W_5	W_6
9	7	6	3	2	1

OPT	sequences for which this is the case						Pr[these sequences]
$= W_1$	hidden	*	*	*	*	*	½
$= W_2$	visible	hidden	*	*	*	*	½ · ½
$= W_3$							

Proof: Formula for OPT

$j^\star = 3$

W_1	W_2	W_3	W_4	W_5	W_6
9	7	6	3	2	1

$= W_1$	hidden	*	*	*	*	*	½
$= W_2$	visible	hidden	*	*	*	*	½ · ½
$= W_3$	visible	visible	hidden	*	*	*	½ · ½

OPT sequences for which this is the case Pr[these sequences]

Analysis: Lower Bound for ALG

Lemma [Rubinstein, Wang, Weinberg 2020]

It holds that

$$\mathbb{E}[ALG_\tau] \geq \left(\sum_{j=1}^{j^\star - 2} \frac{W_j}{2^{j+1}} \right) + \frac{W_{j^\star - 1}}{2^{j^\star - 1}} .$$

Analysis: Lower Bound for ALG

Lemma [Rubinstein, Wang, Weinberg 2020]

It holds that

$$\mathbb{E}[ALG_\tau] \geq \left(\sum_{j=1}^{j^\star-2} \frac{W_j}{2^{j+1}} \right) + \frac{W_{j^\star-1}}{2^{j^\star-1}} .$$

Recall:

$$\mathbb{E}[OPT] = \left(\sum_{j=1}^{j^\star-1} \frac{W_j}{2^j} \right) + \frac{W_{j^\star}}{2^{j^\star-1}}$$

Comparison:

1. For $j \leq j^\star - 2$: Get W_j w.p. $\frac{1}{2^{j+1}}$ instead of $\frac{1}{2^j}$
2. For $j = j^\star - 1$: Get W_j w.p. $\frac{1}{2^j}$ (as before)
3. For $j = j^\star$: Get W_j w.p. 0 instead of $\frac{1}{2^{j-1}}$

Proof: Lower Bound for ALG

$j^\star = 3$

W_1	W_2	W_3	W_4	W_5	W_6
9	7	6	3	2	1

$= W_1$	hidden	*	*	*	*	*	½
$= W_2$	visible	hidden	*	*	*	*	½ · ½
$= W_3$	visible	visible	hidden	*	*	*	½ · ½

OPT sequences for which this is the case   Pr[these sequences]

Proof: Lower Bound for ALG

$j^\star = 3$

	W_1	W_2	W_3	W_4	W_5	W_6	
	9	7	6	3	2	1	
$= W_1$	hidden	*	*	*	*	*	½
$= W_2$	visible	hidden	*	*	*	*	½ · ½
$= W_3$	visible	visible	hidden	*	*	*	½ · ½
OPT	sequences for which this is the case						Pr[these sequences]

Proof: Lower Bound for ALG

$j^\star = 3$

	W_1	W_2	W_3	W_4	W_5	W_6	
	9	7	6	3	2	1	
$= W_1$	visible	*	*	*	*	*	½
$= W_2$	hidden	visible	*	*	*	*	½ · ½
$= W_3$	hidden	hidden	visible	*	*	*	½ · ½

OPT sequences for which this is the case   Pr[these sequences]

Proof: Lower Bound for ALG

$j^\star = 3$

W_1	W_2	W_3	W_4	W_5	W_6
9	7	6	3	2	1

ALG_τ	sequences for which this is the case						Pr[these sequences]
$= 0$	visible	*	*	*	*	*	½
$= W_1$	hidden	visible	*	*	*	*	½ · ½
$\geq W_2$	hidden	hidden	visible	*	*	*	½ · ½

Putting Everything Together

Proof:
(of the theorem)

$$\mathbb{E}[ALG_\tau] \geq \left(\sum_{j=1}^{j^\star-2} \frac{W_j}{2^{j+1}} \right) + \frac{W_{j^\star-1}}{2^{j^\star-1}}$$

(by previous lemma)

Putting Everything Together

Proof:
(of the theorem)

$$\mathbb{E}[ALG_\tau] \geq \left(\sum_{j=1}^{j^\star-2} \frac{W_j}{2^{j+1}}\right) + \frac{W_{j^\star-1}}{2^{j^\star-1}}$$

(by previous lemma)

Recall:

$$\mathbb{E}[OPT] = \left(\sum_{j=1}^{j^\star-1} \frac{W_j}{2^j}\right) + \frac{W_{j^\star}}{2^{j^\star-1}}$$

Putting Everything Together

Proof:
(of the theorem)

$$\mathbb{E}[ALG_\tau] \geq \left(\sum_{j=1}^{j^\star - 2} \frac{W_j}{2^{j+1}} \right) + \frac{W_{j^\star - 1}}{2^{j^\star - 1}}$$

(by previous lemma)

$$= \frac{1}{2} \cdot \left(\sum_{j=1}^{j^\star - 2} \frac{W_j}{2^j} \right) + \frac{W_{j^\star - 1}}{2^{j^\star}} + \frac{W_{j^\star - 1}}{2^{j^\star}}$$

Recall:

$$\mathbb{E}[OPT] = \left(\sum_{j=1}^{j^\star - 1} \frac{W_j}{2^j} \right) + \frac{W_{j^\star}}{2^{j^\star - 1}}$$

Putting Everything Together

Proof:
(of the theorem)

$$\mathbb{E}[ALG_\tau] \geq \left(\sum_{j=1}^{j^\star-2} \frac{W_j}{2^{j+1}}\right) + \frac{W_{j^\star-1}}{2^{j^\star-1}}$$

(by previous lemma)

$$= \frac{1}{2} \cdot \left(\sum_{j=1}^{j^\star-2} \frac{W_j}{2^j}\right) + \frac{1}{2} \cdot \frac{W_{j^\star-1}}{2^{j^\star-1}} + \frac{W_{j^\star-1}}{2^{j^\star}}$$

Recall:

$$\mathbb{E}[OPT] = \left(\sum_{j=1}^{j^\star-1} \frac{W_j}{2^j}\right) + \frac{W_{j^\star}}{2^{j^\star-1}}$$

Putting Everything Together

Proof:
(of the theorem)

$$\mathbb{E}[ALG_\tau] \geq \left(\sum_{j=1}^{j^\star-2} \frac{W_j}{2^{j+1}}\right) + \frac{W_{j^\star-1}}{2^{j^\star-1}}$$

(by previous lemma)

$$= \frac{1}{2} \cdot \left(\sum_{j=1}^{j^\star-2} \frac{W_j}{2^j}\right) + \frac{1}{2} \cdot \frac{W_{j^\star-1}}{2^{j^\star-1}} + \frac{W_{j^\star-1}}{2^{j^\star}}$$

$$\geq \frac{1}{2} \cdot \left(\sum_{j=1}^{j^\star-2} \frac{W_j}{2^j}\right) + \frac{1}{2} \cdot \frac{W_{j^\star-1}}{2^{j^\star-1}} + \frac{W_{j^\star}}{2^{j^\star}}$$

Recall:
$$\mathbb{E}[OPT] = \left(\sum_{j=1}^{j^\star-1} \frac{W_j}{2^j}\right) + \frac{W_{j^\star}}{2^{j^\star-1}}$$

Putting Everything Together

Proof:
(of the theorem)

$$\mathbb{E}[ALG_\tau] \geq \left(\sum_{j=1}^{j^\star-2} \frac{W_j}{2^{j+1}}\right) + \frac{W_{j^\star-1}}{2^{j^\star-1}}$$ (by previous lemma)

$$= \frac{1}{2}\cdot\left(\sum_{j=1}^{j^\star-2} \frac{W_j}{2^j}\right) + \frac{1}{2}\cdot\frac{W_{j^\star-1}}{2^{j^\star-1}} + \frac{W_{j^\star-1}}{2^{j^\star}}$$

$$\geq \frac{1}{2}\cdot\left(\sum_{j=1}^{j^\star-2} \frac{W_j}{2^j}\right) + \frac{1}{2}\cdot\frac{W_{j^\star-1}}{2^{j^\star-1}} + \frac{1}{2}\cdot\frac{W_{j^\star}}{2^{j^\star-1}}$$

Recall:

$$\mathbb{E}[OPT] = \left(\sum_{j=1}^{j^\star-1} \frac{W_j}{2^j}\right) + \frac{W_{j^\star}}{2^{j^\star-1}}$$

Putting Everything Together

Proof:
(of the theorem)

$$\mathbb{E}[ALG_\tau] \geq \left(\sum_{j=1}^{j^\star-2} \frac{W_j}{2^{j+1}}\right) + \frac{W_{j^\star-1}}{2^{j^\star-1}}$$

(by previous lemma)

$$= \frac{1}{2} \cdot \left(\sum_{j=1}^{j^\star-2} \frac{W_j}{2^j}\right) + \frac{1}{2} \cdot \frac{W_{j^\star-1}}{2^{j^\star-1}} + \frac{W_{j^\star-1}}{2^{j^\star}}$$

$$\geq \frac{1}{2} \cdot \left(\sum_{j=1}^{j^\star-2} \frac{W_j}{2^j}\right) + \frac{1}{2} \cdot \frac{W_{j^\star-1}}{2^{j^\star-1}} + \frac{1}{2} \cdot \frac{W_{j^\star}}{2^{j^\star-1}}$$

$$\geq \frac{1}{2} \cdot \left(\sum_{j=1}^{j^\star-1} \frac{W_j}{2^j}\right) + \frac{1}{2} \cdot \frac{W_{j^\star}}{2^{j^\star-1}}$$

Recall: $$\mathbb{E}[OPT] = \left(\sum_{j=1}^{j^\star-1} \frac{W_j}{2^j}\right) + \frac{W_{j^\star}}{2^{j^\star-1}}$$

Putting Everything Together

Proof:
(of the theorem)

$$\mathbb{E}[ALG_\tau] \geq \left(\sum_{j=1}^{j^\star-2} \frac{W_j}{2^{j+1}}\right) + \frac{W_{j^\star-1}}{2^{j^\star-1}}$$

(by previous lemma)

$$= \frac{1}{2} \cdot \left(\sum_{j=1}^{j^\star-2} \frac{W_j}{2^j}\right) + \frac{1}{2} \cdot \frac{W_{j^\star-1}}{2^{j^\star-1}} + \frac{W_{j^\star-1}}{2^{j^\star}}$$

$$\geq \frac{1}{2} \cdot \left(\sum_{j=1}^{j^\star-2} \frac{W_j}{2^j}\right) + \frac{1}{2} \cdot \frac{W_{j^\star-1}}{2^{j^\star-1}} + \frac{1}{2} \cdot \frac{W_{j^\star}}{2^{j^\star-1}}$$

$$\geq \frac{1}{2} \cdot \left(\sum_{j=1}^{j^\star-1} \frac{W_j}{2^j}\right) + \frac{1}{2} \cdot \frac{W_{j^\star}}{2^{j^\star-1}} = \frac{1}{2} \cdot \mathbb{E}[OPT]$$

Recall: $\mathbb{E}[OPT] = \left(\sum_{j=1}^{j^\star-1} \frac{W_j}{2^j}\right) + \frac{W_{j^\star}}{2^{j^\star-1}}$

Q.E.D.

Technique: Deferred Decisions

- What we did in the analysis is to defer the decision whether a certain number is a value or sample (i..e., whether it is H(idden) or V(isible)) until we reached its position
- Upon reaching a position we flipped a fair coin
- Coin flips for $j < j^\star$ are independent, outcome on $j^\star$ is deterministic given previous coin tosses

Summary

- To obtain the optimal factor 2 for known distributions, we only need a single sample from each distribution (!)
- The Game of Googol reduction, and the deferred decisions technique are useful more generally

Identical Distributions

[Correa Dütting Schewior Fischer 2019, Rubinstein Wang Weinberg 2020]

(and lots of follow-up work)

The Problem

samples known to gambler

- Given an unknown distribution $\mathcal{D}$ over (non-negative) values and limited access to $\mathcal{D}$ through k samples $s_1 \sim \mathcal{D}, \ldots, s_k \sim \mathcal{D}$:
  - A gambler gets to see realizations $v_i \sim \mathcal{D}$ one-by-one, and needs to immediately and irrevocable decide whether to accept v_i
  - The prophet sees the entire sequence of values $v_1, v_2, \ldots, v_n$ at once, and can simply choose the maximum value
- **Question:** What's the worst-case gap between $\mathbb{E}$[value accepted by gambler] and $\mathbb{E}$[value accepted by prophet]?

$= \mathbb{E}[\max_i v_i]$

$=: \mathbb{E}[ALG]$

Solution for Known Distribution

Theorem [Hill Kertz '82, Correa-Foncea-Hoeksma-Oosterwijk-Vredeveld '17]

For every known distribution $\mathcal{D}$, and n draws $v_i \sim \mathcal{D}$ there exists an algorithm ALG such that

$$\mathbb{E}[ALG] \geq 0.745 \cdot \mathbb{E}[\max_i v_i] \quad ,$$

and this is best possible.

- There is a sequence of increasing "quantiles" $q_1 \leq q_2 \leq \ldots \leq q_n$ (independent of the distribution)
- The algorithm sets a sequence of decreasing thresholds $\tau_1 \geq \tau_2 \geq \cdots \geq \tau_n$ where $\Pr[v_i \geq \tau_i] = q_i$, and accepts the first $v_i \geq \tau_i$

Question

Which fraction of $\mathbb{E}[\max_i v_i]$ can we achieve with k samples?

A Baseline

Proposition [Correa Dütting Schewior Fischer 2019]

There exists an algorithm *ALG* that requires no samples, and achieves

$$\mathbb{E}[ALG] \geq \frac{1}{e}\,\mathbb{E}[\max_i v_i]\ .$$

A Baseline

Proposition [Correa Dütting Schewior Fischer 2019]

There exists an algorithm *ALG* that requires no samples, and achieves

$$\mathbb{E}[ALG] \geq \frac{1}{e}\, \mathbb{E}[\max_i v_i] \ .$$

Proof: Simple reduction to Secretary Problem.

Algorithm: Skip the first $\approx n/e$ values, accept the first value in the remainder that is larger than the max in this prefix.

Main Result

Theorem [Correa Dütting Schewior Fischer 2019]

For any $\delta > 0$ and any algorithm ALG without samples, there exists a distribution $\mathcal{D}$ such that

$$\mathbb{E}[ALG] \leq \left(\frac{1}{e} + \delta\right) \mathbb{E}[\max_i v_i] \, .$$

Main Result

Theorem [Correa Dütting Schewior Fischer 2019]

For any $\delta > 0$ and any algorithm *ALG* without samples, there exists a distribution $\mathcal{D}$ such that

$$\mathbb{E}[ALG] \leq \left(\frac{1}{e} + \delta\right) \mathbb{E}[\max_i v_i]\,.$$

(continues to hold if *ALG* has access to $o(n)$ samples)

Proof Strategy

- Establish the existence of an infinite-size subset of the natural numbers on which the given algorithm is (essentially) rank-based
- Define a distribution on a finite subset of this set
- Argue that if we could get $\frac{1}{\alpha} > \frac{1}{e}$, then we would get a better algorithm for the Secretary Problem

Key Structural Lemma

Lemma [Correa Dütting Schewior Fischer 2019]

For every $\epsilon > 0$ and every algorithm ALG without samples, there exists an infinite-size set $S \subseteq \mathbb{N}$ with the following property:

For all steps $i \in [n]$ there exists a probability $p_i \in [0,1]$ such that for all distinct values $v_1, \dots, v_i \in S$ seen until then,

$$\Pr[ALG \text{ accepts } v_i \mid v_i = \max\{v_1, \dots, v_i\}] \in [p_i - \epsilon,\ p_i + \epsilon] .$$

Key Structural Lemma

Lemma [Correa Dütting Schewior Fischer 2019]

For every $\epsilon > 0$ and every algorithm ALG without samples, there exists an infinite-size set $S \subseteq \mathbb{N}$ with the following property:

For all steps $i \in [n]$ there exists a probability $p_i \in [0,1]$ such that for all distinct values $v_1, \dots, v_i \in S$ seen until then,

$$\Pr[ALG \text{ accepts } v_i \mid v_i = \max\{v_1, \dots, v_i\}] \in [p_i - \epsilon,\ p_i + \epsilon]\,.$$

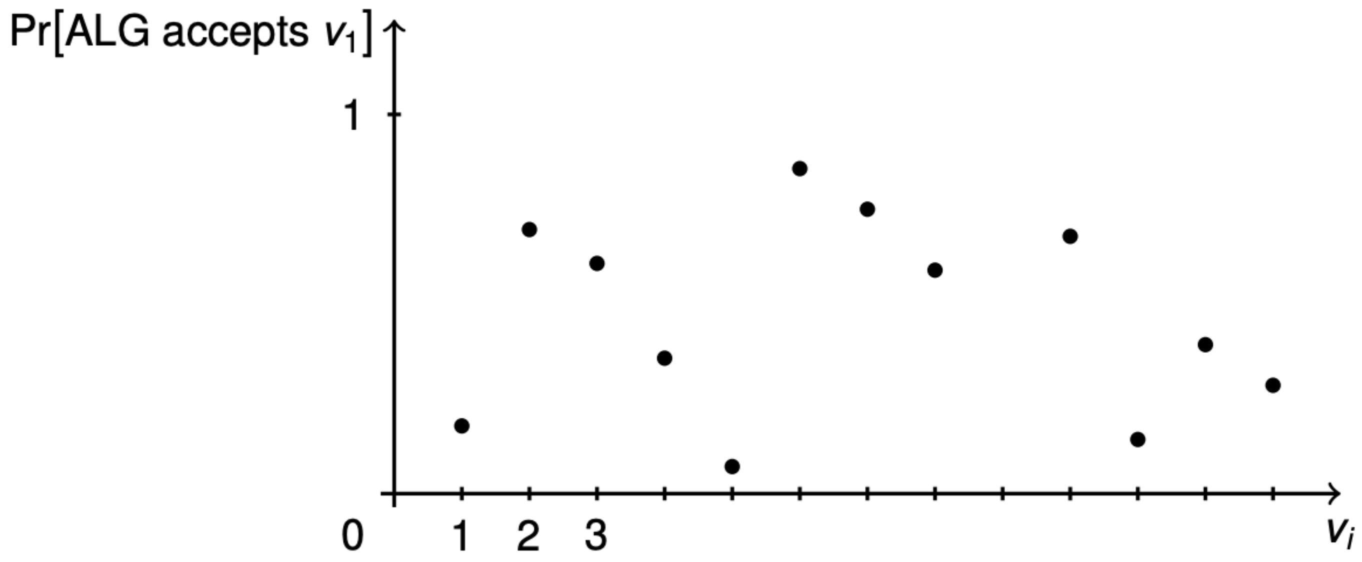
Proof plan: Fixing $\epsilon > 0$ and ALG, establish the existence of $S_1 \supseteq S_2 \supseteq \dots \supseteq S_n$ such that S_i satisfies the property for all $i' \leq i$

Establishing the Property

For all steps $i \in [n]$ there exists a probability $p_i \in [0,1]$ such that for all distinct values $v_1, \ldots, v_i \in S$ seen until then,

$$\Pr[ALG \text{ accepts } v_i \mid v_i = \max\{v_1, \ldots, v_i\}] \in [p_i - \epsilon,\ p_i + \epsilon]\ .$$

For $i = 1$:

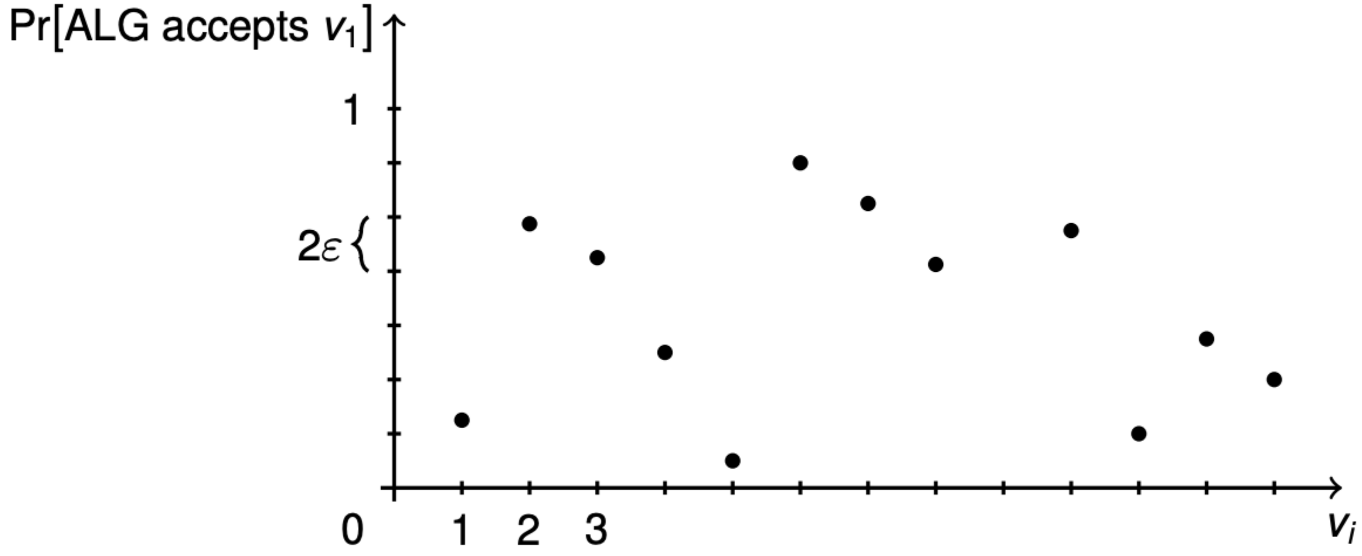


Establishing the Property

For all steps $i \in [n]$ there exists a probability $p_i \in [0,1]$ such that for all distinct values $v_1, \dots, v_i \in S$ seen until then,

$$\Pr[ALG \text{ accepts } v_i \mid v_i = \max\{v_1, \dots, v_i\}] \in [p_i - \epsilon,\ p_i + \epsilon]\,.$$

For $i = 1$:

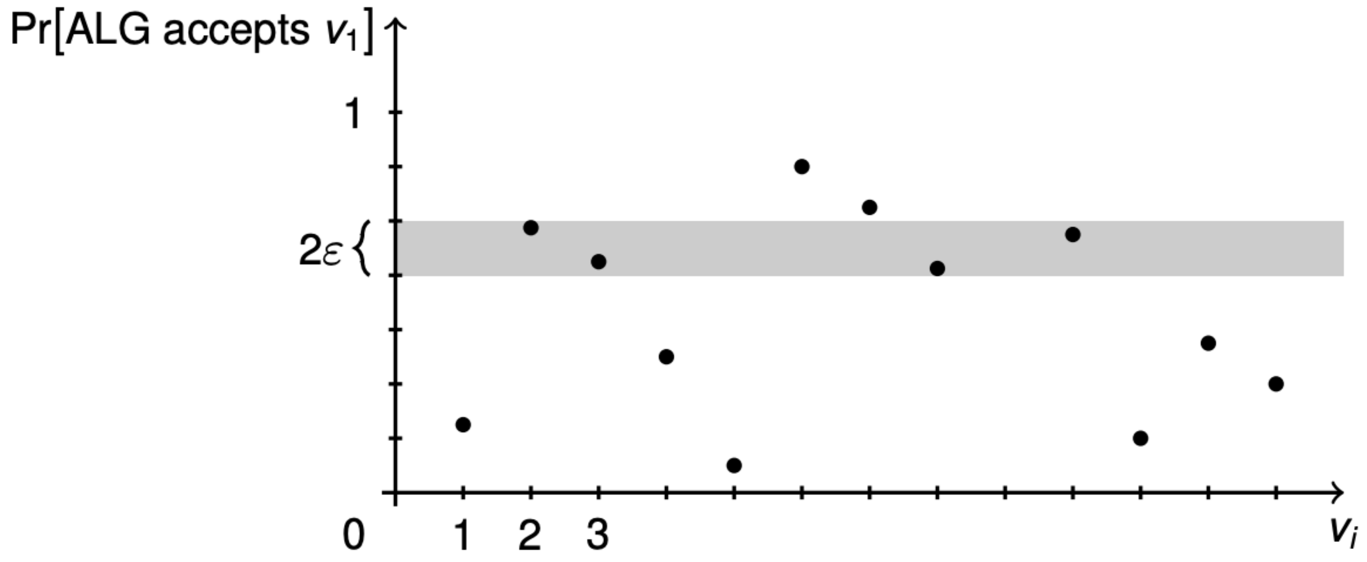


Establishing the Property

For all steps $i \in [n]$ there exists a probability $p_i \in [0,1]$ such that for all distinct values $v_1, \dots, v_i \in S$ seen until then,

$$\Pr[ALG \text{ accepts } v_i \mid v_i = \max\{v_1, \dots, v_i\}] \in [p_i - \epsilon,\ p_i + \epsilon].$$

For $i = 1$:

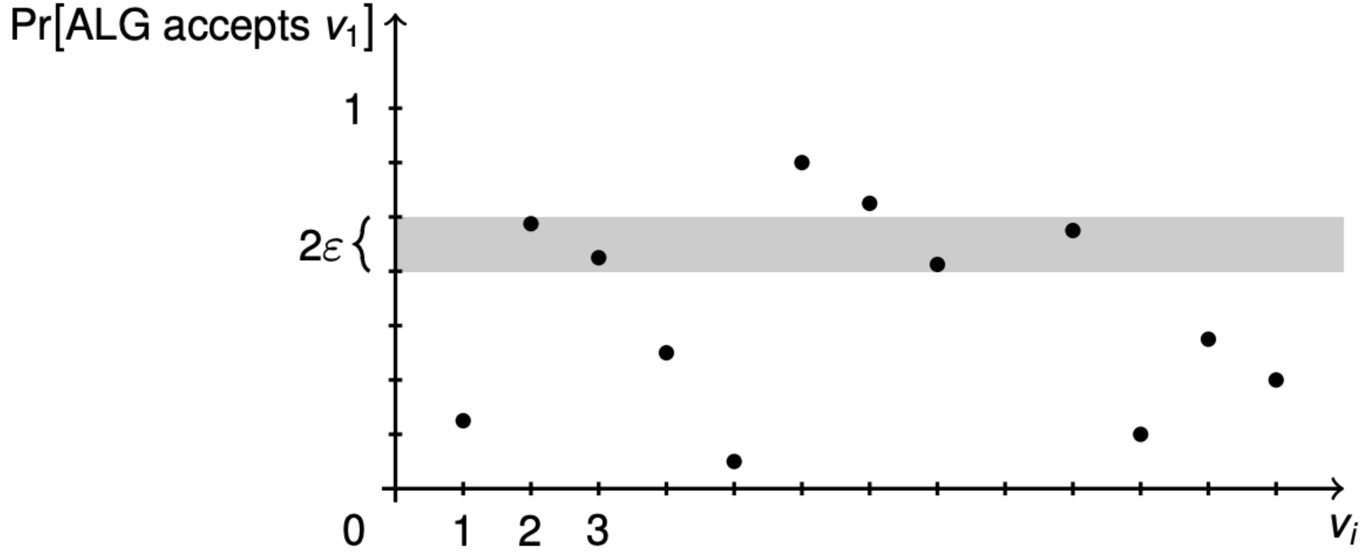


Establishing the Property

For all steps $i \in [n]$ there exists a probability $p_i \in [0,1]$ such that for all distinct values $v_1, \dots, v_i \in S$ seen until then,

$$\Pr[ALG \text{ accepts } v_i \mid v_i = \max\{v_1, \dots, v_i\}] \in [p_i - \epsilon,\ p_i + \epsilon]\,.$$

For $i = 1$:



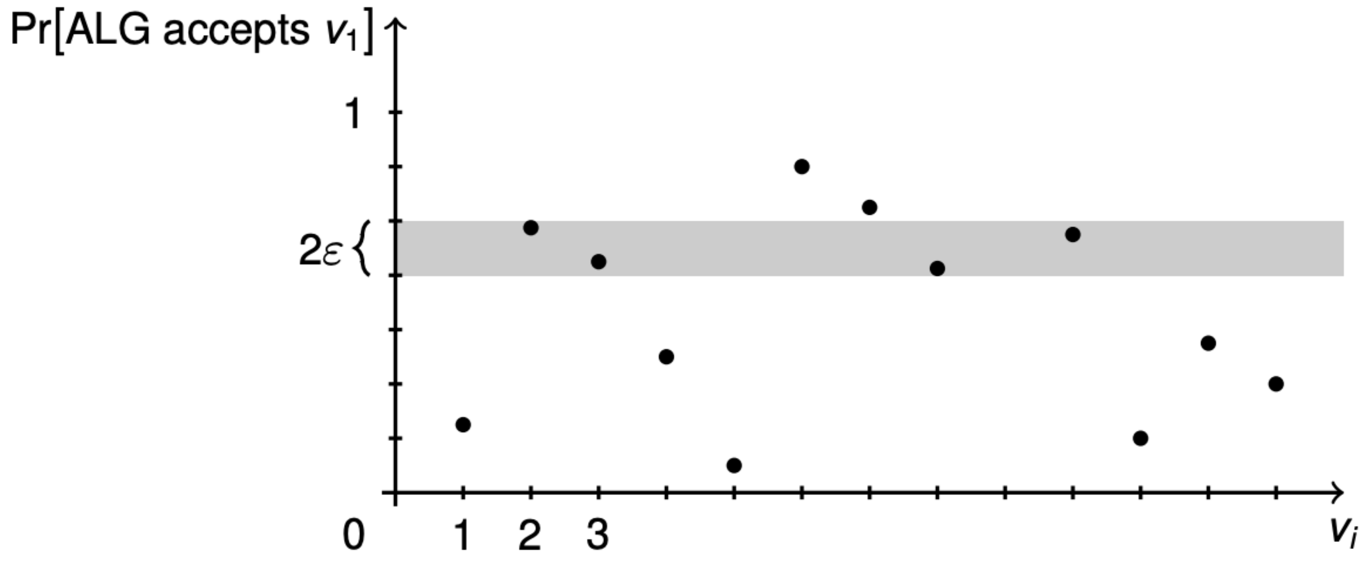
Bolzano-Weiertrass:
There exists a stripe of width 2ϵ that contains infinitely many points

Establishing the Property

For all steps $i \in [n]$ there exists a probability $p_i \in [0,1]$ such that for all distinct values $v_1, \dots, v_i \in S$ seen until then,

$$\Pr[ALG \text{ accepts } v_i \mid v_i = \max\{v_1, \dots, v_i\}] \in [p_i - \epsilon,\ p_i + \epsilon]\,.$$

For $i = 1$:



Bolzano-Weiertrass:
There exists a stripe of width 2ϵ that contains infinitely many points

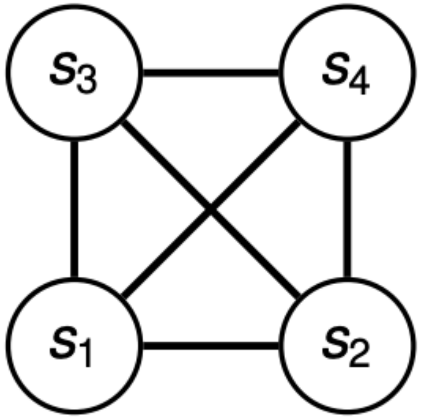
Make this (infinite-size) set S_1

Establishing the Property

For all steps $i \in [n]$ there exists a probability $p_i \in [0,1]$ such that for all distinct values $v_1, \dots, v_i \in S$ seen until then,

$$\Pr[ALG \text{ accepts } v_i \mid v_i = \max\{v_1, \dots, v_i\}] \in [p_i - \epsilon,\ p_i + \epsilon]\,.$$

For $i = 2$:



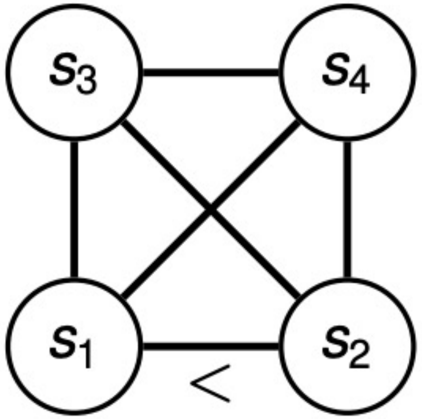
(infinite-size) complete graph on S_1

Establishing the Property

For all steps $i \in [n]$ there exists a probability $p_i \in [0,1]$ such that for all distinct values $v_1, \dots, v_i \in S$ seen until then,

$$\Pr[ALG \text{ accepts } v_i \mid v_i = \max\{v_1, \dots, v_i\}] \in [p_i - \epsilon,\ p_i + \epsilon] .$$

For $i = 2$:



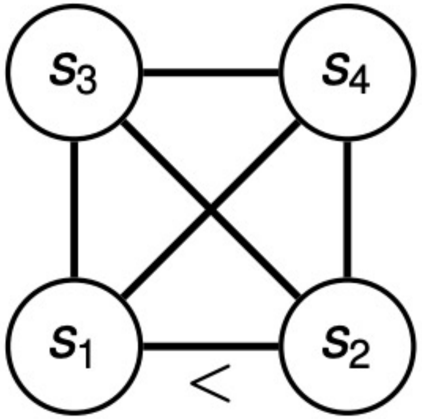
(infinite-size) complete graph on S_1

Establishing the Property

For all steps $i \in [n]$ there exists a probability $p_i \in [0,1]$ such that for all distinct values $v_1, \ldots, v_i \in S$ seen until then,

$$\Pr[ALG \text{ accepts } v_i \mid v_i = \max\{v_1, \ldots, v_i\}] \in [p_i - \epsilon, \ p_i + \epsilon] \, .$$

For $i = 2$:



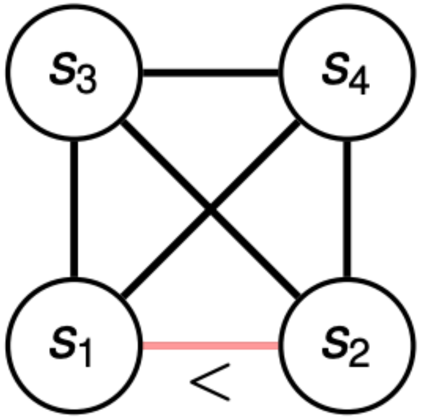
(infinite-size) complete graph on S_1

Establishing the Property

For all steps $i \in [n]$ there exists a probability $p_i \in [0,1]$ such that for all distinct values $v_1, \dots, v_i \in S$ seen until then,

$$\Pr[ALG \text{ accepts } v_i \mid v_i = \max\{v_1, \dots, v_i\}] \in [p_i - \epsilon,\ p_i + \epsilon]\,.$$

For $i = 2$:



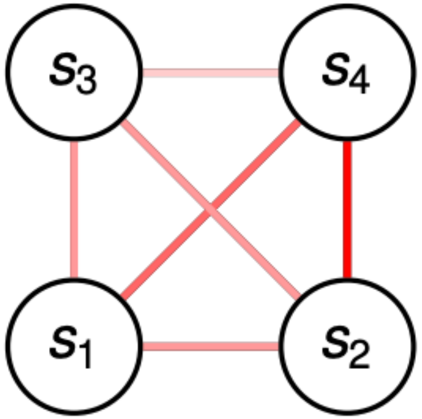
(infinite-size) complete graph on S_1

Establishing the Property

For all steps $i \in [n]$ there exists a probability $p_i \in [0,1]$ such that for all distinct values $v_1, \ldots, v_i \in S$ seen until then,

$$\Pr[ALG \text{ accepts } v_i \mid v_i = \max\{v_1, \ldots, v_i\}] \in [p_i - \epsilon,\ p_i + \epsilon] \,.$$

For $i = 2$:



(infinite-size) complete graph on S_1

Ramsey:
There exists an infinite-size monochromatic induced subgraph

Make this (infinite-size) set S_2

Establishing the Property

For all steps $i \in [n]$ there exists a probability $p_i \in [0,1]$ such that for all distinct values $v_1, \dots, v_i \in S$ seen until then,

$$\Pr[ALG \text{ accepts } v_i \mid v_i = \max\{v_1, \dots, v_i\}] \in [p_i - \epsilon,\ p_i + \epsilon] .$$

For $i \to i+1$:

Same argument as $(i = 1) \to (i = 2)$ except that we construct a hypergraph and use the hypergraph-version of Ramsey's theorem

Establishing the Property

For all steps $i \in [n]$ there exists a probability $p_i \in [0,1]$ such that for all distinct values $v_1, \dots, v_i \in S$ seen until then,

$$\Pr[ALG \text{ accepts } v_i \mid v_i = \max\{v_1, \dots, v_i\}] \in [p_i - \epsilon,\ p_i + \epsilon] .$$

<u>For $i \to i + 1$:</u>

Same argument as $(i = 1) \to (i = 2)$ except that we construct a hypergraph and use the hypergraph-version of Ramsey's theorem

Q.E.D.

Hard Instance

- Let $w_1, \dots, w_{n^3}, u \in S$ be such that $u \geq n^3 \cdot \max\{w_1, \dots, w_{n^3}\}$
- Define distribution $\mathcal{D}$ such that

$$v_i = \begin{cases} u & \text{with probability } \frac{1}{n^2} \\ w_j & \text{with probability } \frac{1}{n^3}\left(1 - \frac{1}{n^2}\right) \qquad (\forall\, j = 1, \dots, n^3) \end{cases}$$

(this completes the proof sketch for the main result)

Linear Number of Samples

Theorem [Correa Dütting Schewior Fischer 2019]

There exists an algorithm ALG that requires $n-1$ samples, and ensures that

$$\mathbb{E}[ALG] \geq \left(1 - \frac{1}{e}\right) \cdot \mathbb{E}[\max_i v_i]$$

Linear Number of Samples

Theorem [Correa Dütting Schewior Fischer 2019]

There exists an algorithm ALG that requires $n-1$ samples, and ensures that

$$\mathbb{E}[ALG] \geq \left(1 - \frac{1}{e}\right) \cdot \mathbb{E}[\max_i v_i]$$

(in the paper: nearly tight impossibility of $\ln(2) \approx 0.693$)

(in particular: unlike in the case of non-identical distributions, a single sample from each distribution is <u>not</u> sufficient to match the best-possible guarantee for a known distribution)

Warm-Up: Quadratic Number of Samples

Algorithm:

- Use $n-1$ fresh samples $s_1^{(i)}, \ldots, s_{n-1}^{(i)}$ for each step i
- Set $\max\{s_1^{(i)}, \ldots, s_{n-1}^{(i)}\}$ as threshold in step i

Warm-Up: Quadratic Number of Samples

Algorithm:

- Use $n-1$ fresh samples $s_1^{(i)}, \dots, s_{n-1}^{(i)}$ for each step i
- Set $\max\{s_1^{(i)}, \dots, s_{n-1}^{(i)}\}$ as threshold in step i

Analysis:

$$\mathbb{E}[ALG] = \sum_{i=1}^{n} \left(1 - \frac{1}{n}\right)^{i-1} \cdot \frac{1}{n} \cdot \mathbb{E}\left[v_i \mid v_i \geq \max\{s_1^{(i)}, \dots, s_{n-1}^{(i)}\}\right]$$

Warm-Up: Quadratic Number of Samples

Algorithm:

- Use $n - 1$ fresh samples $s_1^{(i)}, \ldots, s_{n-1}^{(i)}$ for each step i
- Set $\max\{s_1^{(i)}, \ldots, s_{n-1}^{(i)}\}$ as threshold in step i

Analysis:

$$\mathbb{E}[ALG] = \underbrace{\sum_{i=1}^{n} \left(1 - \frac{1}{n}\right)^{i-1} \cdot \frac{1}{n}}_{\to 1 - \frac{1}{e}} \cdot \underbrace{\mathbb{E}\left[v_i \mid v_i \geq \max\{s_1^{(i)}, \ldots, s_{n-1}^{(i)}\}\right]}_{= \mathbb{E}[\max\{v_1, \ldots, v_n\}]}$$

Warm-Up: Quadratic Number of Samples

Algorithm:

- Use $n-1$ fresh samples $s_1^{(i)}, \dots, s_{n-1}^{(i)}$ for each step i
- Set $\max\{s_1^{(i)}, \dots, s_{n-1}^{(i)}\}$ as threshold in step i

Analysis:

$$\mathbb{E}[ALG] = \underbrace{\sum_{i=1}^{n} \left(1 - \frac{1}{n}\right)^{i-1} \cdot \frac{1}{n}}_{\to 1 - \frac{1}{e}} \cdot \underbrace{\mathbb{E}\left[v_i \mid v_i \geq \max\{s_1^{(i)}, \dots, s_{n-1}^{(i)}\}\right]}_{= \mathbb{E}[\max\{v_1, \dots, v_n\}]}$$

Q.E.D.

(with $O(n^2)$ samples)

Fresh Looking Samples

Algorithm':

- At each time step i, select a uniform random subsample $\mathcal{S}_i$ of size $n-1$ of
  - $n-1$ original samples $s_1, \ldots, s_{n-1}$
  - all values $v_1, \ldots, v_{i-1}$ seen so far
- Set $\max \mathcal{S}_i$ as threshold in step i

Key Lemma

Lemma [Correa Dütting Schewior Fischer 2019]

Conditioned on arriving at step i, the distribution of $\mathcal{S}_i$ (as a set) is the one of $n - 1$ fresh samples,

Proof (of theorem): Analogous to previous argument. **Q.E.D.**

Additional Results

Theorem [Correa Dütting Schewior Fischer 2019]

For every $\epsilon > 0$, there exists an algorithm ALG that requires access to $O_\epsilon(n^2)$ samples, and ensures that

$$\mathbb{E}[ALG] \geq (0.745 - \epsilon) \cdot \mathbb{E}[\max_i v_i]$$

Additional Results

Theorem [Correa Dütting Schewior Fischer 2019]

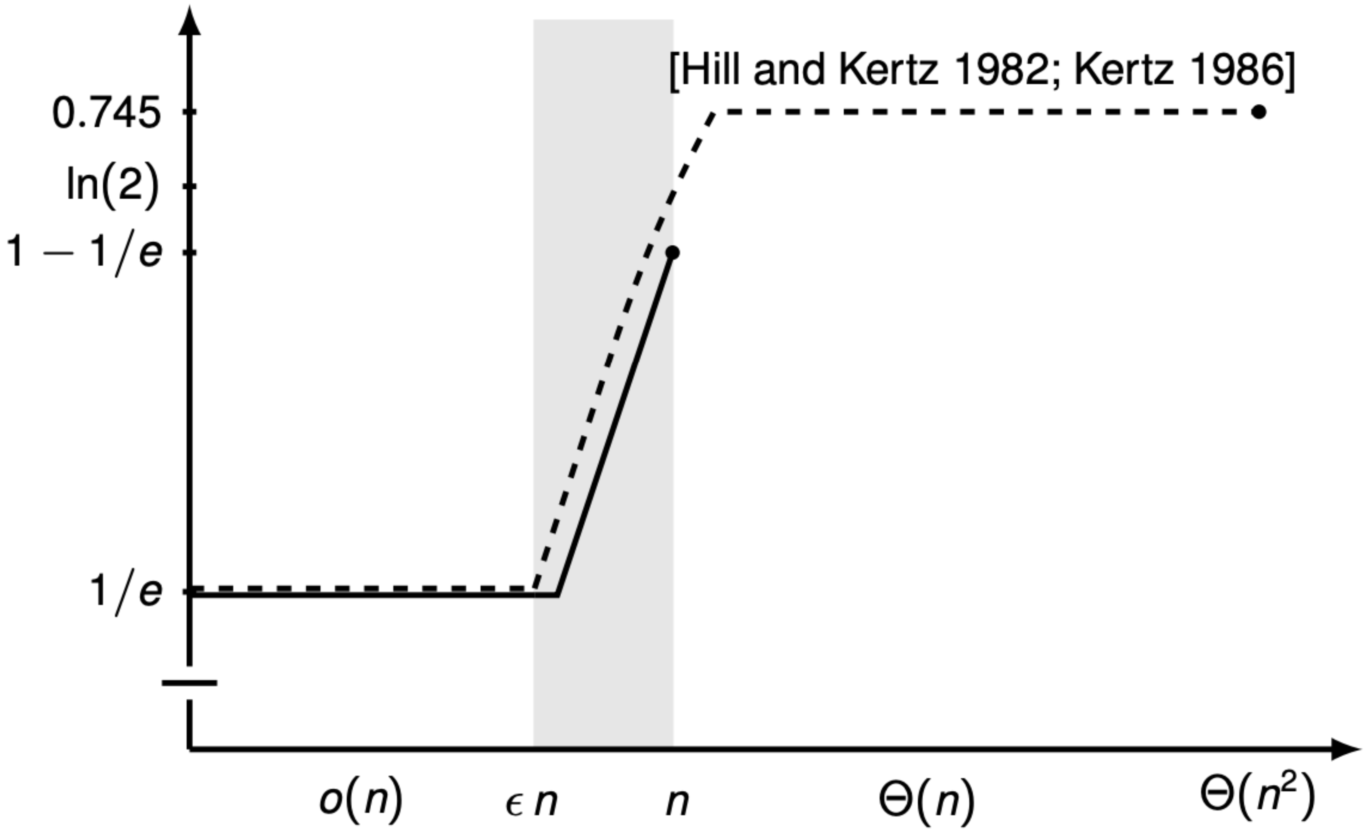
For every $\epsilon > 0$, there exists an algorithm ALG that requires access to $O_\epsilon(n^2)$ samples, and ensures that

$$\mathbb{E}[ALG] \geq (0.745 - \epsilon) \cdot \mathbb{E}[\max_i v_i]$$

(improved to $O_\epsilon(n)$ [Rubinstein Wang Weinberg '20])

(i.e., a constant number of samples per step $i \in [n]$ is sufficient to recover the optimal guarantee for a known distribution)

Summary



Follow-Up Work

- [Kaplan Naor Raz 2020]: Improved bounds for $k < n-1$ samples, same bound of $1-1/e$ for $k = n-1$ samples
- [Correa Cristi Epstein Soto 2020] Game of Googol (with random order) yields improved bound of 0.635 with $k = n$ samples
- [Correa Dütting Schewior Fischer Ziliotto 2021] Choose sets $\mathcal{S}_i$ of varying size; optimal choice yields improved bounds, tight for $k \leq \beta \cdot n$ samples and $\beta \leq 1/(e-1) \approx 0.58$, improved bound of 0.649 for $k = n$ samples
- [Correa Epstein Cristi Soto 2024] LP approach that finds optimal ordinal algorithm, yields 0.671 with $k = n$ samples

Beyond Single Item

Two Main Techniques

- Reduction to order-oblivious secretary problem
  [Azar Kleinberg and Weinberg 2014]
- Greedy plus deferred decisions
  [Korula Pal ‘09, Rubinstein Wang Weinberg ’20, Caramanis et al. 2022, Dütting Kesselheim Lucier Reiffenhauser Singla 2024]

Overview of Results

Setting	Guarantee
k-uniform matroid	$\leq \left(1 - O\left(\frac{1}{\sqrt{k}}\right)\right)^{-1}$ [Azar Kleinberg Weinberg ‘14]
Transversal matroid	≤ 16 [Azar Kleinberg Weinberg ‘14] ≤ 8 [Caramanis et el. ‘22]
Graphic matroid	≤ 8 [Azar Kleinberg Weinberg ‘14] ≤ 4 [Caramanis et al. ‘22]
Laminar matroid	$\leq 12 \cdot \sqrt{3}$ [Azar Kleinberg Weinberg ‘14] $\leq 6 \cdot \sqrt{3}$ [Caramanis et al. ‘22]
General matching (edge arrivals)	≤ 16 [Caramanis et el. ‘22] ≤ 11.66 [Kaplan Naor Raz ‘22]
General matching (vertex arrivals)	≤ 8 [Caramanis et el. ‘22] ≤ 5.83 [Kaplan Naor Raz ‘22]
Budget-Additive Combinatorial Auctions	≤ 24 [Caramanis et el. ‘22]
XOS Combinatorial Auctions	≤ 567 [Dütting Kesselheim Lucier Reiffenhauser Singla ‘24]

Discussion

- Almost all problems that admit $O(1)$-approximation in prophet model, also admit $O(1)$-approximation with a single sample (but some evidence that single-sample prophet may be as hard as order-oblivious secretary problem) [Caramanis et al. '22]
- Some, but not all of the aforementioned results also correspond to truthful (price-based) mechanisms

Open questions: Single-sample $O(1)$-approx. for subadditive CAs? Truthful $O(1)$-approx for XOS CAs with polylog samples?

Summary

- A closer look at the results and techniques for the data-driven single-choice prophet inequality problem
  - Non-identical distributions
  - Identical distributions
- A brief discussion of state-of-the art for data-driven combinatorial prophet inequality problems

Thanks! You made it :)

Acknowledgements

Thanks to Jose Correa, Andres Cristi, Michal Feldman, Thomas Kesselheim, and Sahil Singla for sharing their slides.

Thanks to Andres Cristi, Tomer Ezra, Thomas Kesselheim, Sahil Singla, and David Wajc for feedback and pointers.