# Prophet Inequalities with Oracle Calls

and why they're useful

Vasilis Livanos

University of Chile

April 17th, 2024

# Prophet Inequality

Observe realizations $X_1, X_2, \ldots, X_n \sim$ (known) $\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_n$.

- Order is *adversarial*.
- Design algorithm to maximize selected value.
- Compare against all-knowing *prophet*.
- If $\mathcal{D}_1 = \mathcal{D}_2 = \cdots = \mathcal{D}_n \implies$ IID setting.

# Prophet Inequality

Observe realizations $X_1, X_2, \ldots, X_n \sim$ (known) $\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_n$.

- Order is *adversarial*.
- Design algorithm to maximize selected value.
- Compare against all-knowing *prophet*.
- If $\mathcal{D}_1 = \mathcal{D}_2 = \cdots = \mathcal{D}_n \implies$ IID setting.

Objectives:

- Competitive Ratio (CR or "Prophet Objective"):

$$\frac{\mathbb{E}[ALG]}{\mathbb{E}[\max_i X_i]}$$

# Prophet Inequality

Observe realizations $X_1, X_2, \ldots, X_n \sim$ (known) $\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_n$.

- Order is *adversarial*.
- Design algorithm to maximize selected value.
- Compare against all-knowing *prophet*.
- If $\mathcal{D}_1 = \mathcal{D}_2 = \cdots = \mathcal{D}_n \implies$ IID setting.

Objectives:

- Competitive Ratio (CR or "Prophet Objective"):

$$\frac{\mathbb{E}[ALG]}{\mathbb{E}[\max_i X_i]}$$

- Probability of Selecting Maximum Realization (PMax or "Secretary Objective"):

$$\Pr\left[ALG \text{ selects } \max_i X_i\right]$$

# What is known?

Prophet Objective [Krengel, Sucheston, Garling '77, '78]

$\exists$ stopping strategy that achieves $1/2 \cdot \mathbb{E}[\max_i X_i]$, and this is tight.

Secretary Objective
[Esfandiari, HajiAghayi, Lucier, Mitzenmacher '20]

$\exists$ stopping strategy that selects $\max_i X_i$ with probability at least $1/e$, and this is tight.

# What is known?

Prophet Objective [Krengel, Sucheston, Garling '77, '78]

$\exists$ stopping strategy that achieves $1/2 \cdot \mathbb{E}[\max_i X_i]$, and this is tight.

Secretary Objective
[Esfandiari, HajiAghayi, Lucier, Mitzenmacher '20]

$\exists$ stopping strategy that selects $\max_i X_i$ with probability at least $1/e$, and this is tight.

- Idea: Set *threshold* $T$, accept first $X_i \geq T$.

# What is known?

Prophet Objective [Krengel, Sucheston, Garling '77, '78]

$\exists$ stopping strategy that achieves $1/2 \cdot \mathbb{E}[\max_i X_i]$, and this is tight.

Secretary Objective
[Esfandiari, HajiAghayi, Lucier, Mitzenmacher '20]

$\exists$ stopping strategy that selects $\max_i X_i$ with probability at least $1/e$, and this is tight.

- Idea: Set *threshold* $T$, accept first $X_i \geq T$.

  CR: $\Pr[\max_i X_i \geq T] = 1/2$ [Samuel-Cahn '84]

# What is known?

Prophet Objective [Krengel, Sucheston, Garling '77, '78]

$\exists$ stopping strategy that achieves $1/2 \cdot \mathbb{E}[\max_i X_i]$, and this is tight.

Secretary Objective
[Esfandiari, HajiAghayi, Lucier, Mitzenmacher '20]

$\exists$ stopping strategy that selects $\max_i X_i$ with probability at least $1/e$, and this is tight.

- Idea: Set *threshold* $T$, accept first $X_i \geq T$.

  CR: $\Pr[\max_i X_i \geq T] = 1/2$ [Samuel-Cahn '84]

  CR: $T = \frac{\mathbb{E}[\max_i X_i]}{2}$ [Kleinberg, Weinberg '12]

# What is known?

Prophet Objective [Krengel, Sucheston, Garling '77, '78]

$\exists$ stopping strategy that achieves $1/2 \cdot \mathbb{E}[\max_i X_i]$, and this is tight.

Secretary Objective
[Esfandiari, HajiAghayi, Lucier, Mitzenmacher '20]

$\exists$ stopping strategy that selects $\max_i X_i$ with probability at least $1/e$, and this is tight.

- Idea: Set *threshold* $T$, accept first $X_i \geq T$.

  CR: $\Pr[\max_i X_i \geq T] = 1/2$ [Samuel-Cahn '84]

  CR: $T = \frac{\mathbb{E}[\max_i X_i]}{2}$ [Kleinberg, Weinberg '12]

  PMax: $\Pr[\max_i X_i \geq T] = 1 - 1/e$
  [Esfandiari, HajiAghayi, Lucier, Mitzenmacher '20]

# What is known?

# What is known?

Competitive Ratio:

- Non-IID: $1/2$
  [Krengel, Sucheston, Garling '77, '78]

# What is known?

Competitive Ratio:

- Non-IID: $1/2$
  [Krengel, Sucheston, Garling '77, '78]
- IID: $\approx 0.745$
  [Hill-Kertz '82,
  Correa, Foncea, Hoeksma, Oosterwijk and Vredeveld '21]]

# What is known?

Competitive Ratio:

- Non-IID: $1/2$
  [Krengel, Sucheston, Garling '77, '78]
- IID: $\approx 0.745$
  [Hill-Kertz '82,
  Correa, Foncea, Hoeksma, Oosterwijk and Vredeveld '21]]

Secretary Objective:

- Non-IID: $1/e$
  [Esfandiari, HajiAghayi, Lucier, Mitzenmacher '20]

# What is known?

**Competitive Ratio:**

- Non-IID: $1/2$
  [Krengel, Sucheston, Garling '77, '78]
- IID: $\approx 0.745$
  [Hill-Kertz '82,
  Correa, Foncea, Hoeksma, Oosterwijk and Vredeveld '21]]

**Secretary Objective:**

- Non-IID: $1/e$
  [Esfandiari, HajiAghayi, Lucier, Mitzenmacher '20]
- IID: $\approx 0.58$
  [Gilbert, Mosteller '66]

# Top-1-of-k Model

- $ALG$ selects $S$ with $|S| \leq k$, but only $\max_{X_i \in S} X_i$ matters

# Top-1-of-k Model

- $ALG$ selects $S$ with $|S| \leq k$, but only $\max_{X_i \in S} X_i$ matters

1966: PMax $\geq 1 - e^{-\mathrm{O}(k)}$ for IID.
[Gilbert, Mosteller '66]

# Top-1-of-k Model

- $ALG$ selects $S$ with $|S| \leq k$, but only $\max_{X_i \in S} X_i$ matters

1966: PMax $\geq 1 - e^{-\mathrm{O}(k)}$ for IID.
[Gilbert, Mosteller '66]

2000: CR $\geq 1 - \frac{1}{k+1}$ for Non-IID.
[Assaf, Samuel-Cahn '00]

# Top-1-of-k Model

- $ALG$ selects $S$ with $|S| \le k$, but only $\max_{X_i \in S} X_i$ matters

1966: PMax $\ge 1 - e^{-\mathrm{O}(k)}$ for IID.
[Gilbert, Mosteller '66]

2000: CR $\ge 1 - \frac{1}{k+1}$ for Non-IID.
[Assaf, Samuel-Cahn '00]

2002: CR: Numerical improvements for Non-IID, based on recursive function, but no asymptotic form.
[Assaf, Goldstein, Samuel-Cahn '02]

# Top-1-of-k Model

- $ALG$ selects $S$ with $|S| \leq k$, but only $\max_{X_i \in S} X_i$ matters

1966: PMax $\geq 1 - e^{-\mathrm{O}(k)}$ for IID.
[Gilbert, Mosteller '66]

2000: CR $\geq 1 - \frac{1}{k+1}$ for Non-IID.
[Assaf, Samuel-Cahn '00]

2002: CR: Numerical improvements for Non-IID, based on recursive function, but no asymptotic form.
[Assaf, Goldstein, Samuel-Cahn '02]

2018: $1 - e^{-k/6} \leq$ CR $\leq 1 - k^{-2k}$ for Non-IID.
[Ezra, Feldman, Nehama '18]

# Top-1-of-k Model

- $ALG$ selects $S$ with $|S| \leq k$, but only $\max_{X_i \in S} X_i$ matters

1966: PMax $\geq 1 - e^{-\mathrm{O}(k)}$ for IID.
[Gilbert, Mosteller '66]

2000: CR $\geq 1 - \frac{1}{k+1}$ for Non-IID.
[Assaf, Samuel-Cahn '00]

2002: CR: Numerical improvements for Non-IID, based on recursive function, but no asymptotic form.
[Assaf, Goldstein, Samuel-Cahn '02]

2018: $1 - e^{-k/6} \leq$ CR $\leq 1 - k^{-2k}$ for Non-IID.
[Ezra, Feldman, Nehama '18]

Motivation: Auctions/Hiring with *overbooking*.

# Oracle-Augmented Prophet Inequalities

# Oracle-Augmented Prophet Inequalities

$\underline{\mathcal{O}_k:}$ Assume $ALG$ has $k$ calls to $\mathcal{O}$, who knows $X_1, \dots, X_n$.

# Oracle-Augmented Prophet Inequalities

$\mathcal{O}_k$ : Assume $ALG$ has $k$ calls to $\mathcal{O}$, who knows $X_1, \dots, X_n$.

- Step $i$:
  - 👍 $X_i \geq \max_{j=i+1}^{n} X_j \implies ALG$ selects $X_i$
  - 👎 $X_i < \max_{j=i+1}^{n} X_j \implies ALG$ rejects $X_i$
- Generalization of standard PI.
- Allows for simpler analysis since $ALG$ always selects one value.
- "Algorithms with predictions".

# Oracle-Augmented Prophet Inequalities

$\mathcal{O}_k$ : Assume $ALG$ has $k$ calls to $\mathcal{O}$, who knows $X_1, \dots, X_n$.

- Step $i$:
  - 👍 $X_i \geq \max_{j=i+1}^n X_j \implies ALG$ selects $X_i$
  - 👎 $X_i < \max_{j=i+1}^n X_j \implies ALG$ rejects $X_i$
- Generalization of standard PI.
- Allows for simpler analysis since $ALG$ always selects one value.
- "Algorithms with predictions".

Q.: Is $\mathcal{O}_k$ equivalent to Top-1-of-$(k+1)$?

# Oracle-Augmented Prophet Inequalities

$\mathcal{O}_k$ : Assume $ALG$ has $k$ calls to $\mathcal{O}$, who knows $X_1, \dots, X_n$.

- Step $i$:
  - 👍 $X_i \geq \max_{j=i+1}^{n} X_j \Longrightarrow ALG$ selects $X_i$
  - 👎 $X_i < \max_{j=i+1}^{n} X_j \Longrightarrow ALG$ rejects $X_i$
- Generalization of standard PI.
- Allows for simpler analysis since $ALG$ always selects one value.
- "Algorithms with predictions".

Q.: Is $\mathcal{O}_k$ equivalent to Top-1-of-$(k+1)$?

A.: YES and NO.

# Let's Play!

$$k = 1$$

| | | | |
|---|---|---|---|
| $\mathcal{U}[13, 14]$ | $\mathcal{U}[7, 16]$ | $\mathcal{U}[0, 20]$ | $\begin{cases} 1000 & \text{w.p. } \frac{1}{100} \\ 0 & \text{otherwise} \end{cases}$ |

# Let's Play!

$$k = 1$$

| | | | |
|---|---|---|---|
| $\mathcal{U}[13, 14]$ | $\mathcal{U}[7, 16]$ | $\mathcal{U}[0, 20]$ | $\begin{cases} 1000 & \text{w.p. } \frac{1}{100} \\ 0 & \text{otherwise} \end{cases}$ |

$X_1 = 13.67$

# Let's Play!

$$k = 1$$

| | | | |
|---|---|---|---|
| $\mathcal{U}[13, 14]$ | $\mathcal{U}[7, 16]$ | $\mathcal{U}[0, 20]$ | $\begin{cases} 1000 & \text{w.p. } \frac{1}{100} \\ 0 & \text{otherwise} \end{cases}$ |
| $X_1 = 13.67$ | $X_2 = 8.59$ | | |

# Let's Play!

$$k = 1$$

| | | | |
|---|---|---|---|
| $\mathcal{U}[13, 14]$ | $\mathcal{U}[7, 16]$ | $\mathcal{U}[0, 20]$ | $\begin{cases} 1000 & \text{w.p. } \frac{1}{100} \\ 0 & \text{otherwise} \end{cases}$ |
| $X_1 = 13.67$ | $X_2 = 8.59$ | $X_3 = 2.82$ | |

# Let's Play!

$$k = 1$$

| $\mathcal{U}[13, 14]$ | $\mathcal{U}[7, 16]$ | $\mathcal{U}[0, 20]$ | $\begin{cases} 1000 & \text{w.p. } \frac{1}{100} \\ 0 & \text{otherwise} \end{cases}$ |
|---|---|---|---|
| $X_1 = 13.67$ | $X_2 = 8.59$ | $X_3 = 2.82$ | $X_4 = 0$ |

$$\mathbb{E}[\max\{X_1, X_2, X_3, X_4\}] \approx 24.66$$

$$\mathbb{E}[OPTALG\{X_1, X_2, X_3, X_4\} \text{ for } k = 0] \approx 13.37$$

$$\mathbb{E}[OPTALG\{X_1, X_2, X_3, X_4\} \text{ for } k = 1] \approx 19.9$$

Optimal strategy was to ignore $X_1, X_2$ and query at $X_3$.

# $\mathcal{O}_k \equiv$ Top-1-of-$(k+1)$ for PMax

# $\mathcal{O}_k \equiv$ Top-1-of-$(k+1)$ for PMax

Assume algorithm $\mathcal{A}$ for Top-1-of-$(k+1)$.

- $X_i \to \mathcal{A}$. If $\mathcal{A}$ selects $X_i$, we use $\mathcal{O}$.
- If we're out of oracle calls, select $X_i$.

# $\mathcal{O}_k \equiv$ Top-1-of-$(k+1)$ for PMax

Assume algorithm $\mathcal{A}$ for Top-1-of-$(k+1)$.

- $X_i \to \mathcal{A}$. If $\mathcal{A}$ selects $X_i$, we use $\mathcal{O}$.
- If we're out of oracle calls, select $X_i$.

$$\Pr[\text{We select } \max_i X_i] = \Pr[\mathcal{A} \text{ selects } \max_i X_i].$$

# $\mathcal{O}_k \equiv$ Top-1-of-$(k+1)$ for PMax

Assume algorithm $\mathcal{A}$ for Top-1-of-$(k+1)$.

- $X_i \to \mathcal{A}$. If $\mathcal{A}$ selects $X_i$, we use $\mathcal{O}$.
- If we're out of oracle calls, select $X_i$.

$$\Pr[\text{We select } \max_i X_i] = \Pr[\mathcal{A} \text{ selects } \max_i X_i].$$

Assume algorithm $\mathcal{B}$ for $\mathcal{O}_k$.

- $X_i \to \mathcal{B}$. If $\mathcal{B}$ expends oracle call or selects $X_i$, we select $X_i$.

# $\mathcal{O}_k \equiv$ Top-1-of-$(k+1)$ for PMax

Assume algorithm $\mathcal{A}$ for Top-1-of-$(k+1)$.

- $X_i \to \mathcal{A}$. If $\mathcal{A}$ selects $X_i$, we use $\mathcal{O}$.
- If we're out of oracle calls, select $X_i$.

$$\Pr[\text{We select } \max_i X_i] = \Pr[\mathcal{A} \text{ selects } \max_i X_i].$$

Assume algorithm $\mathcal{B}$ for $\mathcal{O}_k$.

- $X_i \to \mathcal{B}$. If $\mathcal{B}$ expends oracle call or selects $X_i$, we select $X_i$.

$$\Pr[\text{We select } \max_i X_i] = \Pr[\mathcal{B} \text{ selects } \max_i X_i].$$

# $\mathcal{O}_k \not\equiv$ Top-1-of-$(k+1)$ for CR

Guarantee $\alpha$ for $\mathcal{O}_k \implies \alpha$ for Top-1-of-$(k+1)$.

# $\mathcal{O}_k \not\equiv$ Top-1-of-$(k+1)$ for CR

Guarantee $\alpha$ for $\mathcal{O}_k \implies \alpha$ for Top-1-of-$(k+1)$.
Fix $\varepsilon > 0, k = 1$, and let

$$X_1 = 1, \quad X_2 = \begin{cases} 1+\varepsilon & \text{w.p. } 1/2 - \varepsilon \\ 0 & \text{w.p. } 1/2 + \varepsilon \end{cases}, \quad X_3 = \begin{cases} 1/\varepsilon & \text{w.p. } \varepsilon \\ 0 & \text{w.p. } 1 - \varepsilon \end{cases}$$

# $\mathcal{O}_k \not\equiv$ Top-1-of-$(k+1)$ for CR

Guarantee $\alpha$ for $\mathcal{O}_k \implies \alpha$ for Top-1-of-$(k+1)$.
Fix $\varepsilon > 0, k = 1$, and let

$$X_1 = 1, \quad X_2 = \begin{cases} 1+\varepsilon & \text{w.p. } 1/2 - \varepsilon \\ 0 & \text{w.p. } 1/2 + \varepsilon \end{cases}, \quad X_3 = \begin{cases} 1/\varepsilon & \text{w.p. } \varepsilon \\ 0 & \text{w.p. } 1-\varepsilon \end{cases}$$

- $\mathbb{E}[\max\{X_1, X_2, X_3\}] \to 2$.
- Let $\mathcal{A}$ for Top-1-of-2 that always selects $X_1$ and $X_3$.
  $\mathbb{E}[\mathcal{A}] \to 2$.

# $\mathcal{O}_k \not\equiv$ Top-1-of-$(k+1)$ for CR

Guarantee $\alpha$ for $\mathcal{O}_k \implies \alpha$ for Top-1-of-$(k+1)$.
Fix $\varepsilon > 0, k = 1$, and let

$$X_1 = 1, \quad X_2 = \begin{cases} 1+\varepsilon & \text{w.p. } 1/2 - \varepsilon \\ 0 & \text{w.p. } 1/2 + \varepsilon \end{cases}, \quad X_3 = \begin{cases} 1/\varepsilon & \text{w.p. } \varepsilon \\ 0 & \text{w.p. } 1 - \varepsilon \end{cases}$$

- $\mathbb{E}[\max\{X_1, X_2, X_3\}] \to 2$.
- Let $\mathcal{A}$ for Top-1-of-2 that always selects $X_1$ and $X_3$.
  $\mathbb{E}[\mathcal{A}] \to 2$.
- Let $\mathcal{B}$ for $\mathcal{O}_1$.
  - If $\mathcal{B}$ doesn't query at $X_1 \implies \mathbb{E}[\mathcal{B}] = \mathbb{E}[\max\{X_2, X_3\}] \to 3/2$.

# $\mathcal{O}_k \not\equiv$ Top-1-of-$(k+1)$ for CR

Guarantee $\alpha$ for $\mathcal{O}_k \implies \alpha$ for Top-1-of-$(k+1)$.
Fix $\varepsilon > 0, k = 1$, and let

$$X_1 = 1, \quad X_2 = \begin{cases} 1+\varepsilon & \text{w.p. } 1/2 - \varepsilon \\ 0 & \text{w.p. } 1/2 + \varepsilon \end{cases}, \quad X_3 = \begin{cases} 1/\varepsilon & \text{w.p. } \varepsilon \\ 0 & \text{w.p. } 1 - \varepsilon \end{cases}$$

- $\mathbb{E}[\max\{X_1, X_2, X_3\}] \to 2$.
- Let $\mathcal{A}$ for Top-1-of-2 that always selects $X_1$ and $X_3$.
  $\mathbb{E}[\mathcal{A}] \to 2$.
- Let $\mathcal{B}$ for $\mathcal{O}_1$.
  - If $\mathcal{B}$ doesn't query at $X_1 \implies \mathbb{E}[\mathcal{B}] = \mathbb{E}[\max\{X_2, X_3\}] \to 3/2$.
  - If $\mathcal{B}$ queries at $X_1 \implies$

$$\mathbb{E}[\mathcal{B}] = \left(\frac{1}{2}+\varepsilon\right)(1-\varepsilon)\cdot 1 + \left(\frac{1}{2}-\varepsilon\right)\cdot(1+\varepsilon) + \left(\frac{1}{2}+\varepsilon\right)\varepsilon\cdot\frac{1}{\varepsilon} \to 3/2.$$

# PMax for IID

# PMax for IID

Theorem [Har-Peled, Harb, L. '24]

Single-threshold* algorithm $\mathcal{A}$ achieves

$$\text{PMax}(\mathcal{A}) \geq 1 - \mathrm{O}\left(k^{-k/5}\right).$$

# PMax for IID

**Theorem** [Har-Peled, Harb, L. '24]

Single-threshold* algorithm $\mathcal{A}$ achieves

$$\text{PMax}(\mathcal{A}) \geq 1 - \mathrm{O}\left(k^{-k/5}\right).$$

For every algorithm $ALG$,

$$\text{PMax}(ALG) \leq 1 - \mathrm{O}\left(k^{-k}\right).$$

# PMax for IID

Theorem [Har-Peled, Harb, L. '24]

Single-threshold* algorithm $\mathcal{A}$ achieves

$$\text{PMax}(\mathcal{A}) \geq 1 - \mathrm{O}\left(k^{-k/5}\right).$$

For every algorithm $ALG$,

$$\text{PMax}(ALG) \leq 1 - \mathrm{O}\left(k^{-k}\right).$$

- Almost tight bound, up to the exponent.
- New upper bound.
  Asymptotical improvement on lower bound of [Gilbert, Mosteller '66]

# PMax for IID

# PMax for IID

- Idea: Set threshold $T$ such that $\Pr[X_i \geq T] = L$ for some $L$.

  Use Chernoff bound to argue that $\leq k$ of the $X_i$'s are above $T$ and we don't run out of oracle calls.

# PMax for IID

- Idea: Set threshold $T$ such that $\Pr[X_i \geq T] = L$ for some $L$.

  Use Chernoff bound to argue that $\leq k$ of the $X_i$'s are above $T$ and we don't run out of oracle calls.

- Observation: Only use oracle if $X_{i_1} \leq X_{i_2} \leq \cdots \leq X_{i_m}$!

# PMax for IID

- ~~<u>Idea:</u> Set threshold $T$ such that $\Pr[X_i \geq T] = L$ for some $L$. Use Chernoff bound to argue that $\leq k$ of the $X_i$'s are above $T$ and we don't run out of oracle calls.~~

- <u>Observation:</u> Only use oracle if $X_{i_1} \leq X_{i_2} \leq \cdots \leq X_{i_m}$!

- <u>Better idea:</u> Set threshold $T$ such that $\Pr[X_i \geq T] = \frac{e^{\sqrt{k}}}{n}$.

  Use Chernoff bound to argue that $\mathrm{MAX}$ of sequence of $X_i$'s above $T$ changes $\leq k$ times $\implies$ We don't run out of queries.

# PMax for IID

- ~~Idea: Set threshold $T$ such that $\Pr[X_i \geq T] = L$ for some $L$. Use Chernoff bound to argue that $\leq k$ of the $X_i$'s are above $T$ and we don't run out of oracle calls.~~

- Observation: Only use oracle if $X_{i_1} \leq X_{i_2} \leq \cdots \leq X_{i_m}$!

- Better idea: Set threshold $T$ such that $\Pr[X_i \geq T] = \frac{e^{\sqrt{k}}}{n}$.

  Use Chernoff bound to argue that $\mathrm{MAX}$ of sequence of $X_i$'s above $T$ changes $\leq k$ times $\implies$ We don't run out of queries.

- Upper Bound: $\mathcal{U}[0,1]$ with roughly $k \log k$ of the $X_i$'s in $[1 - {}^{k \log k}/_{n}, 1]$.

  Every algorithm will either miss $\max_i X_i$ or run out of oracle calls w.p. $k^{-k}$.

# CR for Non-IID

# CR for Non-IID

Theorem [Har-Peled, Harb, L. '24]

$\exists$ sequence $\{\xi_k\}_{k\in\mathbb{N}}$ and single-threshold* algorithm $\mathcal{A}$ such that

$$\mathsf{CR}(\mathcal{A}) \geq 1 - \mathrm{O}\left(e^{-\xi_k}\right) = 1 - \mathrm{O}\left(e^{-k/e+\mathrm{o}(k)}\right).$$

# CR for Non-IID

Theorem [Har-Peled, Harb, L. '24]

$\exists$ sequence $\{\xi_k\}_{k\in\mathbb{N}}$ and single-threshold* algorithm $\mathcal{A}$ such that

$$\mathsf{CR}(\mathcal{A}) \geq 1 - \mathrm{O}\left(e^{-\xi_k}\right) = 1 - \mathrm{O}\left(e^{-k/e + \mathrm{o}(k)}\right).$$

For every algorithm $ALG$,

$$\mathsf{CR}(ALG) \leq 1 - \mathrm{O}\left(e^{-\xi_k}\right).$$

# CR for Non-IID

Theorem [Har-Peled, Harb, L. '24]

$\exists$ sequence $\{\xi_k\}_{k\in\mathbb{N}}$ and single-threshold* algorithm $\mathcal{A}$ such that

$$\text{CR}(\mathcal{A}) \geq 1 - \mathrm{O}\left(e^{-\xi_k}\right) = 1 - \mathrm{O}\left(e^{-k/e+\mathrm{o}(k)}\right).$$

For every algorithm $ALG$,

$$\text{CR}(ALG) \leq 1 - \mathrm{O}\left(e^{-\xi_k}\right).$$

- Tight bound for the oracle model.
- Asymptotical improvement on both upper and lower bounds of [Ezra, Feldman, Nehama '18].

# CR for Non-IID

# CR for Non-IID

- $1 - e^{-\xi_k} = e^{-\xi_k} \sum_{i=0}^{k} \frac{(\xi_k)^i}{i!} \iff e^{-\xi_k} = e^{-\xi_k} \sum_{i=k+1}^{\infty} \frac{(\xi_k)^i}{i!}$

  $\xi_k$ : Exponent sequence.

  Intuition: Balances $\Pr[\text{Poi}(\xi_k) \geq k+1] = \Pr[\text{Poi}(\xi_k) = 0]$.

# CR for Non-IID

- $1 - e^{-\xi_k} = e^{-\xi_k} \sum_{i=0}^{k} \frac{(\xi_k)^i}{i!} \iff e^{-\xi_k} = e^{-\xi_k} \sum_{i=k+1}^{\infty} \frac{(\xi_k)^i}{i!}$

  $\xi_k$ : Exponent sequence.

  Intuition: Balances $\Pr[\text{Poi}(\xi_k) \geq k+1] = \Pr[\text{Poi}(\xi_k) = 0]$.

- Idea: Set threshold $T$ such that $\Pr[\max_i X_i \geq T] = 1 - e^{-\xi_k}$.

  For $x \geq T$, analyze $\Pr[ALG \geq x]$ using Poissonization technique of [Harb '23].

# CR for Non-IID

- $1 - e^{-\xi_k} = e^{-\xi_k} \sum_{i=0}^{k} \frac{(\xi_k)^i}{i!} \iff e^{-\xi_k} = e^{-\xi_k} \sum_{i=k+1}^{\infty} \frac{(\xi_k)^i}{i!}$

  $\xi_k$ : Exponent sequence.

  Intuition: Balances $\Pr[\text{Poi}(\xi_k) \geq k+1] = \Pr[\text{Poi}(\xi_k) = 0]$.

- Idea: Set threshold $T$ such that $\Pr[\max_i X_i \geq T] = 1 - e^{-\xi_k}$.

  For $x \geq T$, analyze $\Pr[ALG \geq x]$ using Poissonization technique of [Harb '23].

  *Like in standard PI, we set* $\Pr[\max_i X_i \geq T] = p$ *and get a CR of* $p$*.*

# CR for Non-IID

- Upper Bound:

$$X_1 = 1, \quad X_2, \dots, X_{n-1} \sim \mathsf{Poi}(\xi_k), \quad X_n = \begin{cases} 1/\varepsilon & \text{w.p. } \varepsilon \\ 0 & \text{w.p. } 1-\varepsilon \end{cases}$$

# CR for Non-IID

- <u>Upper Bound:</u>

$$X_1 = 1, \quad X_2, \dots, X_{n-1} \sim \mathsf{Poi}(\xi_k), \quad X_n = \begin{cases} 1/\varepsilon & \text{w.p. } \varepsilon \\ 0 & \text{w.p. } 1-\varepsilon \end{cases}$$

- $$\mathbb{E}[\max_i X_i] = 2.$$

$$\begin{aligned} \mathbb{E}[ALG] &= 2 \cdot \Pr[\leq k \text{ of } X_2, \dots, X_{n-1} \neq 0] \\ &\quad + 1 \cdot (1 - \Pr[\leq k \text{ of } X_2, \dots, X_{n-1} \neq 0]) \\ &= 2\left(1 - e^{-\xi_k}\right), \end{aligned}$$

by the choice of $\xi_k$.

# Open Questions

Top-$1$-of-$k$ Model:

- Optimal asymptotics for PMax in IID setting and CR in Non-IID setting (requires new approach).
- PMax in Non-IID?

  Analyzing oracle model is enough!
- CR in IID?

Oracle Model:

- Incorrect predictions $\implies$ Robustness, ties with ML.

# Questions?