Prophet Inequalities

Jose Correa
Universidad de Chile

Material Perapared with
Andres Cristi (UChile → EPFL)

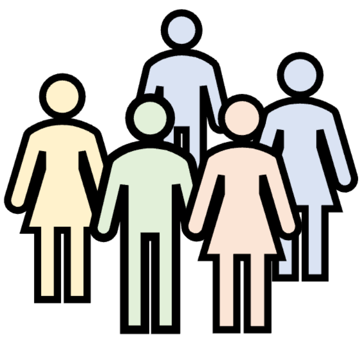
Motivation

Online platforms, e-commerce, etc

Flexible Model:

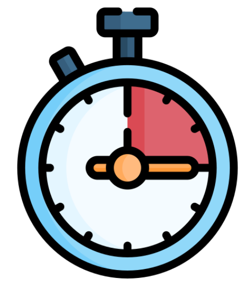
Multiple Goals

Incentives



Limited data

Sequential decisions



Prophet Inequalities – Optimal Stopping

- Johannes Kepler 1613 (who found a wife after interviewing 11 candidates!)

“ Though all Christians start a wedding invitation by solemnly declaring their marriage is due to special Divine arrangement, I as a philosopher, would like to discourse with you, O wisest of men, in greater detail about this. Was it Divine Providence or my own moral guilt which, for two years or longer, tore me in so many different directions and made m consider the possibility of such different unions?”

- Arthur Cayley 1875

4528. (Proposed by Professor Cayley) A lottery is arranged as follows: There are n tickets representing $a, b, c, \cdots$ pounds respectively. A person draws once; looks at his ticket; and if he pleases, draws again (out of the remaining $n-1$ tickets); and so on, drawing in all not more than k times; and he receives the value of the last ticket drawn. Supposing that he regulates his drawings in the manner most advantageous to him according to the theory of probabilities, what is the value of his expectation?

Even in the news!

Course Overview

1. Classic single-choice problems:

   The classic prophet inequality, secretary problem, prophet secretary problem, etc

2. Data driven prophet inequalities:

   How can limited amount of data be nearly as useful as full distributional knowledge

3. Combinatorial Prophet Inequalities

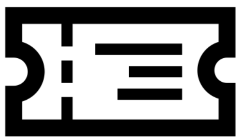
   Many ideas for single choice problems, extend to combinatorial contexts such as k-choice, Matching, hyper graph matching, and beyond

4. Online Combinatorial Auctions

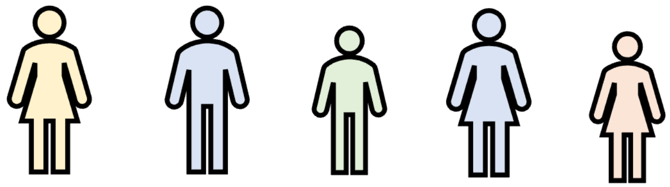
   General Model that encompasses many online selection/allocation problems

1. Classic single-choice problems

The prophet inequality



1 item

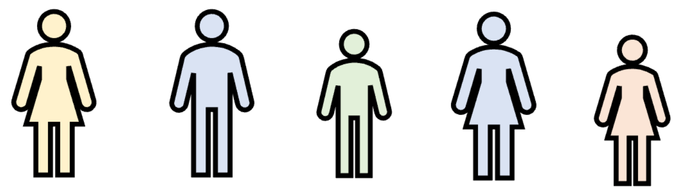


n agents
independent valuations
$v_i \sim F_i$

- We are given $F_1, \dots, F_n$
- Agents arrive sequentially: we observe $v_1 \sim F_1, v_2 \sim F_2, \dots$ one by one
- We (the Decision Maker) decide stop/continue
- We maximize $\mathbb{E}(v_{\text{stop}})$
- Compare against a prophet that can see realizations in advance and thus gets the optimal social welfare $\mathbb{E}\left(\max_i v_i\right)$

The prophet inequality

1 item



n agents
independent valuations
$v_i \sim F_i$

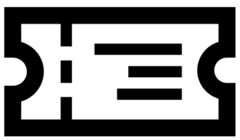
- Optimal online algorithm: Backwards Induction

$$V(v_1, \dots, v_n) = \mathbb{E}[\max\{v_1, V(v_2, \dots, v_n)\}]$$

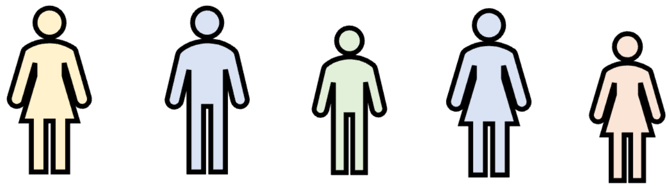
$$= V(v_2, \dots, v_n) + \mathbb{E}\left[\left(v_1 - V(v_2, \dots, v_n)\right)^+\right]$$

- Stop if current value exceeds the expectation of the future.

The prophet inequality



1 item



n agents
independent valuations
$v_i \sim F_i$

- Example: 3 agents, F_i=U[0,1]

DM:

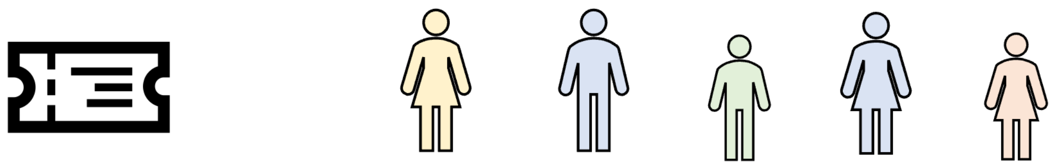
- $V(v_4)$=1/2
- $V(v_3, v_4)$=1/2 + $\mathbb{E}[\left(v_3 - \frac{1}{2}\right)^+]$ = 5/8
- $V(v_2, v_3, v_4)$=5/8 + $\mathbb{E}\left[\left(v_2 - \frac{5}{8}\right)^+\right]$

$$= \frac{5}{8} + \frac{3}{8} * \frac{3}{8} * \frac{1}{2} = \frac{89}{128}$$

Prophet:

- $\mathbb{E}(\max(v_4))$=1/2
- $\mathbb{E}\left(\max(v_3, v_4)\right) = \frac{2}{3}$
- $\mathbb{E}\left(\max(v_2, v_3, v_4)\right) = \frac{3}{4}$

The prophet inequality



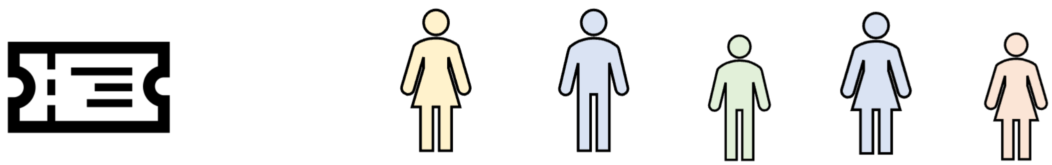
Theorem. [Krengel and Sucheston ‘77] $V(v_1, \ldots, v_n) \geq \frac{1}{2} \max_i v_i$

We can get ½ of the expected optimal welfare.

And ½ is the best possible constant:

$$v_i = 1 \qquad v_2 = \begin{cases} \frac{1}{\varepsilon} & w.p.\ \ \varepsilon, \text{ and} \\ 0 & w.p.\ 1 - \varepsilon \end{cases}$$

The prophet inequality

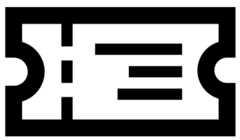


Theorem. [Krengel and Sucheston '77] $V(v_1, \ldots, v_n) \geq \frac{1}{2} \max_i v_i$

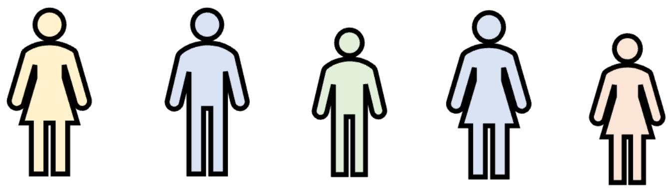
Many proofs and algorithms.

Optimal algorithm by induction	[Hill, Kertz '81]
Median: p such that $\mathbb{P}\left(\max_i v_i \leq p\right) = 1/2$	[Samuel-Cahn '84]
Half of expectation: $p = \frac{1}{2} \cdot \mathbb{E}\left(\max_i v_i\right)$	[Kleinberg, Weinberg STOC'12]
Max of samples: draw $v_i' \sim F_i$ and post $p = \max_i v_i'$	[Rubinstein, Wang, Weinberg ITCS'20]

The secretary problem



1 item

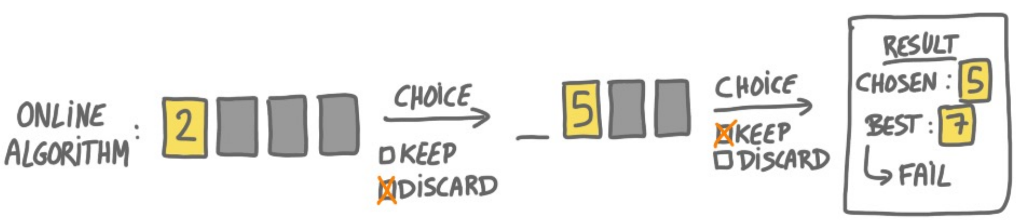


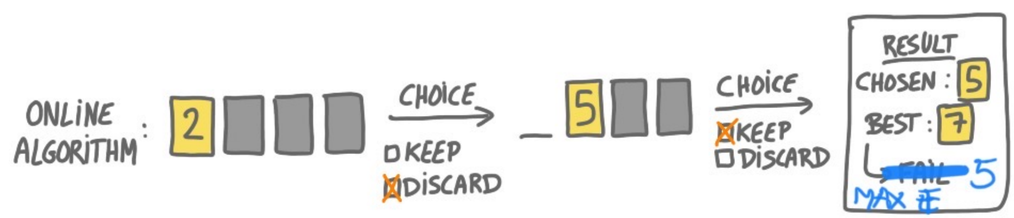
n agents
arbitrary valuations v_i
Random order

- Agents arrive sequentially: we observe $v_1, v_2, \ldots$ one by one
- We (the Decision Maker) decide to stop/continue
- We maximize $\mathbb{P}\left(v_{\text{stop}} = \max_i v_i\right)$

Secretary

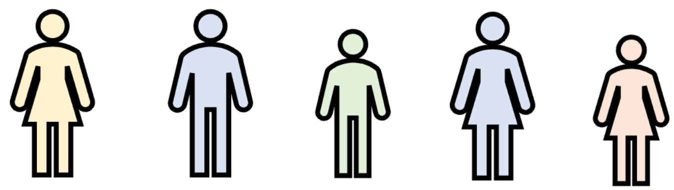
Prophet





The secretary problem

1 item



n agents
arbitrary valuations v_i
Random order

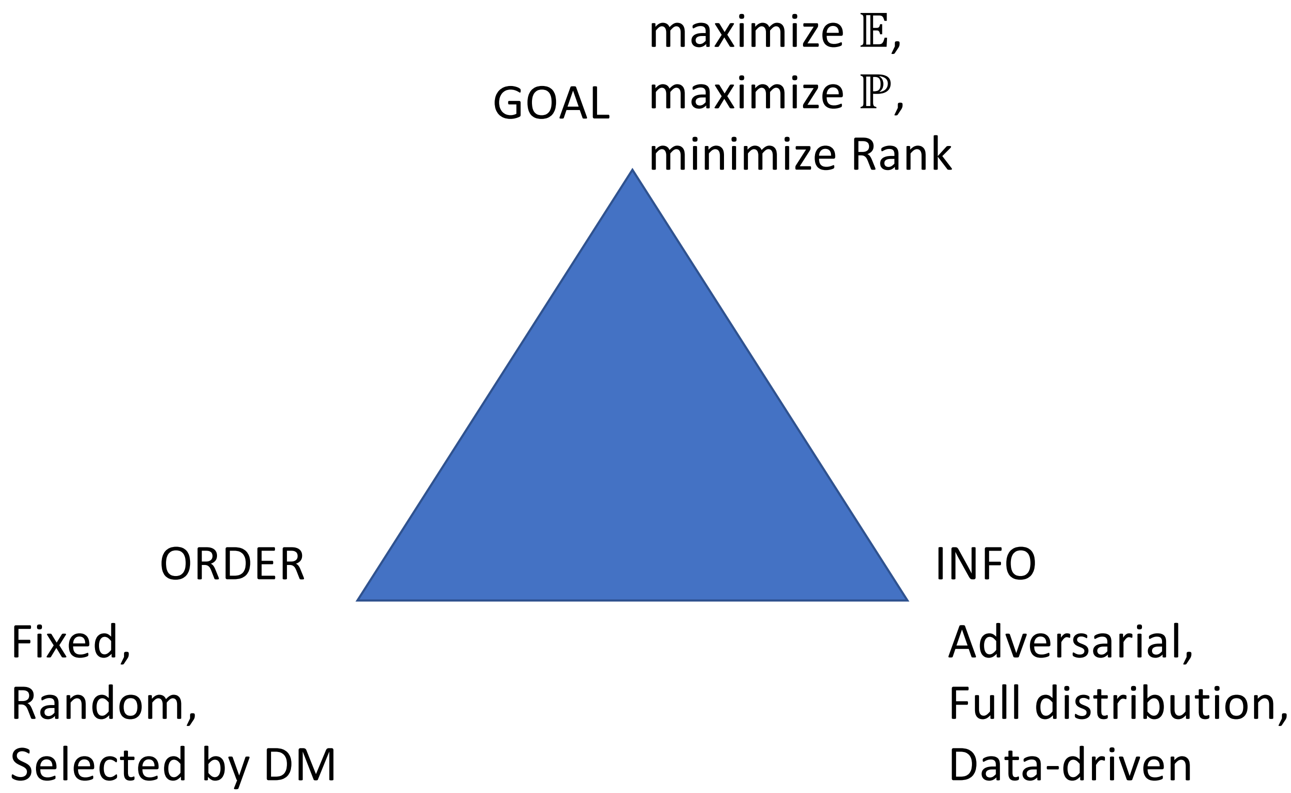
Theorem [Lindley, 1961] The n/e -threshold strategy is optimal and picks the maximum value with probability $1/e$.

The secretary problem

n/e -threshold strategy:

1. Discard the n/e first values but remember the maximum M in this segment.
2. Keep the first value that exceeds M.

Single selection Optimal Stopping



Zoo of Optimal Stopping problems

- Classic Secretary Problem: Pick the best, Random order, Adversarial values [Gardner 60]
  SOLVED: Best algorithms gets 1/e. [Lindley 61, Ferguson 89]
- Full information secretary: Pick the best, (Random order), i.i.d. values with known distribution [Gilbert and Mosteller 66]
  SOLVED: Best algorithm gets 0.58 [Samuels 82]
- Min rank Secretary problem: Minimize expected ranking, Random order
  SOLVED: Best expected ranking is 3.87 [Chow, Moriguti, Robbins, Samuels 64]
- Full information min rank: Minimize expected ranking, Random order, i.i.d. values with known distribution
  OPEN: Best expected ranking is close to 2 but unknown (This is known as Robbins problem)
- Secretary with general objective: Minimize a function of the ranking, Random order
  SOLVED: Optimal algorithm for arbitrary objective [Mucci 73]

Zoo of Optimal Stopping problems

- Classic Prophet Inequality: Max expectation, Adversarial order, independent values with known distribution [Krengel and Sucheston 77]
  SOLVED: Best algorithm gets ½ of the prophet [KS and Garling 78]
- IID prophet: : Max expectation, (Random order), i.i.d. values with known distribution [Hill and Kertz 82]
  SOLVED: Best algorithm gets 0.745 of the prophet [Kertz 86] [Correa et al EC 2017]
- Free order prophet: Max expectation, DM selects the order, independent values with known distribution [Hill 83]
  OPEN: Best known algorithm gets 0.725 of the prophet [Peng and Tang FOCS 2022]
- Prophet Secretary: Max expectation, Random order, independent values with known distribution [Esfandiari et al ESA 2015]
  OPEN: Best known algorithm gets 0.669 of the prophet [C. Saona, Ziliotto SODA 2019]
    No possible to go above 0.725 [Bubna, Chiplunkar 2023+, Mallman-Trenn, Saona 2023+]

Some proofs

- Proof for the Secretary Problem

[Bruss 84]

- Three proofs of the prophet inequality
  - Induction

[Hill and Kertz 81]

  - Balanced prices

[Samuel-Cahn 84]

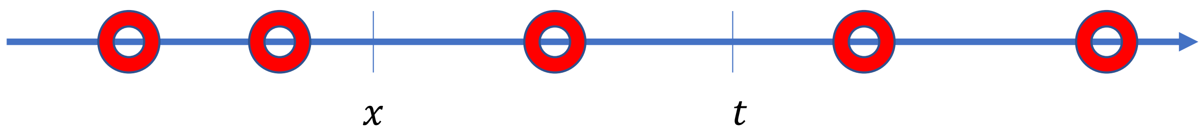
  - Stochastic dominance

[Kleinberg and Weinberg STOC 2012]

- Proof for single threshold prophet secretary

[Eshani et al SODA 2018]
[C. Saona, Ziliotto SODA 2019]

Secretary



1. Assign uniform [0,1] random times to the elements
2. ALG: Do nothing until time x. Then accept first value larger than what you have seen

$$\mathbb{P}(\,STOP \leq t) = 1 - x/t \qquad \text{(STOP by } t \text{ if and only if max is after } x\text{)}$$

$$\mathbb{P}(ALG = \max \mid STOP\ at\ t) = t$$

$$\mathbb{P}(ALG = \max) = \int_x^1 \frac{x}{t^2}\, t dt = -x\ln(x) = 1/e \qquad (\text{ at } x = \frac{1}{e}\,)$$

Prophet Inequality: Induction

- Consider an instance $v_1, \ldots, v_n$and call $V(v_1, \ldots, v_n)$ the optimal value for the DM (Gambler)

  Goal: Lower Bound $$\frac{V(v_1,\ldots,v_n)}{\mathbb{E}\left(\max_i v_i\right)}$$

- Two steps in the proof:
  1. Instance $v_1, \ldots, v_n$ has a larger ratio than instance $\lambda, v_2, \ldots, v_n$ for some constant λ.
  2. Instance $\lambda, v_2, \ldots, v_n$ has a larger ratio than instance $\lambda, v_2, \ldots, v_{n-2}, L$, for some "Long-Shot" L.

Long-shot: random variable that is very large with tiny probability and zero otherwise

- Conclusion then follows easily.
  - Inductively we reduce to two random variables.

$$\frac{V(v_1, v_2)}{\mathbb{E}(\max(v_1, v_2))} \geq \frac{\max(\mathbb{E}(v_1), \mathbb{E}(v_2))}{\mathbb{E}(v_1) + \mathbb{E}(v_2)} \geq \frac{1}{2}$$

Tight Instance

The proof shows that the worst case occurs with two values.

The first is deterministic $v_1 = 1$

And the second is a long-shot $v_2 = \begin{cases} \frac{1}{\varepsilon} & \text{w.p.} \quad \varepsilon \\ 0 & \text{w.p.} \quad 1 - \varepsilon \end{cases}$

The DM gets 1 while the prophet gets $\varepsilon\left(\frac{1}{\varepsilon}\right) + (1 - \varepsilon) \approx 2$

Prophet Inequality: "Balanced prices"

Pick a threshold T and accept first value above it.
Let r be the index of the first v_i above T
$p = \mathbb{P}(\max v_i > T)$

DM $= \mathbb{E}(v_r)$= Revenue + Utility $= pT + \mathbb{E}\big((v_r - T)_+\big)$

Utility

$$\mathbb{E}\big((v_r - T)_+\big) = \sum_{i=1}^{n} \mathbb{E}\big((v_i - T)_+ \mid r = i\big)\, \mathbb{P}(r = i)$$

$$= \sum_{i=1}^{n} \mathbb{E}(v_i - T \mid v_i > T)\mathbb{P}(v_i > T) \prod_{j<i} \mathbb{P}\big(v_j \leq T\big)$$

$$\geq (1-p) \sum_{i=1}^{n} \mathbb{E}\big((v_i - T)_+\big) \geq (1-p)\mathbb{E}(\max(v_i - T)_+)$$

Prophet Inequality: “Balanced prices”

Pick a threshold T and accept first value above it.
Let r be the index of the first v_i above T
$p = \mathbb{P}(\max v_i > T)$

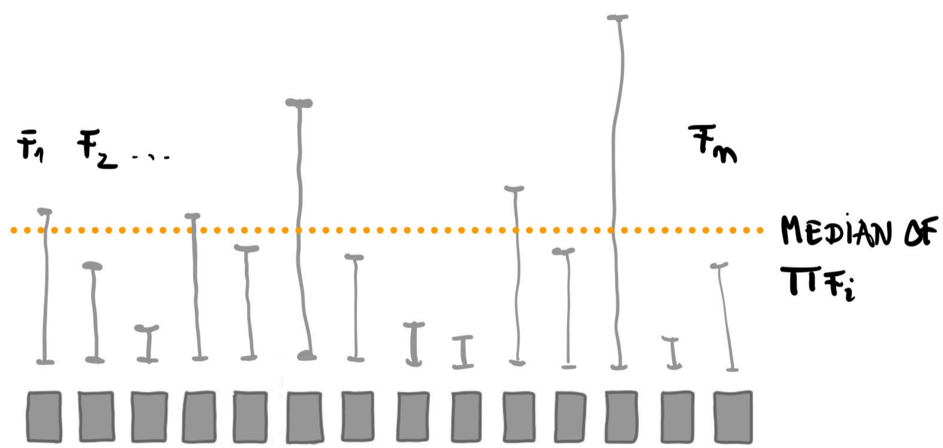
DM gets $\mathbb{E}(v_r) =$ Revenue + Utility

$$= pT + (1-p)\mathbb{E}(\max(v_i - T)$$

$$\geq pT + (1-p)(\mathbb{E}(\max v_i) - T)$$

$T = \frac{\mathbb{E}(\max v_i)}{2}$ or $T = median$

$$\geq \frac{\mathbb{E}(\max v_i)}{2}$$



Prophet Inequality: Stochastic dominance

Again pick a threshold T and accept first value above it.
Let r be the index of the first v_i above T
$p = \mathbb{P}(\max v_i > T)$

Note that:

$$\mathbb{P}(v_r > x) \geq \begin{cases} p & x \leq T \\ (1-p)\mathbb{P}(\max v_i > x) & x > T \end{cases}$$

$$\mathbb{P}(v_r > x) = \sum_{i=1}^{n} \mathbb{P}(v_i > x) \prod_{j<i} \mathbb{P}\left(v_j \leq T\right)$$

$$\geq \ (1-p)\ \sum_{i=1}^{n} \mathbb{P}(v_i > x) \geq (1-p)\mathbb{P}(\max v_i > x)$$

Prophet Inequality: Stochastic dominance

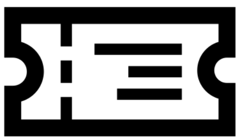
Again pick a threshold T and accept first value above it.
Let r be the index of the first v_i above T
$p = \mathbb{P}(\max v_i > T)$

$$\mathbb{E}(v_r) = \int_0^\infty \mathbb{P}(v_r > x) \geq \int_0^T p dx + \int_T^\infty (1-p)\mathbb{P}(\max v_i > x)$$

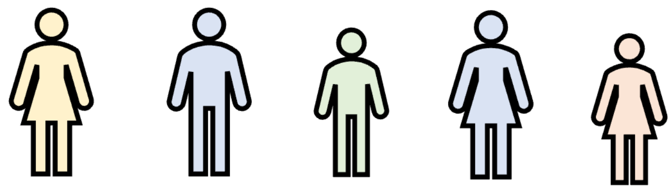
$$\mathbb{E}(\max v_i) = \int_0^\infty \mathbb{P}(\max v_i > x) \leq T + \int_T^\infty \mathbb{P}(\max v_i > x)$$

$$\geq pT + (1-p)(\mathbb{E}(\max v_i) - T) \underbrace{\geq \frac{\mathbb{E}(\max v_i)}{2}}_{T = \frac{\mathbb{E}(\max v_i)}{2} \text{ or } T = median}$$

Prophet Secretary



1 item



n agents
independent valuations
$v_i \sim F_i$

- Arrival order is random and unknown to the DM
- Optimal online algorithm: Large Dynamic program (unknown complexity!)
- Best possible prophet inequality also open.

$$\mathbb{E}[\text{Best online algorithm}] \geq 0.669\ \mathbb{E}[\max\{v_1, \dots, v_n\}]$$

- Constant cannot be better than 0.7251.

Prophet Secretary: Sinlge Threshold

- Consider n random variables.
- $n-1$ are deterministic equal to 1.
- The other gives n with probability $1/n$ and 0 otherwise.

$$\mathbb{E}[\max\{v_1, \dots, v_n\}] = n \times \frac{1}{n} + 1 \times \left(1 - \frac{1}{n}\right) \approx 2$$

- Now fix a threshold T.

  - If $T < 1$, DM gets n w.p. $1/n^2$, and 1 otherwise. So $\mathbb{E}[DM] \approx 1$
  - If $T > 1$, DM gets n w.p. $1/n$. So $\mathbb{E}[DM] \approx 1$

- So ½ is best possible.

Prophet Secretary: Single Threshold

- Consider n random variables.
- $n - 1$ are deterministic equal to 1.
- The other gives n with probability $1/n$ and 0 otherwise.

$$\mathbb{E}[\max\{v_1, \dots, v_n\}] = n \times \frac{1}{n} + 1 \times \left(1 - \frac{1}{n}\right) \approx 2$$

- To beat ½ RANDOM threshold T.
  - Set $T = 1$, and break ties at random. Accept a 1 w.p. $1/n$

$$\mathbb{P}(\text{DM gets something}) = 1 - \left(1 - \frac{1}{n}\right)^n \approx 1 - \frac{1}{e}$$

$$\mathbb{E}(\text{DM} \mid \text{DM gets something}) = n \times \frac{1}{n} + 1 \times \left(1 - \frac{1}{n}\right) \approx 2$$

$$\Longrightarrow \quad \mathbb{E}(\text{DM}) \approx 2 \times \left(1 - \frac{1}{e}\right) \approx \left(1 - \frac{1}{e}\right) \mathbb{E}[\max v_i]$$

Prophet Secretary: Single Threshold

- Can get $1 - 1/e$ in general
- Proof for continuous and strictly increasing distributions
- Set T such that $\mathbb{P}(\max v_i > T) = 1 - \frac{1}{e}$
- Let r be the index of the first v_i above T

  - Theorem: $\mathbb{E}[\, v_r] \geq \left(1 - \frac{1}{e}\right) \mathbb{E}[\max v_i]$

- We prove that $\mathbb{P}(v_r > x) \geq (1 - 1/e)\mathbb{P}(\max v_i > x)$

- Then conclude by integrating.

- Note that if $x \leq T$, $\mathbb{P}(v_r > x) = \mathbb{P}(v_r > T) \geq \mathbb{P}(\max v_i > T) = 1 - 1/e$

Prophet Secretary: Single Threshold

- Take $x > T$, and let σ be a random permutation. v_i comes at time $\sigma(i)$.

$$\frac{\mathbb{P}\big(r = \sigma(i)\big)}{\mathbb{P}(v_i > T)} = \mathbb{P}(r = \sigma(i) | v_i > T) \;\; = \sum_{S \subset [n] \setminus i} \frac{1}{1 + |S|} \prod_{j \in S} (1 - F_j(T)) \prod_{j \notin S, i \neq i} F_j(T)$$

$$\geq \left(1 - \frac{1}{e}\right)$$ Schur-convex function

$$\mathbb{P}(v_r > x) = \sum_{i=1}^{n} \mathbb{P}(v_i > x |\, r = \sigma(i)) \;\; \mathbb{P}(r = \sigma(i))$$

$$\geq \sum_{i=1}^{n} \mathbb{P}(v_i > x | v_i > T) \left(1 - \frac{1}{e}\right) \mathbb{P}(v_i > T) = \left(1 - \frac{1}{e}\right) \sum_{i=1}^{n} \frac{\mathbb{P}(v_i > x)}{\mathbb{P}(v_i > T)} \mathbb{P}(v_i > T)$$

$$\geq \left(1 - \frac{1}{e}\right) \mathbb{P}(\max v_i > x)$$