The Multisecretary Problem with many types

Omar Besbes, Yash Kanoria, **Akshit Kumar**

Columbia Business School

» Multi-secretary Problem

Problem Statement

Given a sequence of T secretaries and a hiring budget B, a decision maker (DM) wants to hire the top B secretaries in terms of their ability.

» Multi-secretary Problem

Problem Statement

Given a sequence of T secretaries and a hiring budget B, a decision maker (DM) wants to hire the top B secretaries in terms of their ability.

Details and Assumptions

* The secretaries arrive in an online fashion.

» Multi-secretary Problem

Problem Statement

Given a sequence of T secretaries and a hiring budget B, a decision maker (DM) wants to hire the top B secretaries in terms of their ability.

Details and Assumptions

* The secretaries arrive in an online fashion.
* The DM makes **irrevocable** *hire* or *reject* decisions.

» Multi-secretary Problem

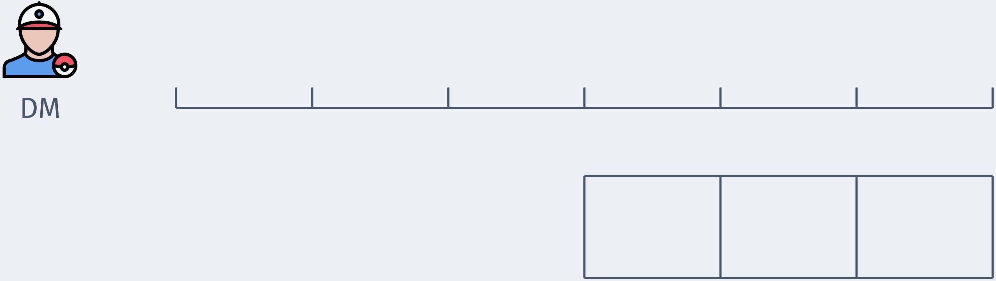
Problem Statement

Given a sequence of T secretaries and a hiring budget B, a decision maker (DM) wants to hire the top B secretaries in terms of their ability.

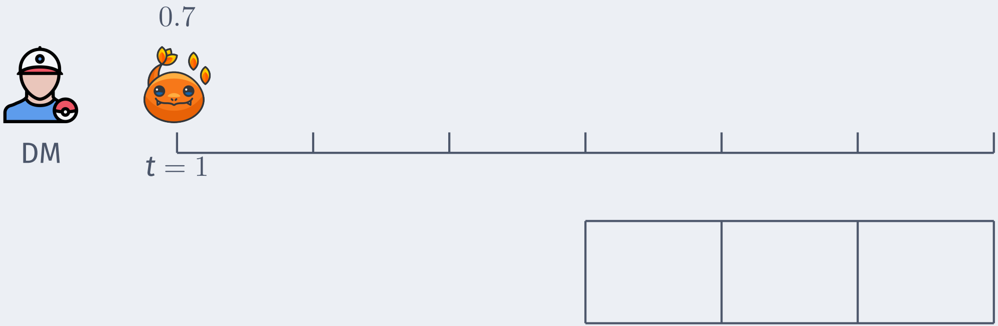
Details and Assumptions

- The secretaries arrive in an online fashion.
- The DM makes **irrevocable** *hire* or *reject* decisions.
- The abilities (types) of the secretaries are drawn **independently** from a **common** and **known** distribution F over $[0, 1]$.

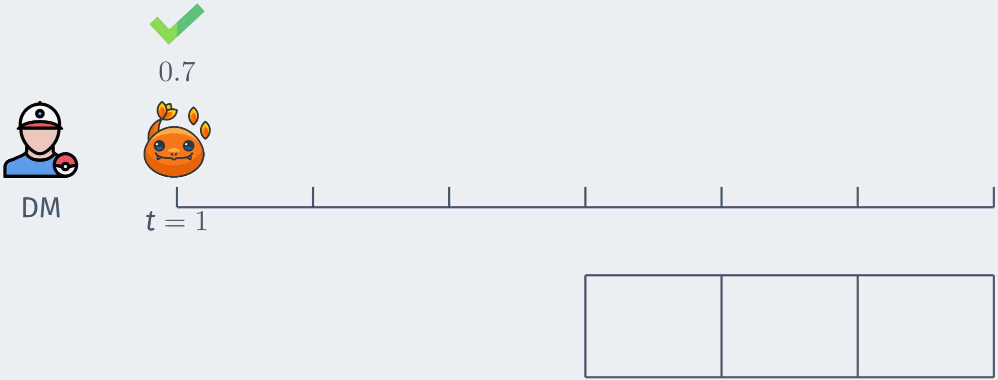
» Gotta catch'em ~~all~~ some



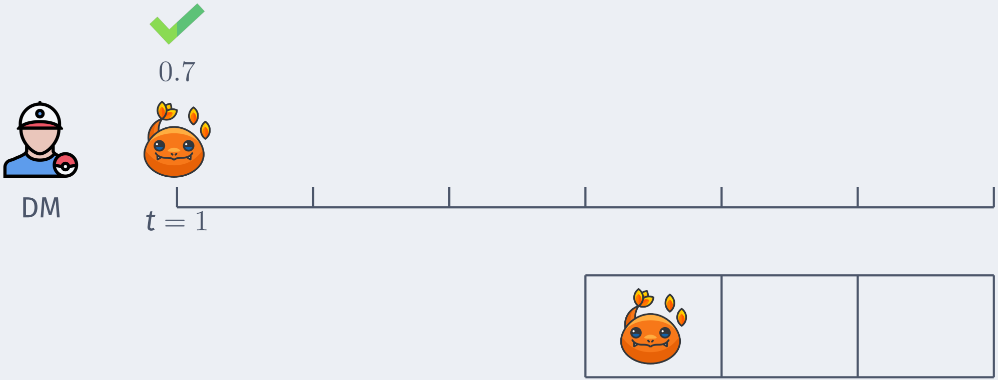
» Gotta catch'em ~~all~~ some



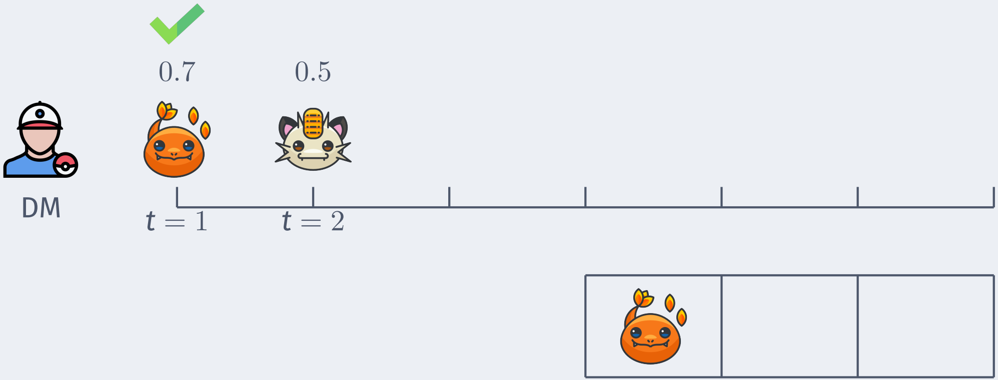
» Gotta catch'em ~~all~~ some



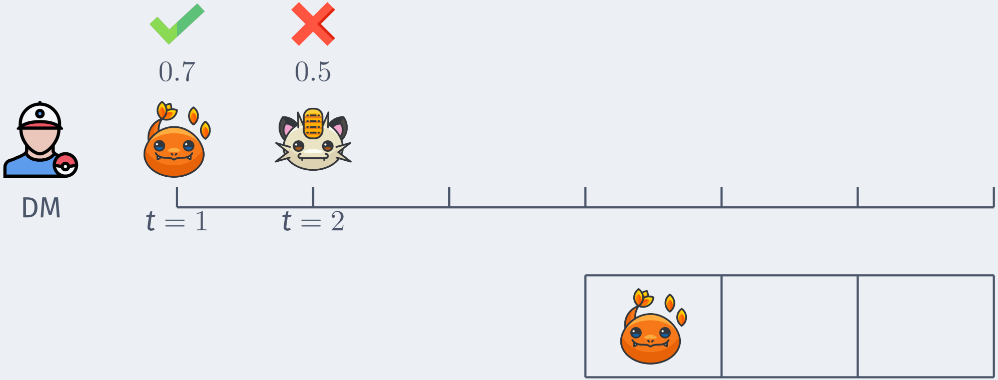
» Gotta catch'em ~~all~~ some



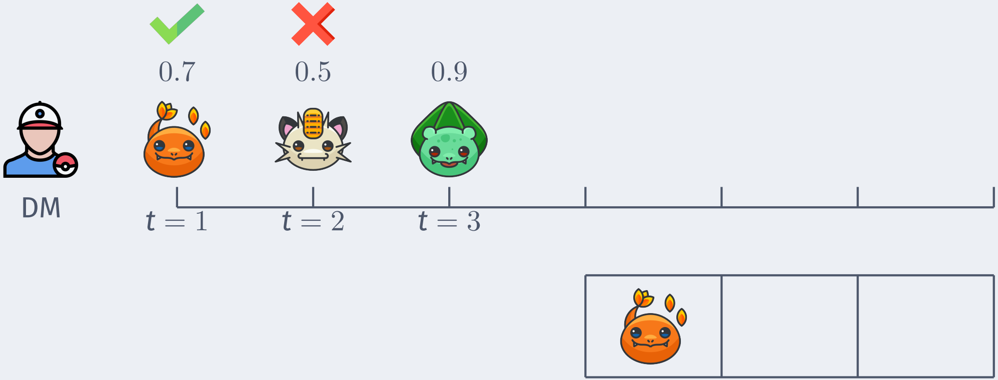
» Gotta catch'em ~~all~~ some



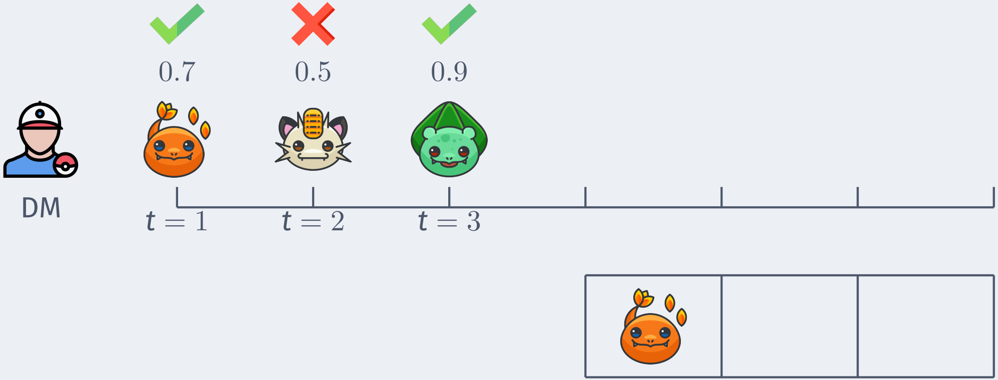
» Gotta catch'em ~~all~~ some



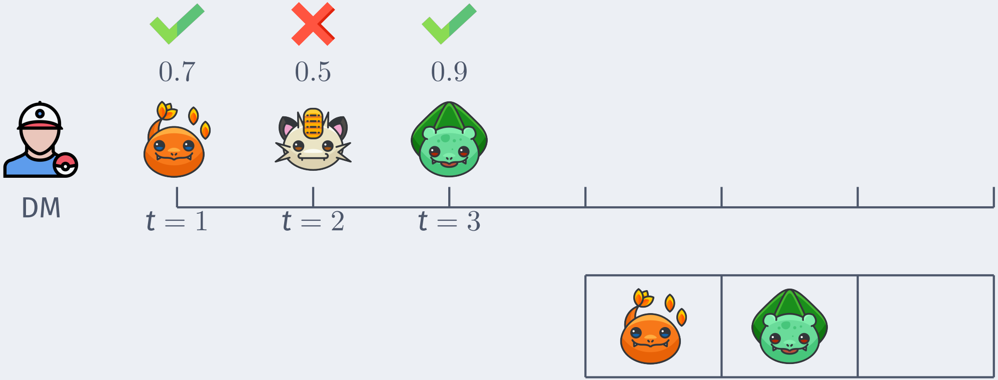
» Gotta catch'em ~~all~~ some



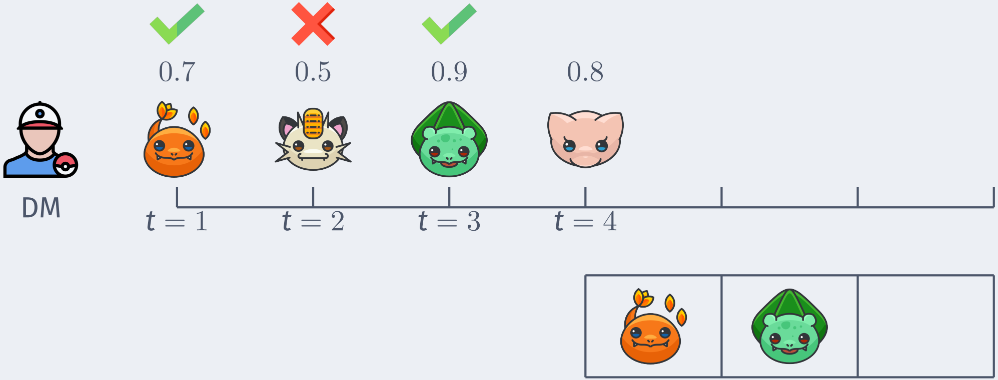
» Gotta catch'em ~~all~~ some



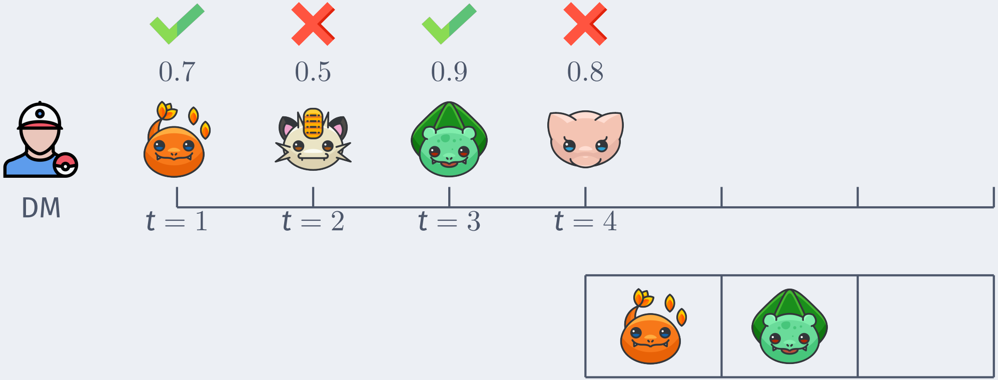
» Gotta catch'em ~~all~~ some



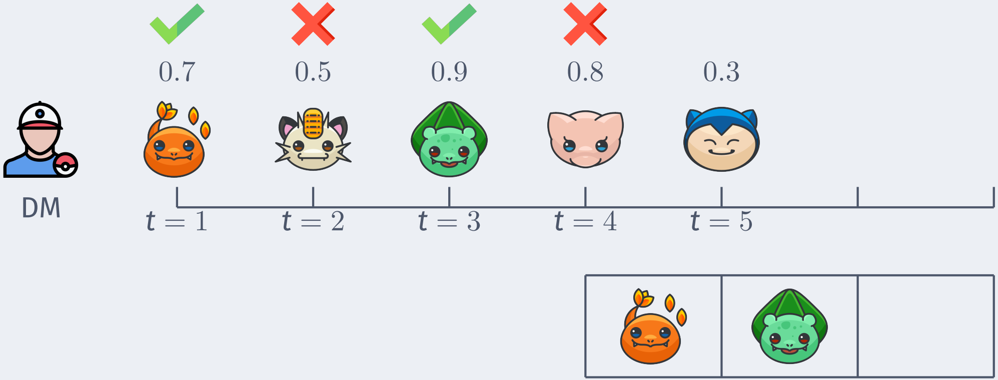
» Gotta catch'em ~~all~~ some



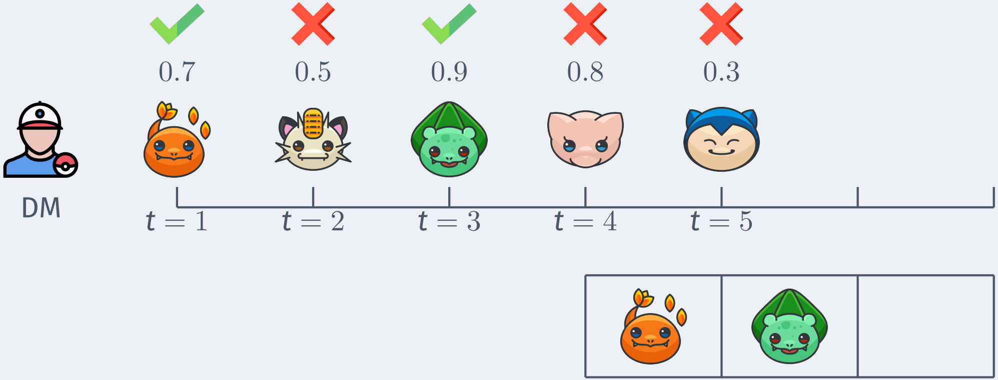
» Gotta catch'em ~~all~~ some



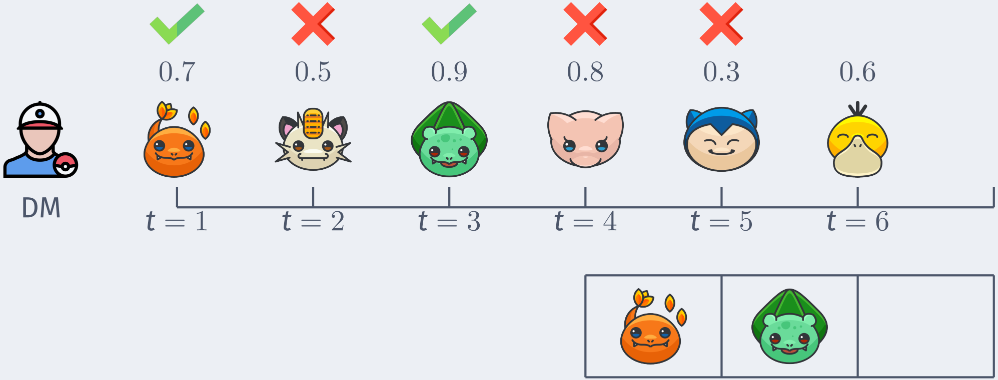
» Gotta catch'em ~~all~~ some



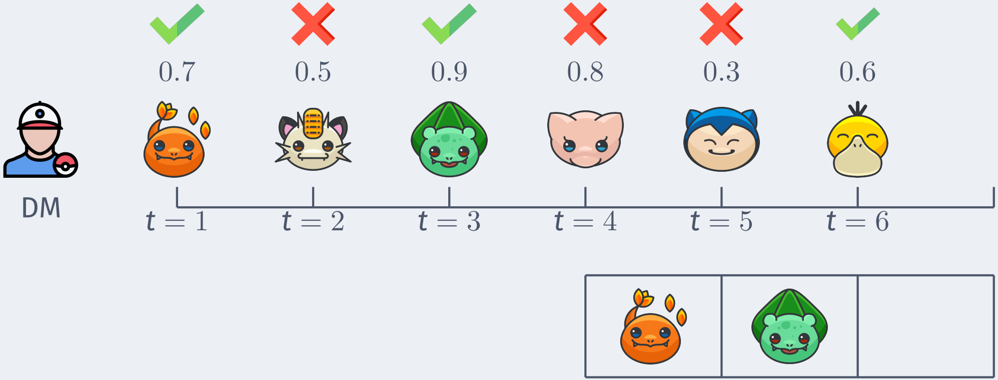
» Gotta catch’em ~~all~~ some



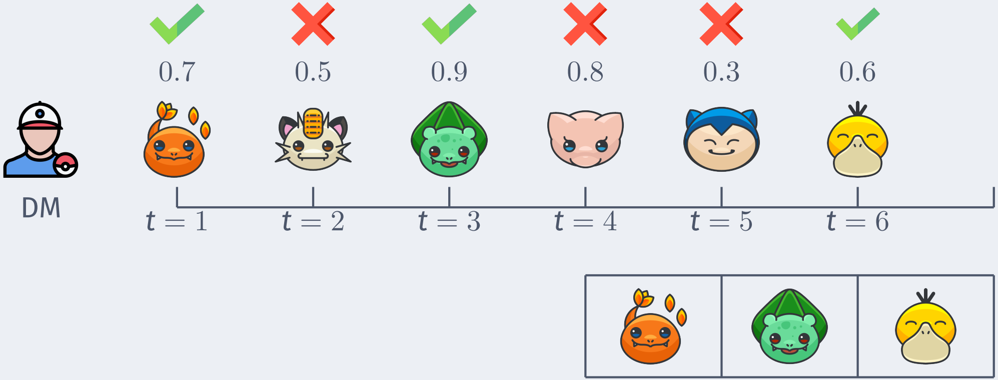
» Gotta catch'em ~~all~~ some



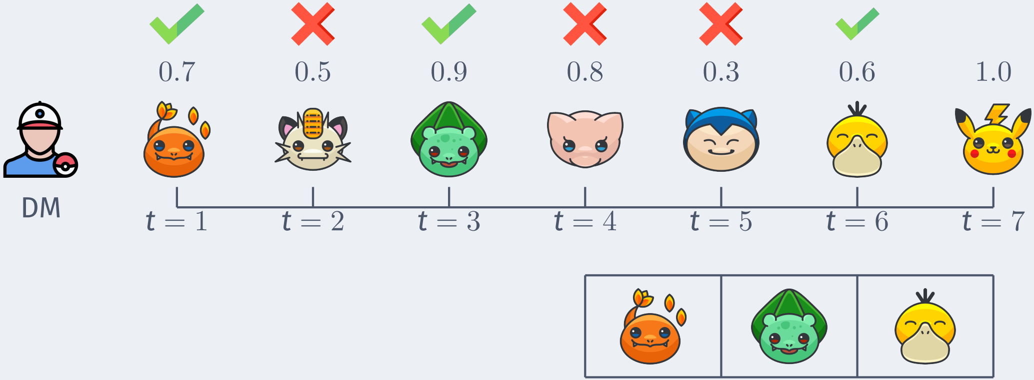
» Gotta catch'em ~~all~~ some



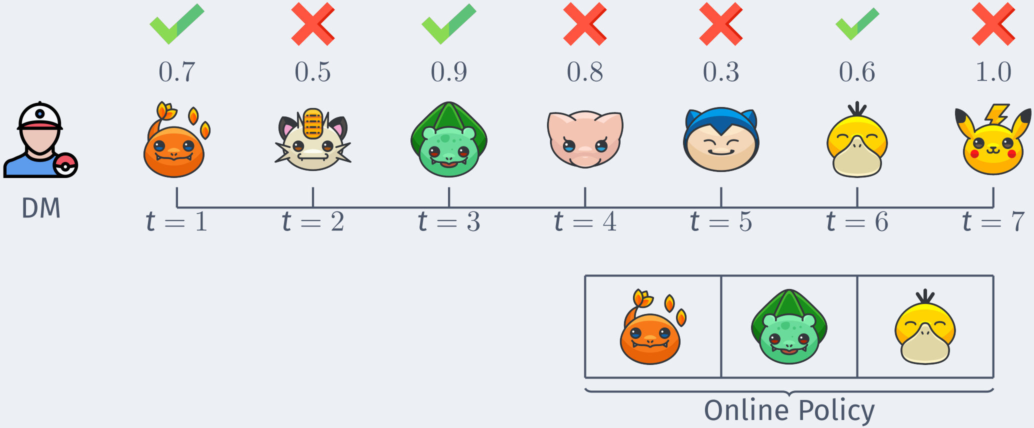
» Gotta catch'em ~~all~~ some



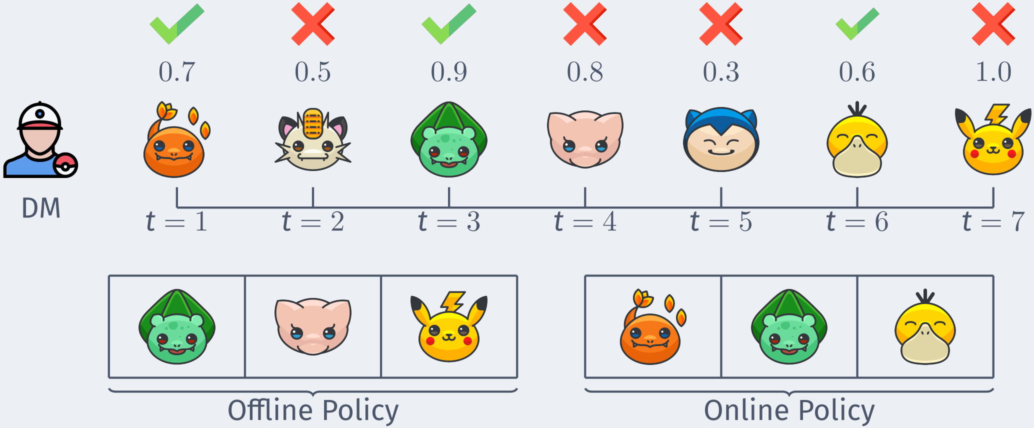
» Gotta catch'em ~~all~~ some



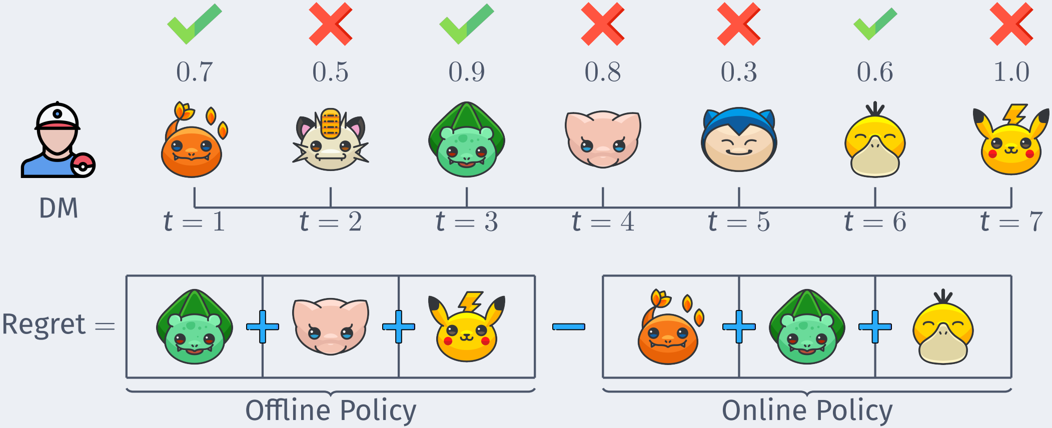
» Gotta catch'em ~~all~~ some



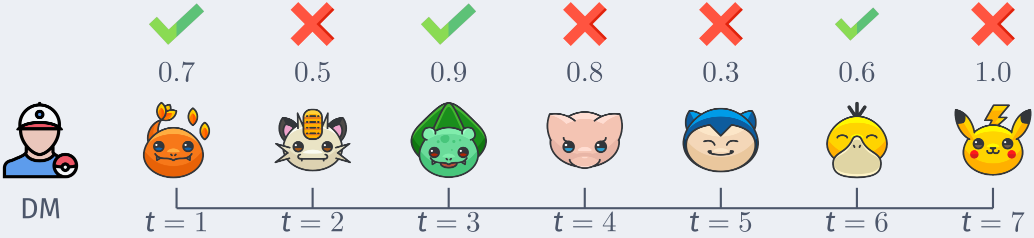
» Gotta catch'em ~~all~~ some



» Gotta catch'em ~~all~~ some

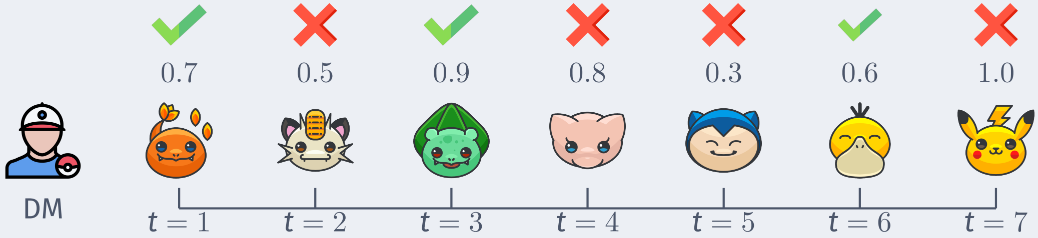


» Gotta catch'em ~~all~~ some



$$\text{Regret} = \underbrace{\boxed{0.9 + 0.8 + 1.0}}_{\text{Offline Policy}} - \underbrace{\boxed{0.7 + 0.9 + 0.6}}_{\text{Online Policy}} = 0.5$$

» Gotta catch'em ~~all~~ some



$$\text{Regret} = \underbrace{\boxed{0.9 + 0.8 + 1.0}}_{\text{Offline Policy}} - \underbrace{\boxed{0.7 + 0.9 + 0.6}}_{\text{Online Policy}} = 0.5$$

We will focus on $\mathbb{E}\,[\text{Regret}]$

» Motivation and Related Applications

* Online Knapsack Problem [Kleywegt and Papastavrou, 1998]
    * Unit weights and *i.i.d* rewards

» Motivation and Related Applications

* Online Knapsack Problem [Kleywegt and Papastavrou, 1998]
  * Unit weights and *i.i.d* rewards
* Single-leg Revenue Management [Talluri et al., 2004]
  * Selling horizon T, flight capacity B and types are different fare classes.

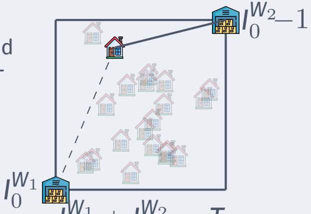
» Motivation and Related Applications

* Online Knapsack Problem [Kleywegt and Papastavrou, 1998]
  * Unit weights and *i.i.d* rewards
* Single-leg Revenue Management [Talluri et al., 2004]
  * Selling horizon T, flight capacity B and types are different fare classes.
* Order Fulfillment Problem [Lei et al., 2018]

» Motivation and Related Applications

- Online Knapsack Problem [Kleywegt and Papastavrou, 1998]
  - Unit weights and *i.i.d* rewards
- Single-leg Revenue Management [Talluri et al., 2004]
  - Selling horizon T, flight capacity B and types are different fare classes.
- Order Fulfillment Problem [Lei et al., 2018]

Spatially distributed demand
2 WHs with total inventory T
Unit demand at each time
Fulfill from one of the WHs
Inventory depletes by 1
min total matching cost



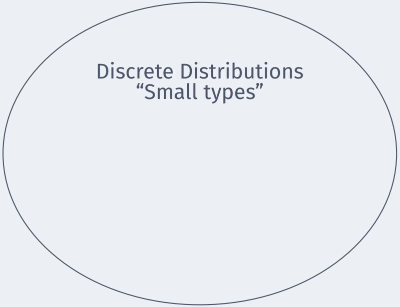
$I_0^{W_1} + I_0^{W_2} = T$

Reduces to multi-secretary problem

hire $\iff$ match to W_2
reject $\iff$ match to W_1

» What's known and research question?

» What's known and research question?



» What's known and research question?

Discrete Distributions
"Small types"

» What's known and research question?

Discrete Distributions
"Small types"

Regret $= \Theta(1)$
[Arlotto and Gurvich, 2019]

» What's known and research question?

Discrete Distributions
"Small types"

Regret $= \Theta(1)$

[Arlotto and Gurvich, 2019]

Regret can scale with # of types

» What's known and research question?

Discrete Distributions
"Small types"

Regret $= \Theta(1)$

[Arlotto and Gurvich, 2019]

Regret can scale with # of types

Continuous Distributions
"Many Types"

» What's known and research question?

Discrete Distributions
"Small types"

Regret $= \Theta(1)$
[Arlotto and Gurvich, 2019]
Regret can scale with # of types

Continuous Distributions
"Many Types"

» What's known and research question?

Discrete Distributions
"Small types"

Regret $= \Theta(1)$
[Arlotto and Gurvich, 2019]
Regret can scale with # of types

Continuous Distributions
"Many Types"

Regret $= \Theta(\log T)$
[Bray, 2022], [Lueker, 1998]

» What's known and research question?

Discrete Distributions
"Small types"

Regret $= \Theta(1)$
[Arlotto and Gurvich, 2019]
Regret can scale with # of types

Continuous Distributions
"Many Types"

Regret $= \Theta(\log T)$
[Bray, 2022], [Lueker, 1998]
Need density to be bounded below

» What's known and research question?

Discrete Distributions
"Small types"

Regret $= \Theta(1)$

[Arlotto and Gurvich, 2019]

Regret can scale with # of types

Continuous Distributions
"Many Types"

Regret $= \Theta(\log T)$

[Bray, 2022], [Lueker, 1998]

Need density to be bounded below

» What's known and research question?

Discrete Distributions
“Small types”

Regret $= \Theta(1)$

[Arlotto and Gurvich, 2019]

Regret can scale with # of types

Continuous Distributions
“Many Types”

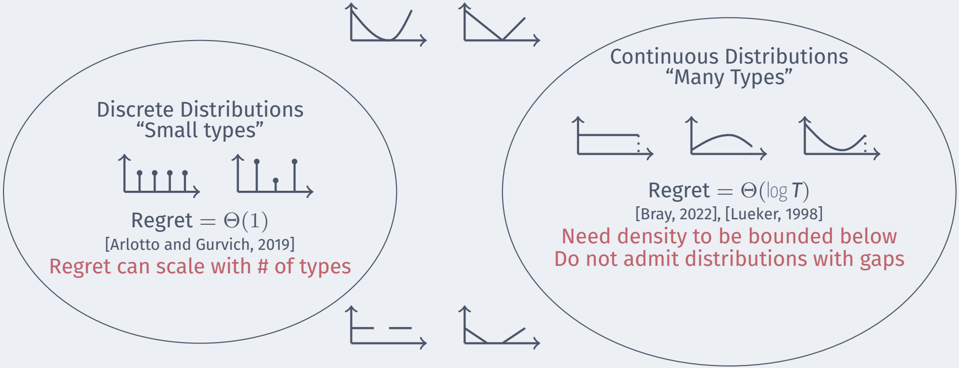
Regret $= \Theta(\log T)$

[Bray, 2022], [Lueker, 1998]

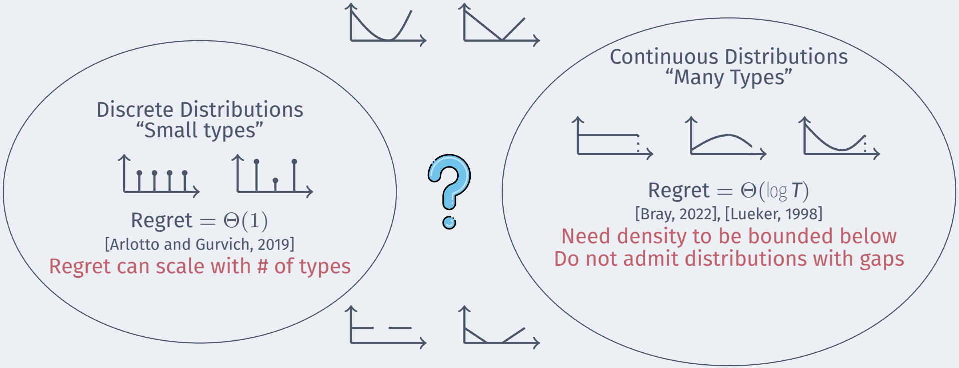
Need density to be bounded below

Do not admit distributions with gaps

» What's known and research question?



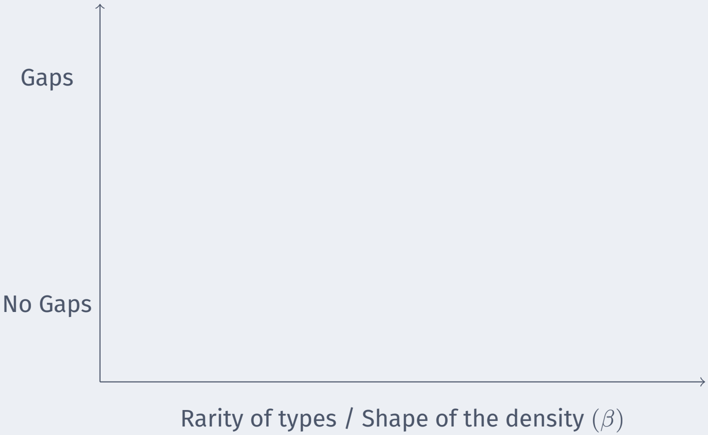
» What's known and research question?



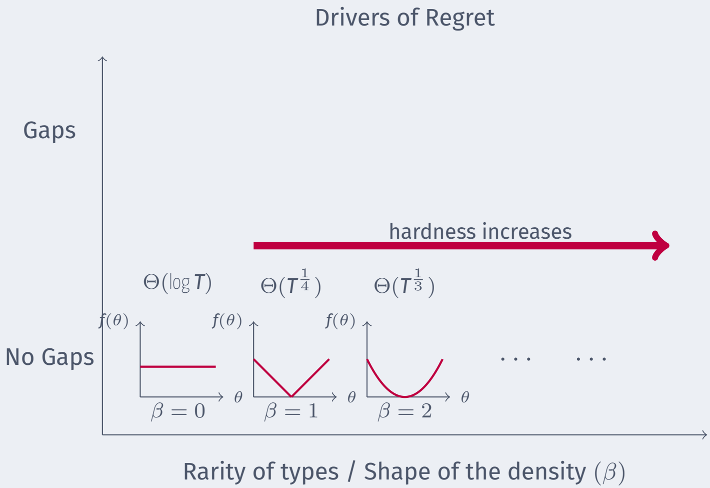
How to interpolate between these distributions in a unified manner?
How to deal with gaps ? What are the possible regret scalings ?

Punchline and Overview

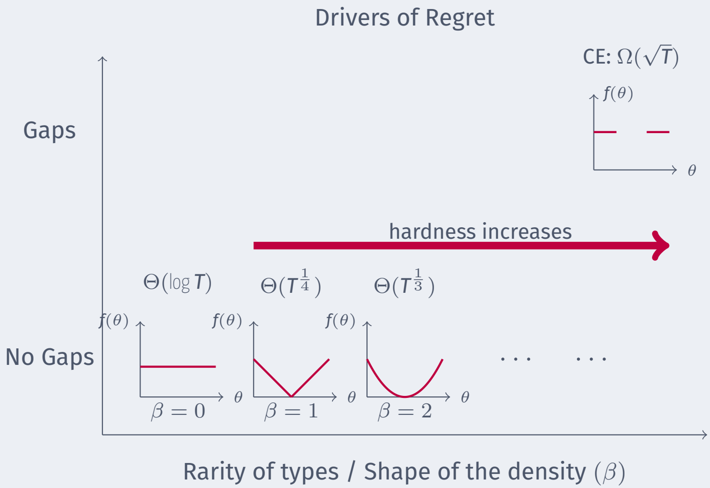
Drivers of Regret



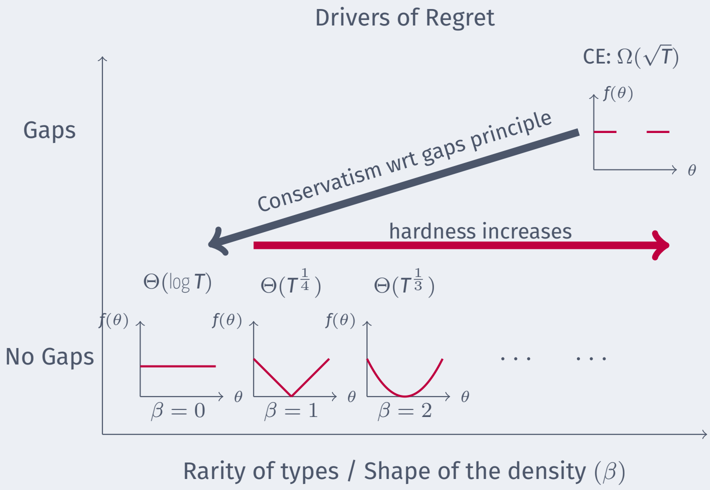
Punchline and Overview



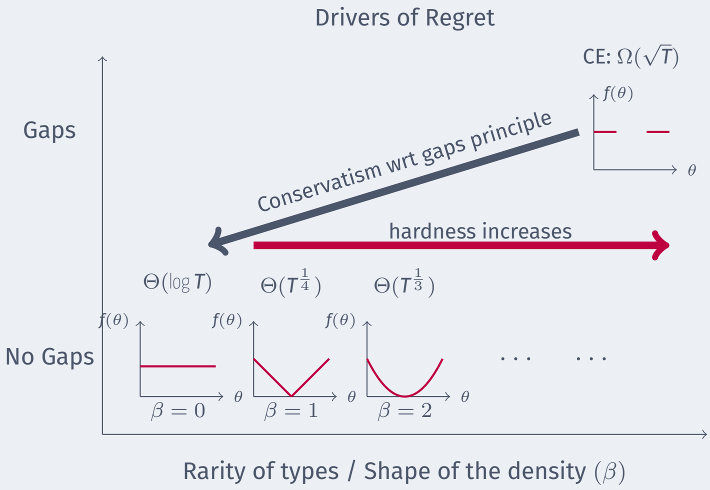
» Punchline and Overview



» Punchline and Overview



» Punchline and Overview



- Shape is a fundamental driver of regret
- Dealing with gaps is an algorithmic challenge.
- New principle: Conservatism wrt gaps
- If close to the gap, use the gap.

» First step towards unified approach

Discrete Distributions
“Small types”

Regret $= \Theta(1)$
[Arlotto and Gurvich, 2019]
Regret can scale with # of types

Continuous Distributions
“Many Types”

Regret $= \Theta(\log T)$
[Bray, 2022], [Lueker, 1998]
~~Need density to be bounded below~~
Do not admit distributions with gaps

» First step towards unified approach

Discrete Distributions
“Small types”

Regret $= \Theta(1)$
[Arlotto and Gurvich, 2019]
Regret can scale with # of types

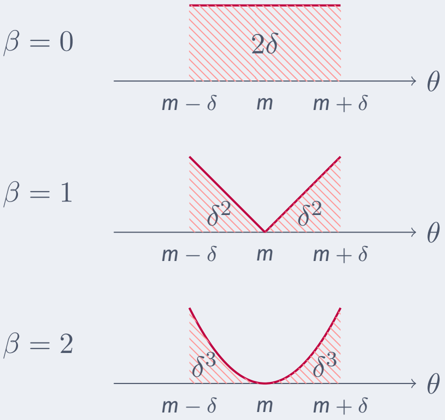
$(\beta, 1)$-clustered distributions

Continuous Distributions
“Many Types”

Regret $= \Theta(\log T)$
[Bray, 2022], [Lueker, 1998]
~~Need density to be bounded below~~
Do not admit distributions with gaps

» $(\beta, 1)$-clustered distribution

Definition



* Consider a continuous distribution with finite many points at which the density is zero.
* β characterizes the rate of mass accumulation around the points of zero density.
* Consider m such that $f(m) = 0$, then F is $(\beta, 1)$-clustered if

$$|F(m \pm \delta) - F(m)| \geq \delta^{\beta+1}$$

» $(\beta, 1)$-clustered distributions

Optimal Regret Scaling

Universal Lower Bound

For every $\beta \in [0, \infty)$, there exists a distribution F_β such that

$$\sup_{B \in [T]} \mathbb{E}_{F_\beta}[\text{Regret}] = \begin{cases} \Omega(\log T), & \beta = 0, \\ \Omega\left(T^{\frac{1}{2} - \frac{1}{2(1+\beta)}}\right), & \beta > 0. \end{cases}$$

» $(\beta, 1)$-clustered distributions

Optimal Regret Scaling

Universal Lower Bound

For every $\beta \in [0, \infty)$, there exists a distribution F_β such that

$$\sup_{B \in [T]} \mathbb{E}_{F_\beta}[\text{Regret}] = \begin{cases} \Omega(\log T), & \beta = 0, \\ \Omega\left(T^{\frac{1}{2} - \frac{1}{2(1+\beta)}}\right), & \beta > 0. \end{cases}$$

Regret of Certainty Equivalent

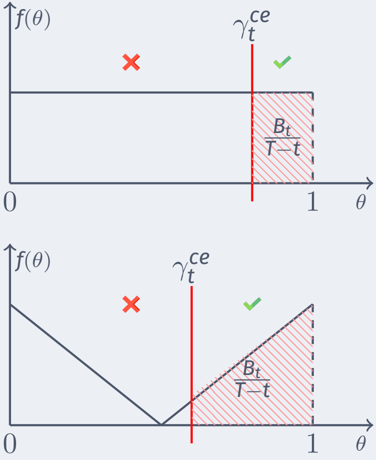
If F is a $(\beta, 1)$-clustered distribution, then for all $B \in [T]$,

$$\mathbb{E}[\text{Regret}(CE)] = \begin{cases} \mathcal{O}(\log T), & \beta = 0, \\ \mathcal{O}\left(T^{\frac{1}{2} - \frac{1}{2(1+\beta)}}\right), & \beta > 0. \end{cases}$$

» Certainty Equivalent Control

For $(\beta, 1)$-continuous distributions

* Let B_t be the remaining budget at time t
* Compute the budget ratio
  $br_t = \frac{\text{Remaining Budget}}{\text{Remaining Time}} = \frac{B_t}{T-t}$
* Define a quantile threshold $p_t^{ce} = 1 - br_t$
* Define a ability threshold $\gamma_t^{ce} = F^{-1}(p_t^{ce})$
* hire $\iff \theta_t \geq \gamma_t^{ce}$



» $(\beta, 1)$-clustered distributions

- Relax the assumption of density being positively lower bounded for continuous distribution.
- The lower bound along with the upper bound show that β (shape of the density) is a fundamental driver of regret.
- Entire spectrum of regret scaling is possible: $\Theta(\log T)$ or $\Theta(T^\alpha)$ where $\alpha \in (0, \frac{1}{2})$.
- A simple heuristic like CE is optimal.

» Gaps in the distribution

Discrete Distributions
"Small types"

Regret $= \Theta(1)$
[Arlotto and Gurvich, 2019]
Regret can scale with # of types

Continuous Distributions
"Many Types"

Regret $= \Theta(\log T)$
[Bray, 2022], [Lueker, 1998]
~~Need density to be bounded below~~
Do not admit distributions with gaps

» Gaps in the distribution

Discrete Distributions
“Small types”

Regret $= \Theta(1)$

[Arlotto and Gurvich, 2019]

Regret can scale with # of types

$(\beta, 1)$-clustered distributions

Continuous Distributions
“Many Types”

Regret $= \Theta(\log T)$

[Bray, 2022], [Lueker, 1998]

~~Need density to be bounded below~~

Do not admit distributions with gaps

» Gaps in the distribution

Discrete Distributions
"Small types"

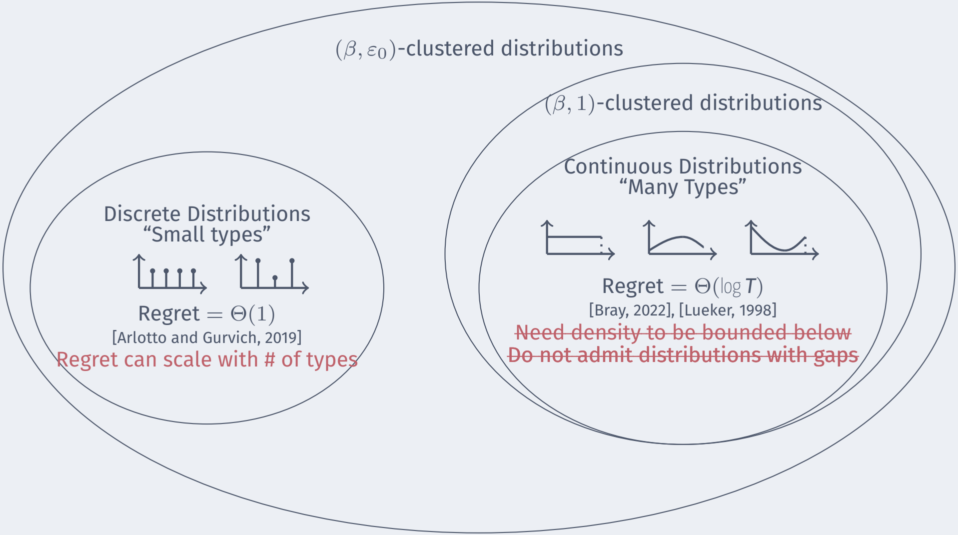
Regret $= \Theta(1)$
[Arlotto and Gurvich, 2019]
Regret can scale with # of types

$(\beta, 1)$-clustered distributions

Continuous Distributions
"Many Types"

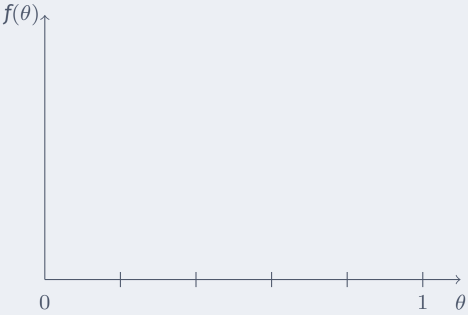
Regret $= \Theta(\log T)$
[Bray, 2022], [Lueker, 1998]
~~Need density to be bounded below~~
~~Do not admit distributions with gaps~~

» Gaps in the distribution



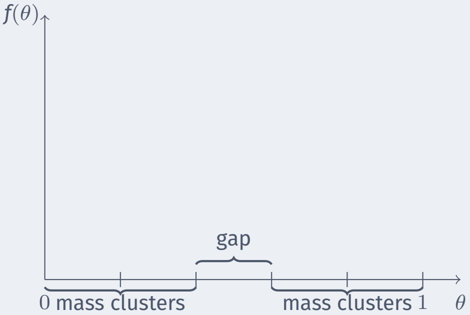
» (β, ε_0)-clustered distribution

Examples



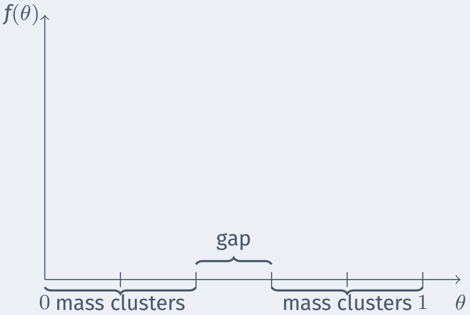
» (β, ε_0)-clustered distribution

Examples



» (β, ε_0)-clustered distribution

Examples

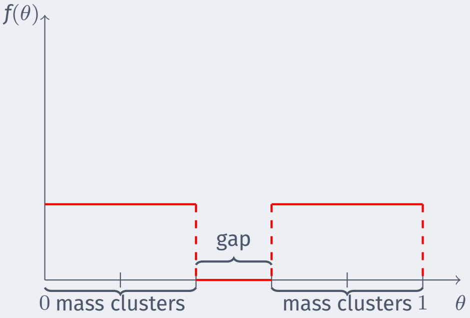


Gap $\equiv$ intervals of positive length with zero mass

mass cluster $\equiv$ interval with positive mass

» (β, ε_0)-clustered distribution

Examples

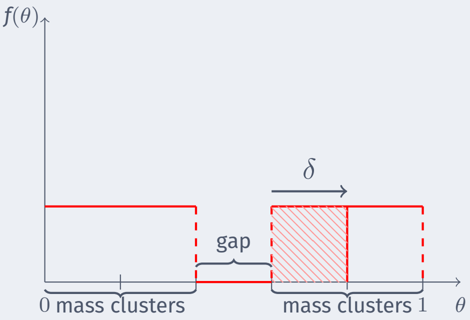


Gap $\equiv$ intervals of positive length with zero mass

mass cluster $\equiv$ interval with positive mass

$\beta = 0$ (mass accumulation around gaps)

» (β, ε_0)-clustered distribution



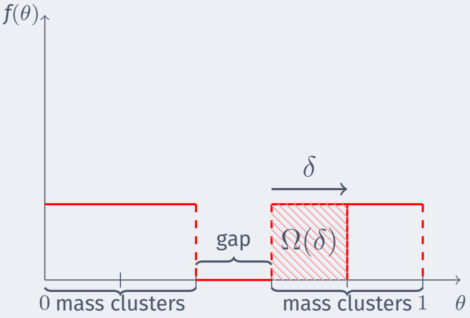
Gap $\equiv$ intervals of positive length with zero mass

mass cluster $\equiv$ interval with positive mass

$\beta = 0$ (mass accumulation around gaps)

» (β, ε_0)-clustered distribution

Examples



Gap $\equiv$ intervals of positive length with zero mass

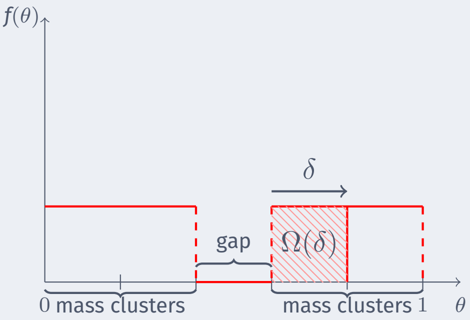
mass cluster $\equiv$ interval with positive mass

$\beta = 0$ (mass accumulation around gaps)

$|F(m + \delta) - F(m)| \geq \delta$ on the same mass cluster

» (β, ε_0)-clustered distribution

Examples



Gap $\equiv$ intervals of positive length with zero mass

mass cluster $\equiv$ interval with positive mass

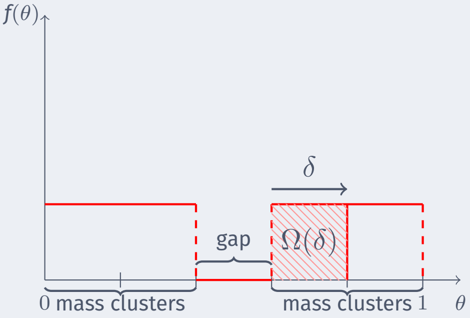
$\beta = 0$ (mass accumulation around gaps)

$|F(m+\delta) - F(m)| \geq \delta$ on the same mass cluster

$\mu(\text{mass clusters}) \geq \varepsilon_0$

» (β, ε_0)-clustered distribution

Examples



Gap $\equiv$ intervals of positive length with zero mass

mass cluster $\equiv$ interval with positive mass

$\beta = 0$ (mass accumulation around gaps)

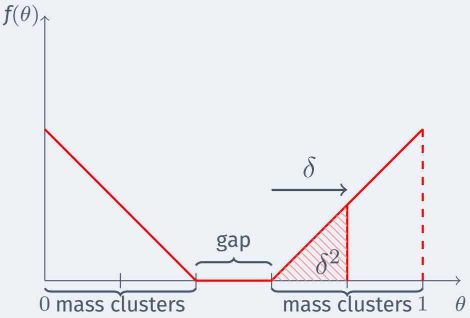
$|F(m+\delta) - F(m)| \geq \delta$ on the same mass cluster

μ(mass clusters) $\geq \varepsilon_0$

For discrete distrbutions, $\beta = 0, \varepsilon_0 = \min_j p_j$

» (β, ε_0)-clustered distribution

Examples



Gap $\equiv$ intervals of positive length with zero mass

mass cluster $\equiv$ interval with positive mass

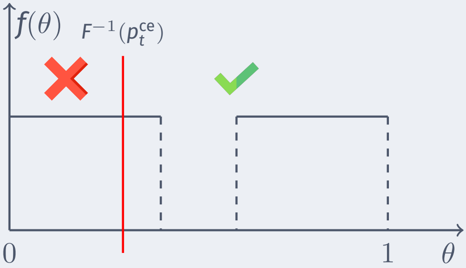
$\beta = 1$ (mass accumulation around gaps)

$|F(m+\delta) - F(m)| \geq \delta^2$ on the same mass cluster

μ(mass clusters) $\geq \varepsilon_0$

» Certainty Equivalent Control

For Bi-modal Uniform Distribution



Let B_t be the remaining budget at time t

$$\text{Budget Ratio} = \frac{\text{Remaining Budget}}{\text{Remaining Time}} = \frac{B_t}{T-t}$$

$$\text{CE Quantile Threshold} = 1 - \frac{B_t}{T-t} \triangleq p_t^{ce}$$

$$\text{Decision: hire} \iff \theta_t \geq F^{-1}(p_t^{ce})$$

» Failure of Certainty Equivalent Control

Insufficiency of Certainty Equivalent Control

Assume that $F = \text{Unif}([0, \frac{1}{4}] \cup [\frac{3}{4}, 1])$, for $B = T/2$, we have

$$\mathbb{E}\left[\text{Regret}(CE)\right] = \Omega\left(\sqrt{T}\right)$$

» Failure of Certainty Equivalent Control

Regret Lower Bound

Insufficiency of Certainty Equivalent Control

Assume that $F = \text{Unif}([0, \frac{1}{4}] \cup [\frac{3}{4}, 1])$, for $B = T/2$, we have

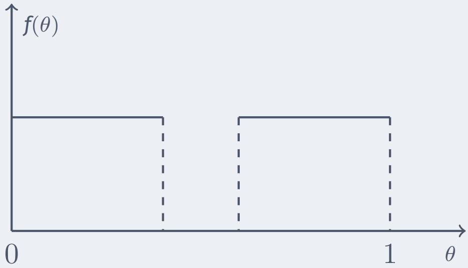
$$\mathbb{E}\left[\text{Regret}(CE)\right] = \Omega\left(\sqrt{T}\right)$$

Remark

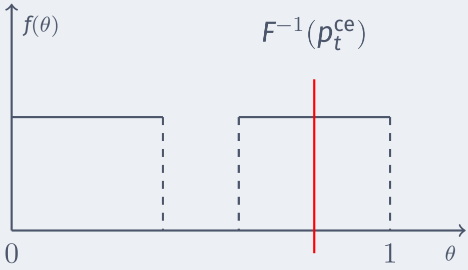
* Same scaling is achievable under a static threshold policy.

» Conservatism with respect to gaps

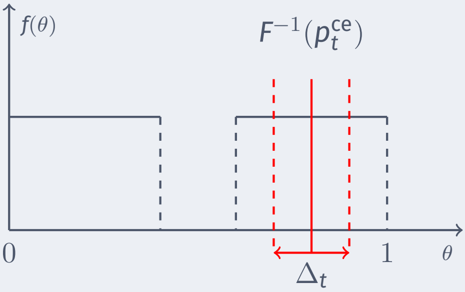
Algorithmic Idea



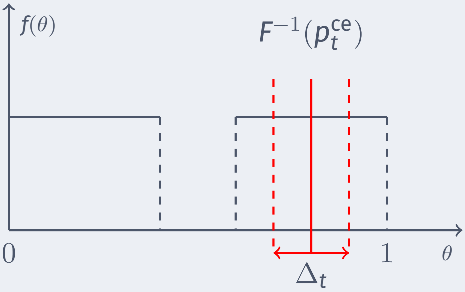
» Conservatism with respect to gaps



» Conservatism with respect to gaps

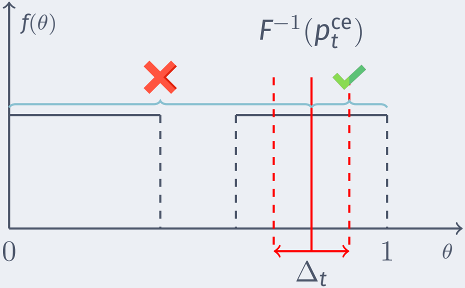


» Conservatism with respect to gaps



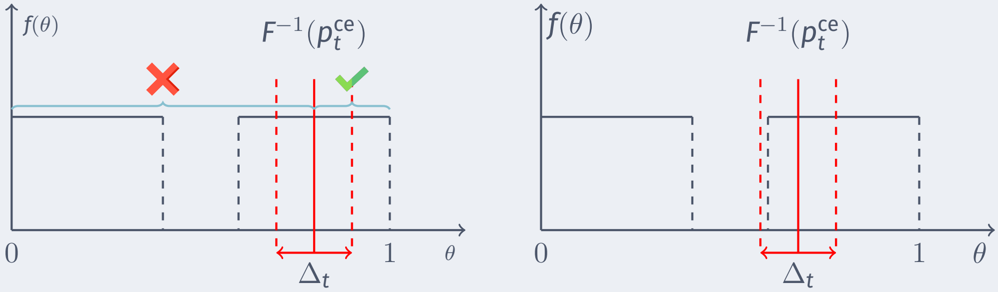
If far from a gap, use the CE threshold

» Conservatism with respect to gaps



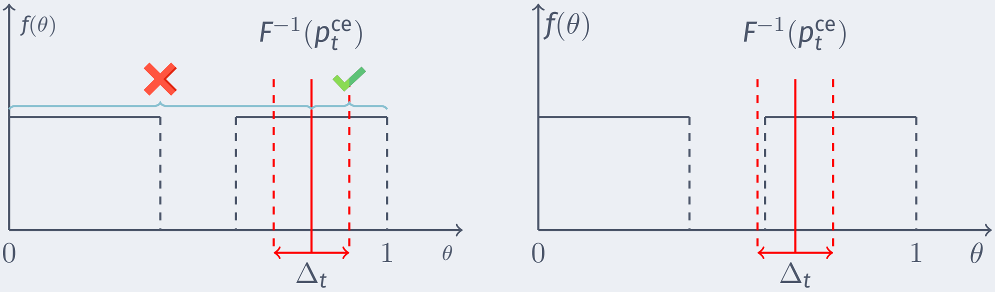
If far from a gap, use the CE threshold

» Conservatism with respect to gaps



If far from a gap, use the CE threshold

» Conservatism with respect to gaps

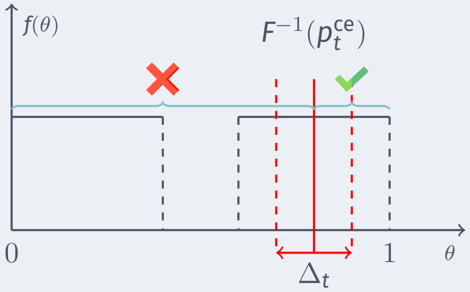


If far from a gap, use the CE threshold

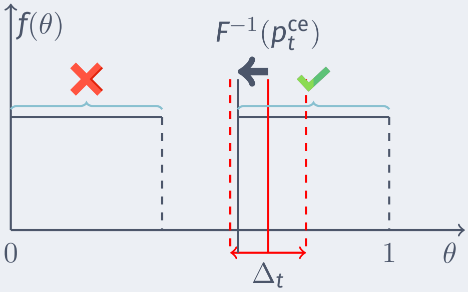
If close to gap, use the gap as threshold

» Conservatism with respect to gaps

Algorithmic Idea



If far from a gap, use the CE threshold



If close to gap, use the gap as threshold

» Conservatism with respect to gaps

Punchline

Regret of CwG Policy

If F is a (β, ε_0)-clustered distribution, then

$$\mathbb{E}\left[\text{Regret}(CwG)\right] = \begin{cases} \mathcal{O}\left((\log T)^2\right), & \beta = 0, \\ \mathcal{O}\left(\text{poly}(\log T) T^{\frac{1}{2} - \frac{1}{2(1+\beta)}}\right), & \beta > 0 \end{cases}$$

If F is a discrete distribution, $\mathbb{E}\left[\text{Regret}(CwG)\right] = \mathcal{O}(1/\varepsilon_0)$

» Conservatism with respect to gaps

Punchline

Regret of CwG Policy

If F is a (β, ε_0)-clustered distribution, then

$$\mathbb{E}\left[\text{Regret}(CwG)\right] = \begin{cases} \mathcal{O}\left((\log T)^2\right), & \beta = 0, \\ \mathcal{O}\left(\text{poly}(\log T) T^{\frac{1}{2} - \frac{1}{2(1+\beta)}}\right), & \beta > 0 \end{cases}$$

If F is a discrete distribution, $\mathbb{E}\left[\text{Regret}(CwG)\right] = \mathcal{O}(1/\varepsilon_0)$

Remark

* $F = \text{Unif}([0, \frac{1}{4}] \cup [\frac{3}{4}, 1])$, CwG ($\mathcal{O}((\log T)^2)$) outperforms CE ($\Omega(\sqrt{T})$).
* Matches the universal lower bound upto polylog factors $\implies$ CwG is near-optimal.

» Conclusion

* *Motivation:* Discrete dist. and continuous dist. with bounded density are well studied but require different algorithmic ideas.

» Conclusion

- *Motivation:* Discrete dist. and continuous dist. with bounded density are well studied but require different algorithmic ideas.
- *Research Ques:* How to unify and interpolate between them?

» Conclusion

- *Motivation:* Discrete dist. and continuous dist. with bounded density are well studied but require different algorithmic ideas.
- *Research Ques:* How to unify and interpolate between them?
- *Challenges:* Unified approach requires dealing with two potential drivers:
  - Rarity of types around points of zero density (β): fundamental driver
  - Presence of gaps: avoidable losses but need smarter algorithms

» Conclusion

- *Motivation:* Discrete dist. and continuous dist. with bounded density are well studied but require different algorithmic ideas.
- *Research Ques:* How to unify and interpolate between them?
- *Challenges:* Unified approach requires dealing with two potential drivers:
  - Rarity of types around points of zero density (β): fundamental driver
  - Presence of gaps: avoidable losses but need smarter algorithms
- *Algorithmic Innovation:* Conservatism with respect to gaps can deal with a broader class of distributions.
  - Recover the guarantees for discrete and bounded continuous distributions.

» Open Problems

* The multi-secretary problem is NRM for $d = 1$, how can we generalize these ideas to higher dimensions?

» Open Problems

* The multi-secretary problem is NRM for $d = 1$, how can we generalize these ideas to higher dimensions?
* How to think of gaps in higher dimensions and how to deal with them?

» Open Problems

* The multi-secretary problem is NRM for $d = 1$, how can we generalize these ideas to higher dimensions?
* How to think of gaps in higher dimensions and how to deal with them?
* Extensions of NRM problems to instances with gaps.

» Open Problems

- The multi-secretary problem is NRM for $d = 1$, how can we generalize these ideas to higher dimensions?
- How to think of gaps in higher dimensions and how to deal with them?
- Extensions of NRM problems to instances with gaps.
  - Application: Order fulfillment problem with more than 2 warehouses.
  - Gaps in spatial demand distributions in practical settings.

Thank you!

» References

Arlotto, A. and Gurvich, I. (2019).
Uniformly bounded regret in the multisecretary problem.
Stochastic Systems, 9(3):231–260.

Bray, R. L. (2022).
The multisecretary problem with continuous valuations.
arXiv preprint arXiv:1912.08917.

Kleywegt, A. J. and Papastavrou, J. D. (1998).
The dynamic and stochastic knapsack problem.
Operations research, 46(1):17–35.

Lei, Y., Jasin, S., and Sinha, A. (2018).
Joint dynamic pricing and order fulfillment for e-commerce retailers.
Manufacturing & Service Operations Management, 20(2):269–284.

Lueker, G. S. (1998).
Average-case analysis of off-line and on-line knapsack problems.