Opening Pandora's Box: the Correlated Case

Evangelia Gergatsouli
UW - Madison

Based on joint work with: Shuchi Chawla, Yifeng Teng, Christos Tzamos, Ruimin Zhang

A Search Problem

Find the best alternative!

- Information is not free
- Explore alternatives (open boxes)
- Stop anytime and take best so far

} **Strategy**

Opening cost: 6
Final option: box 4
Total cost: 12+6

Goal:
Find minimum cost strategy

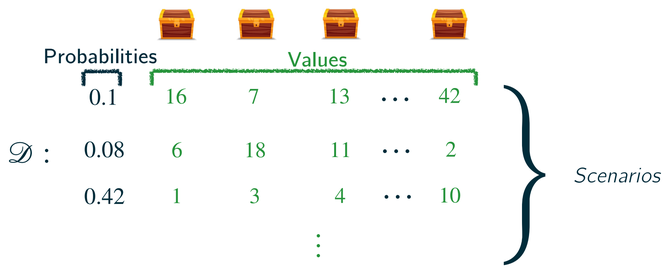
Weitzman's algorithm gives the optimal* [Weitz 1979]

Weitzman's Algorithm:

- Calculate a *reservation value* σ_i for every box i
- Search boxes in order of increasing index until:
  - Current min price seen smaller than index of next box

σ_i: solution to the equation $\mathbb{E}_{\mathscr{D}_i}[(\sigma_i - v)^+] = c_i$

*Optimal for independent distributions! What about correlation?



*Simplifying Assumptions:
1. Explicitly given $\mathscr{D}$. Results hold within $(1+\varepsilon)$ using poly$(n, 1/\varepsilon)$ samples
2. Unit-cost boxes
3. Uniform scenarios

Correlation

Explicitly given $\mathscr{D}$

Can we approximate the optimal? → (probably) Not within constant in poly-time.

Why? Pandora's Box equivalent to Uniform Decision Tree
[Chawla, Gergatsouli, McMahan, Tzamos arXiv 2022]

- $\tilde{O}\left(\log m\right)$ in poly time
- $\tilde{O}(1/\alpha)$ in time $n^{\tilde{O}(m^\alpha)}$, $\alpha < 1$
- Cannot do better than 4 unless P=NP

Sample access to $\mathscr{D}$

Can we approximate the optimal? → No: arbitrarily encode location of box

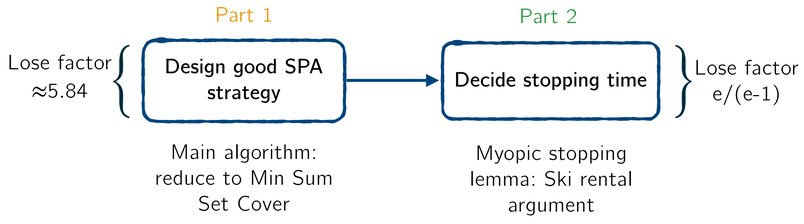
Cannot learn arbitrary mapping with **finitely** many queries!

Optimal strategies are **Fully Adaptive**
→ **adapt** order
→ **adapt** stopping time

Consider **Partially Adaptive** strategies
→ **fix** order
→ **adapt** stopping time

Initial Approach

Space of PA strategies can be large! → Scenario aware PA (**SPA**)
SPA: Fix order → scenario is revealed → decide stopping time



Final Result:
We can approximate the optimal PA within **9.22**

Problem?
Algorithm too **complicated**!

1. Solve LP
2. Remove variables & rescale
3. Round LP

} Part 1

4. Do ski rental

} Part 2

New Approach

Our Algorithm:

- Calculate a *reservation value* σ_i for every box i for every round
- Open box b with $\min \sigma_i$, stop if $v_b \leq \sigma_b$
- Box is now free $c_b = 0$ → Can re-open box for free
- Update Prior with $v_b > \sigma_b$ → Some scenarios not possible anymore

Claim: $\mathbb{E}[(\sigma - v)^+] = c$ is **equivalent** to $\sigma = \min_{A \in S} \frac{c_b |S| + \sum v_b^s}{|A|}$

Note: set A is scenarios with value < σ

Restating the Algorithm

Algorithm:

- R_0 = all scenarios
- While $R_t \neq \emptyset$ do
  - Calculate σ_b for each box
  - Open box b_t with $\min_{i \in \mathscr{B}} \sigma_i$, stop if $v_{b_t} \leq \sigma_{b_t}$
- Box is now free $c_{b_t} = 0$
- Update prior $R_t \leftarrow R_t \backslash A$

Final Result:
We can approximate the optimal PA within **4.42**

Summary of Results