Top Two Algorithms Revisited

Marc Jourdan, Rémy Degenne, Dorian Baudry,
Rianne de Heide and Émilie Kaufmann

October 17, 2022

Motivation

Goal: Identify the item having the highest averaged return.

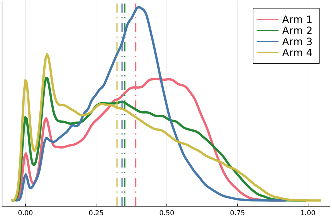
Typical assumptions: Parametric (Bernoulli, Gaussian).

⚠ Too restrictive !

☞ This paper: Bounded distributions.

Crop-management task:

- item = planting date
- observation = yield



Best-arm identification (BAI)

K arms, $F_i \in \mathcal{F}$ bounded distribution of arm $i \in [K]$ with mean μ_i.

Goal: identify $i^\star = \arg\max_i \mu_i$ with confidence $1 - \delta$.

Algorithm: at time n,
• Sequential test: if the stopping time τ_δ is reached, then return the candidate answer $\hat{\imath}_n$.
• **Sampling rule**: pull arm I_n and observe $X_n \sim F_{I_n}$.

Objective: Minimize $\mathbb{E}_{\boldsymbol{F}}[\tau_\delta]$ for δ-correct algorithms

$$\mathbb{P}_{\boldsymbol{F}}\left[\tau_\delta < +\infty, \, \hat{\imath}_{\tau_\delta} \neq i^\star\right] \leq \delta \, .$$

(Garivier and Kaufmann, 2016) For all δ-correct algorithm,

$$\forall \boldsymbol{F} \in \mathcal{F}^K, \quad \liminf_{\delta \to 0} \frac{\mathbb{E}_{\boldsymbol{F}}[\tau_\delta]}{\log(1/\delta)} \geq T^\star(\boldsymbol{F}) \, .$$

δ-correct sequential test

How to obtain a δ-correct sequential test ?

- ☞ recommend the empirical best arm, $\hat{\imath}_n = \arg\max_{i \in [K]} \mu_{n,i}$.
- ☞ calibrated GLR stopping rule

$$\tau_\delta = \inf \left\{ n \in \mathbb{N} \mid \min_{j \neq \hat{\imath}_n} W_n(\hat{\imath}_n, j) > c(n, \delta) \right\} ,$$

where $c(n, \delta)$ is a calibrated threshold and $W_n(i, j)$ is the empirical transportation cost between arms (i, j).

Top Two sampling rule

Family of algorithms:

☞ β proportion of samples to the best arm (Russo, 2016).

1: Choose a **leader** $B_n \in [K]$
2: Choose a **challenger** $C_n \in [K] \setminus \{B_n\}$
3: Sample B_n with probability β, else sample C_n

Theorem

Instantiating the Top Two algorithm with any pair of leader/challenger satisfying some properties yields a δ-correct algorithm which is asymptotically β-optimal for instances having distinct means.

Leader and challenger

How to choose the leader ?

- Thompson Sampling (Russo, 2016), $\arg\max_{i \in [K]} \theta_i$ with $\theta \sim \Pi_{n-1}$ where Π_{n-1} is a sampler on $(0, B)^K$.
- **Empirical Best**, $\hat{\imath}_{n-1}$.

How to choose the challenger ?

- Re-Sampling (Russo, 2016), repeat $\theta \sim \Pi_{n-1}$ until $B_n \notin \arg\max_{i \in [K]} \theta_i$, then $\arg\max_{i \in [K]} \theta_i$.
- Transportation Cost (Shang et al., 2020), $\arg\min_{j \neq B_n} W_{n-1}(B_n, j)$.
- **Transportation Cost Improved**,

$$\arg\min_{j \neq B_n} W_{n-1}(B_n, j) + \log(N_{n-1,j}) \, .$$

Empirical results

Crop-management task: arm = planting date / observation = yield

Moderate regime, $\delta = 0.01$. Top Two algorithms with $\beta = 1/2$.

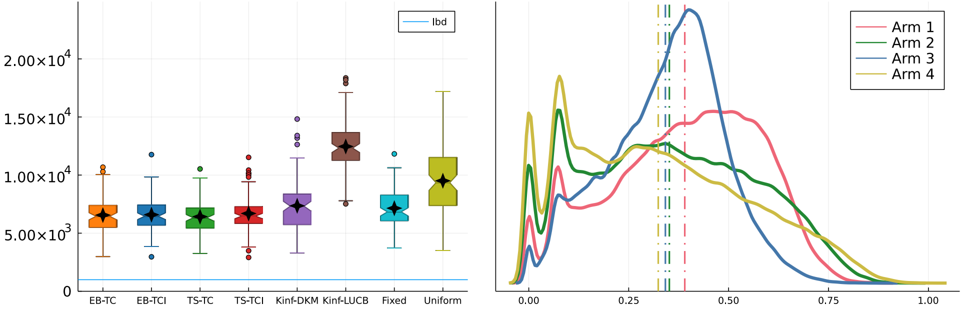


Figure: Empirical stopping time (a) on scaled DSSAT instances with their density and mean (b). Lower bound is $T^\star(\boldsymbol{F}) \log(1/\delta)$.

Conclusion

Paper & Poster

1. Generic and modular analysis of Top Two algorithms.
2. First asymptotically β-optimal instances for bounded distributions.
3. Competitive performance on a real-world non-parametric task.