Top Two Algorithms Revisited

Marc Jourdan[1], Rémy Degenne[1], Dorian Baudry[1], Rianne de Heide[2] and Émilie Kaufmann[1]

[1] Univ. Lille, CNRS, Inria, Centrale Lille, UMR 9198-CRIStAL, F-59000 Lille, France

[2] Vrije Universiteit Amsterdam

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

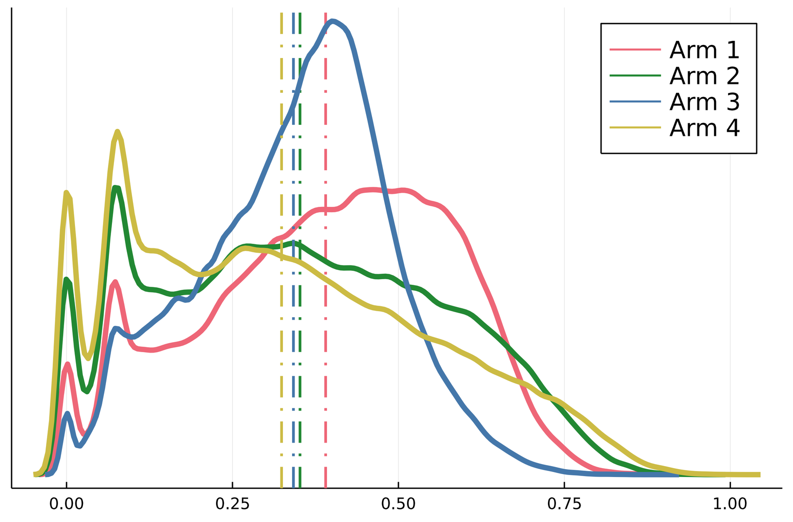

Figure 1: DSSAT instances' density ($K = 4$).

Best-arm identification (BAI)

K arms: $F_i \in \mathcal{F}$ cdf of arm $i \in [K]$ with mean $m(F_i) = \mathbb{E}_{X\sim F_i}[X]$.

Set of distributions $\mathcal{F}$ with set of means $\mathcal{I}$:

(a) **Bounded distributions** in $[0, B]$ and $\mathcal{I} = (0, B)$,

(b) Single parameter exponential families (SPEF) with sub-exponential tails.

Goal: identify $i^\star(\boldsymbol{F}) = \arg\max_{i\in[K]} m(F_i)$ with confidence $1 - \delta \in (0, 1)$.

Algorithm: at time n,

- *Sequential test:* if the stopping time τ_δ is reached, then return the candidate answer $\hat{\imath}_n$, else
- **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, meaning that

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

Sample complexity lower bound

Garivier and Kaufmann (2016), Agrawal et al. (2020): For all δ-correct algorithm,

$$\forall \boldsymbol{F} \in \mathcal{F}^K, \quad \mathbb{E}_{\boldsymbol{F}}[\tau_\delta] \geq T^\star(\boldsymbol{F}) \ln(1/(2.4\delta)) \, .$$

Family of β-algorithms: $\beta \in (0, 1)$ proportion of pulls to the best arm (Russo, 2016).

☞ Example: **Top Two sampling rule.**

The inverse of the β-characteristic time is

$$T^\star_\beta(\boldsymbol{F})^{-1} = \sup_{w\in\triangle_K, w_{i^\star(\boldsymbol{F})}=\beta} \min_{i\neq i^\star(\boldsymbol{F})} \inf_{u\in\mathcal{I}} \left\{\beta\mathcal{K}^-_{\inf}(F_{i^\star(\boldsymbol{F})}, u) + w_i\mathcal{K}^+_{\inf}(F_i, u)\right\} \, .$$

$\triangle_K$ simplex, $\mathcal{K}^{\pm}_{\inf}(F, u) = \inf\left\{\mathrm{KL}(F, G) \mid G \in \mathcal{F},\ m(G) \gtrless u\right\}$ for all $(F, u) \in \mathcal{F}\times\mathcal{I}$.

Properties:

- $T^\star(\boldsymbol{F}) = \min_{\beta\in(0,1)} T^\star_\beta(\boldsymbol{F})$ and $T^\star_{1/2}(\boldsymbol{F}) \leq 2T^\star(\boldsymbol{F})$.
- $T^\star_\beta(\boldsymbol{F})$ is achieved for a unique β-optimal allocation w^β when $i^\star(\boldsymbol{F})$ is unique.

δ-correct sequential test

? How to obtain a δ-correct sequential test ?

☞ recommend the empirical best arm $\hat{\imath}_n = i^\star(\boldsymbol{F}_n)$, where $N_{n,i} = \sum_{t\in[n]} \mathbb{1}(I_t = i)$ and $F_{n,i} = \frac{1}{N_{n,i}} \sum_{t\in[n]} \delta_{X_t}\mathbb{1}(I_t = i)$.

☞ **GLR stopping rule**: given a calibrated threshold $c(n, \delta)$, define

$$\tau_\delta = \inf\{n \in \mathbb{N} \mid \min_{j\neq\hat{\imath}_n} W_n(\hat{\imath}_n, j) > c(n, \delta)\} \, , \qquad (1)$$

where the empirical transportation cost between arms (i, j) is

$$W_n(i, j) = \inf_{u\in\mathcal{I}} \left\{N_{n,i}\mathcal{K}^-_{\inf}(F_{n,i}, u) + N_{n,j}\mathcal{K}^+_{\inf}(F_{n,j}, u)\right\} \, .$$

Top Two sampling rule

Choose a **leader** $B_n \in [K]$

Choose a **challenger** $C_n \neq B_n$

Sample B_n with probability β, else sample C_n

? How to choose the leader ?

☞ *Thompson Sampling* (TS) (Russo, 2016), $\arg\max_{i\in[K]} \theta_i$ with $\theta \sim \Pi_{n-1}$ where Π_{n-1} is a sampler on $\mathcal{I}^K$.

☞ **Empirical Best** (EB), $\hat{\imath}_{n-1}$.

? How to choose the challenger ?

☞ *Re-Sampling* (RS) (Russo, 2016), $\arg\max_{i\in[K]} \theta_i$ where we sample $\theta \sim \Pi_{n-1}$ until $B_n \notin \arg\max_{i\in[K]} \theta_i$.

☞ *Transportation Cost* (TC) (Shang et al., 2020) = [SHKMV20], $\arg\min_{j\neq B_n} W_{n-1}(B_n, j)$.

☞ **Transportation Cost Improved** (TCI),

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

Bounded instances

Calibrated threshold: $c(n, \delta) = \ln(1/\delta) + 2\ln(1 + n/2) + 2 + \ln(K - 1)$.

Computing empirical $\mathcal{K}_{\inf}$: let $(X_{t,i})_{t\in[N_{n,i}]}$ be the samples of arm i, then

$$N_{n,i}\mathcal{K}^+_{\inf}(F_{n,i}, u) = \sup_{\lambda\in[0,1]} \sum_{t\in[N_{n,i}]} \ln\left(1 - \lambda\frac{X_{t,i} - u}{B - u}\right) \, ,$$

which is computed with a zero-order optimization algorithm (e.g. Brent's method).

Computing $W_n(i, j)$: minimizing a univariate function on a bounded interval.

? How to design a sampler over $(0, B)^K$? Riou and Honda (2020)

☞ **Dirichlet sampler:** $\Pi_n = \bigtimes_{i\in[K]} \Pi_{n,i}$ where $\Pi_{n,i}$ uses the empirical cdf $F_{n,i}$ augmented by $\{0, B\}$. The sampler $\Pi_{n,i}$ returns

$$\sum_{t\in[N_{n,i}]} w_t X_{t,i} + Bw_{N_{n,i}+1} \quad \text{with} \quad \boldsymbol{w} \sim \mathrm{Dir}(\mathbf{1}_{N_{n,i}+2}) \, .$$

Sample complexity upper bound

Theorem 1. *Given* (1) *with a calibrated threshold, instantiating the Top Two sampling rule with any pair of leader/challenger satisfying some properties yields a δ-correct algorithm, and for instances $\boldsymbol{F} \in \mathcal{F}^K$ having distinct means it satisfies*

$$\limsup_{\delta\to 0} \frac{\mathbb{E}_{\boldsymbol{F}}[\tau_\delta]}{\log(1/\delta)} \leq T^\star_\beta(\boldsymbol{F}) \, .$$

Table 1: Leaders and challengers satisfying the sufficient properties for **Theorem 1** to hold.

Distributions		TS	EB	RS	TC	TCI
SPEF	Gaussian	[SHKMV20]	✓	[SHKMV20]	[SHKMV20]	✓
	Bernoulli	✓	✓	✓	✓	✓
	sub-Exp	?	✓	?	✓	✓
Bounded		✓	✓	✓	✓	✓

Proof. Convergence time $T^\epsilon_\beta = \inf\{T \mid \forall n \geq T,\ \|N_n/n - w^\beta\|_\infty \leq \epsilon\}$. Under (1),

$$\ln(1/\delta) \approx_{\delta\to 0} c(n, \delta) \geq \min_{j\neq\hat{\imath}_n} W_n(\hat{\imath}_n, j) \approx_{n\geq T^\epsilon_\beta} nT^\star_\beta(\boldsymbol{F})^{-1} \, .$$

Sufficient exploration: $\min_{i\in[K]} N_{n,i} \geq \sqrt{n/K}$ for n large enough.

Let $\psi_{n,i} = \mathbb{P}_{|(n-1)}[I_n = i]$ and $\Psi_{n,i} = \sum_{t\in[n]} \psi_{t,i}$. Then, $(N_{n,i} - \Psi_{n,i})/\sqrt{n}$ are sub-Gaussian random variables and the Top Two sampling rule satisfies

$$\psi_{n,i} = \beta\mathbb{P}_{|(n-1)}[B_n = i] + (1 - \beta)\sum_{j\neq i} \mathbb{P}_{|(n-1)}[B_n = j]\mathbb{P}_{|(n-1)}[C_n = i|B_n = j] \, .$$

Convergence towards w^β**:** showing $\mathbb{E}_{\boldsymbol{F}}[T^\varepsilon_\beta] < +\infty$ for ε small enough

- Leader, $\mathbb{P}_{|n}[B_{n+1} \neq i^\star] = \mathcal{O}(n^{-\alpha})$ for n large enough, with $\alpha > 0$.
- Challenger, for n large enough and all $i \neq i^\star$, $\Psi_{n,i}/n \geq w^\beta_i + \epsilon$ implies that $\mathbb{P}_{|n}[C_{n+1} = i|B_{n+1} = i^\star] = \mathcal{O}(n^{-\alpha})$. ■

Experiments

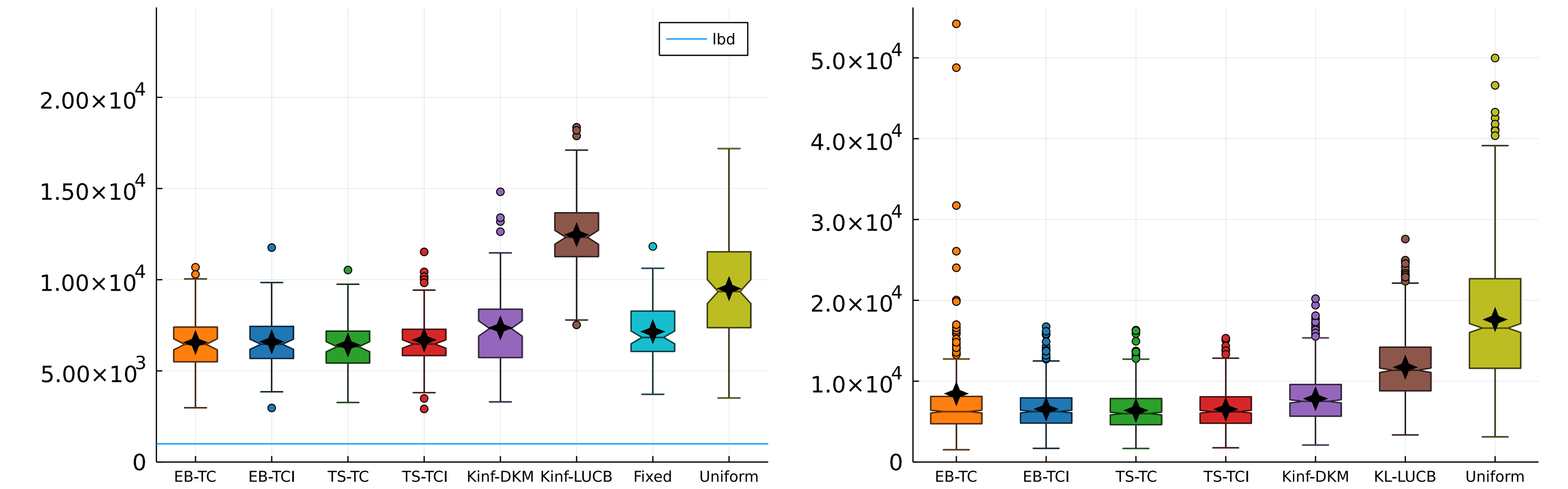

Figure 2: Empirical stopping time for $\delta = 0.01$ on (a) DSSAT instances and (b) random and distinct Bernoulli instances ($K = 10$) with $\mu_1 = 0.6$ and $\mu_i \sim \mathcal{U}([0.2, 0.5])$ for $i \neq 1$. Lower bound is $T^\star(\boldsymbol{F})\ln(1/\delta)$. Top Two algorithms with $\beta = 1/2$.

Conclusion

1. Generic and modular analysis of Top Two algorithms.
2. Proving asymptotic β-optimality of Top Two algorithms as in Table 1.
3. Competitive performance on a real-world non-parametric task.