Characterizing and Approximating the Optimal Allocation for Top-m Arm Identification

Master Thesis

Kevin Klein, Johannes Kirschner, Mojmír Mutný, Prof. Dr. Andreas Krause

Context

- Bandits, arms $[k]$.

Context

- Bandits, arms $[k]$.
- Reward distributions: 1-dimensional exponential family.

Context

- Bandits, arms $[k]$.
- Reward distributions: 1-dimensional exponential family.
- Frequentist setting: true means θ^* and m true best arms $S^* \subset [k]$.

Context

- Bandits, arms $[k]$.
- Reward distributions: 1-dimensional exponential family.
- Frequentist setting: true means θ^* and m true best arms $S^* \subset [k]$.
- Top-m arm identification goal: identifying S^* with high confidence and few samples.

Context

- Bandits, arms $[k]$.
- Reward distributions: 1-dimensional exponential family.
- Frequentist setting: true means θ^* and m true best arms $S^* \subset [k]$.
- Top-m arm identification goal: identifying S^* with high confidence and few samples.
- Confidence: $\max_S \Pr[S \text{ is top-}m | \mathcal{D}_n]$ with $S \subset [k]$ and $|S| = m$.

Context

- Bandits, arms $[k]$.
- Reward distributions: 1-dimensional exponential family.
- Frequentist setting: true means θ^* and m true best arms $S^* \subset [k]$.
- Top-m arm identification goal: identifying S^* with high confidence and few samples.
- Confidence: $\max_S \Pr[S \text{ is top-}m | \mathcal{D}_n]$ with $S \subset [k]$ and $|S| = m$.
- Fixed vs adaptive.

Context

- Bandits, arms $[k]$.
- Reward distributions: 1-dimensional exponential family.
- Frequentist setting: true means θ^* and m true best arms $S^* \subset [k]$.
- Top-m arm identification goal: identifying S^* with high confidence and few samples.
- Confidence: $\max_S \Pr[S \text{ is top-}m | \mathcal{D}_n]$ with $S \subset [k]$ and $|S| = m$.
- Fixed vs adaptive.
- Allocation ψ with $\psi_{n,l} = \Pr[I_n = l]$.

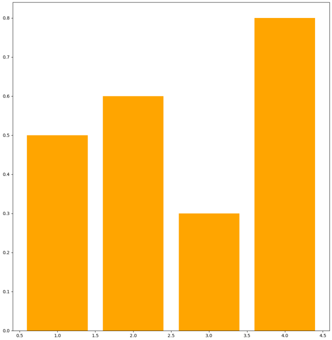

Figure: Arms and their true means θ^*. $k = 4$.

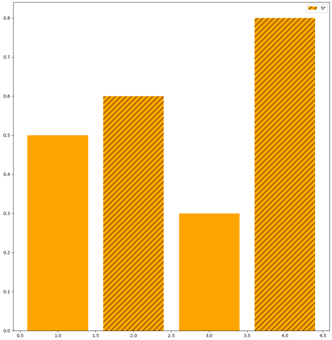

Figure: Arms, their true means θ^* and S^*. $k = 4$, $m = 2$.

Credit

Most of the work is inspired by Daniel Russo's *Simple Bayesian algorithm for Best Arm Identification* [1].

[1] `https://arxiv.org/abs/1602.08448`

Optimality of fixed allocation

- $\Pr[S^* \text{ is top-}m | \mathcal{D}_n] \to 1$ with $n \to \infty$ is not hard.

Optimality of fixed allocation

- $\Pr[S^* \text{ is top-}m|\mathcal{D}_n] \to 1$ with $n \to \infty$ is not hard.
- What's hard: converge optimally fast.

Optimality of fixed allocation

- $\Pr[S^*$ is top-$m | \mathcal{D}_n] \to 1$ with $n \to \infty$ is not hard.
- What's hard: converge optimally fast.
- Goal: Maximize rate at which $\Pr[S^*$ is top-$m | \mathcal{D}_n] \to 1$ with $n \to \infty$.

Optimality of fixed allocation

- $\Pr[S^*$ is top-$m|\mathcal{D}_n] \to 1$ with $n \to \infty$ is not hard.
- What's hard: converge optimally fast.
- Goal: Maximize rate at which $\Pr[S^*$ is top-$m|\mathcal{D}_n] \to 1$ with $n \to \infty$.
- We prove that for any fixed allocation ψ:

$$\Pr[S^* \text{ is top-}m|\mathcal{D}_n] \tag{1}$$

$$\approx 1 - \exp\{-n \min_{i \notin S^*} \min_{j \in S^*} \min_{x \in \mathbb{R}} \psi_i d_{KL}(\theta_i^* || x) + \psi_j d_{KL}(\theta_j^* || x)\} \tag{2}$$

Optimality of fixed allocation

- $\Pr[S^* \text{ is top-}m | \mathcal{D}_n] \to 1$ with $n \to \infty$ is not hard.
- What's hard: converge optimally fast.
- Goal: Maximize rate at which $\Pr[S^* \text{ is top-}m | \mathcal{D}_n] \to 1$ with $n \to \infty$.
- We prove that for any fixed allocation ψ:

$$\Pr[S^* \text{ is top-}m | \mathcal{D}_n] \tag{1}$$

$$\approx 1 - \exp\{-n \min_{i \notin S^*} \min_{j \in S^*} \min_{x \in \mathbb{R}} \psi_i d_{KL}(\theta_i^* || x) + \psi_j d_{KL}(\theta_j^* || x)\} \tag{2}$$

- As a consequence, the optimal minimizes the exponent.

$$\psi^* = \arg\max_{\psi} \min_{i \notin S^*} \min_{j \in S^*} \min_{x \in \mathbb{R}} \psi_i d_{KL}(\theta_i^* || x) + \psi_j d_{KL}(\theta_j^* || x)$$

Optimality of fixed allocation

- $\Pr[S^* \text{ is top-}m|\mathcal{D}_n] \to 1$ with $n \to \infty$ is not hard.
- What's hard: converge optimally fast.
- Goal: Maximize rate at which $\Pr[S^* \text{ is top-}m|\mathcal{D}_n] \to 1$ with $n \to \infty$.
- We prove that for any fixed allocation ψ:

$$\Pr[S^* \text{ is top-}m|\mathcal{D}_n] \tag{1}$$

$$\approx 1 - \exp\{-n \min_{i \notin S^*} \min_{j \in S^*} \min_{x \in \mathbb{R}} \psi_i d_{KL}(\theta_i^* || x) + \psi_j d_{KL}(\theta_j^* || x)\} \tag{2}$$

- As a consequence, the optimal minimizes the exponent.

$$\psi^* = \arg\max_{\psi} \min_{i \notin S^*} \min_{j \in S^*} \min_{x \in \mathbb{R}} \psi_i d_{KL}(\theta_i^* || x) + \psi_j d_{KL}(\theta_j^* || x)$$

Optimality of fixed allocation

- $\Pr[S^* \text{ is top-}m|\mathcal{D}_n] \to 1$ with $n \to \infty$ is not hard.
- What's hard: converge optimally fast.
- Goal: Maximize rate at which $\Pr[S^* \text{ is top-}m|\mathcal{D}_n] \to 1$ with $n \to \infty$.
- We prove that for any fixed allocation ψ:

$$\Pr[S^* \text{ is top-}m|\mathcal{D}_n] \tag{1}$$

$$\approx 1 - \exp\{-n \min_{i \notin S^*} \min_{j \in S^*} \boxed{\min_{x \in \mathbb{R}} \psi_i d_{KL}(\theta_i^* || x) + \psi_j d_{KL}(\theta_j^* || x)}\} \tag{2}$$

- As a consequence, the optimal minimizes the exponent.

$$\psi^* = \arg\max_{\psi} \min_{i \notin S^*} \min_{j \in S^*} \min_{x \in \mathbb{R}} \psi_i d_{KL}(\theta_i^* || x) + \psi_j d_{KL}(\theta_j^* || x)$$

Evidence for distinction

- $C_{j,i}(\psi_j, \psi_i) = \min_{x \in \mathbb{R}} \psi_i d_{KL}(\theta_i^* || x) + \psi_j d_{KL}(\theta_j^* || x)$

Evidence for distinction

- $C_{j,i}(\psi_j, \psi_i) = \min_{x \in \mathbb{R}} \psi_i d_{KL}(\theta_i^* || x) + \psi_j d_{KL}(\theta_j^* || x)$
- Distinction comprises both frequency ψ and distance of values, captured by KL divergence.

Characterization

Theorem (Characterization of optimal allocation)

There is a unique fixed allocation ψ^* maximizing the convergence rate of the posterior mass put on the true best arms. It enforces that:

$$\forall j_1, j_2 \in S^*, \forall i_1, i_2 \notin S^* : C_{j,i}(\psi^*_{j_1}, \psi^*_{i_1}) = C_{j,i}(\psi^*_{j_2}, \psi^*_{i_2})$$

Characterization

Theorem (Characterization of optimal allocation)

There is a unique fixed allocation ψ^* maximizing the convergence rate of the posterior mass put on the true best arms. It enforces that:

$$\forall j_1, j_2 \in S^*, \forall i_1, i_2 \notin S^* : C_{j,i}(\psi^*_{j_1}, \psi^*_{i_1}) = C_{j,i}(\psi^*_{j_2}, \psi^*_{i_2})$$

This leads to a slightly simpler convergence rate for the optimal allocation:

$$\Pr[S^* \text{ is top-}m|\mathcal{D}_n] \approx 1 - \exp\{-n\, C_{j,i}(\psi^*_j, \psi^*_i)\}$$

for any $j \in S^*$, $i \notin S^*$.

Example: Setup

- Top-4.

Example: Setup

- Top-4.
- Given true underlying means:

$$\theta^1 = [.1,\ .2,\ .3,\ .4,\ .5,\ \underbrace{.6,\ .7,\ .8,\ .9}_{S^*}] \tag{3}$$

$$\theta^2 = [.4,\ .425,\ .45,\ .475,\ .5,\ \underbrace{.525,\ .55,\ .575,\ .6}_{S^*}] \tag{4}$$

Example: Setup

- Top-4.
- Given true underlying means:

$$\theta^1 = [.1,\ .2,\ .3,\ .4,\ .5,\ \underbrace{.6,\ .7,\ .8,\ .9}_{S^*}] \tag{3}$$

$$\theta^2 = [.4,\ .425,\ .45,\ .475,\ .5,\ \underbrace{.525,\ .55,\ .575,\ .6}_{S^*}] \tag{4}$$

- Overdetermined system of equations: Find ψ such that $\forall j_1, j_2 \in S^*, \forall i_1, i_2 \notin S^* : C_{j,i}(\psi_{j_1}, \psi_{i_1}) = C_{j,i}(\psi_{j_2}, \psi_{i_2})$.

Example: Resulting allocation

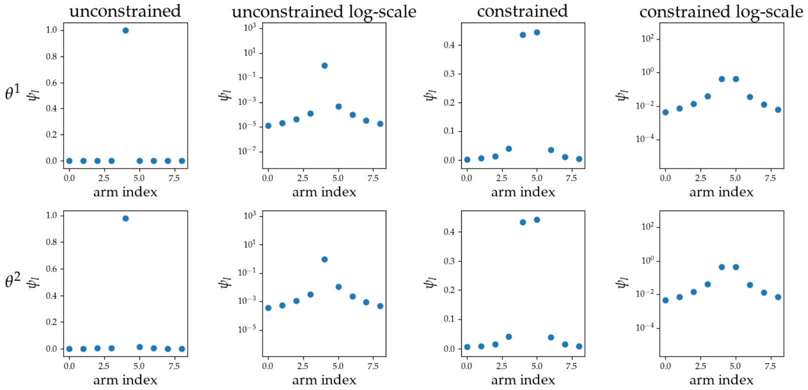

Figure: Unconstrained and constrained optimal allocation for θ^1 and θ^2, top-4.

Top-2m XOR Thompson sampling

- Adaptive.

Top-2m XOR Thompson sampling

- Adaptive.
- Simple.

Top-2m XOR Thompson sampling

- Adaptive.
- Simple.
- Bayesian (remember: confidence).

Top-2m XOR Thompson sampling

- Adaptive.
- Simple.
- Bayesian (remember: confidence).
- Asymptotic consistency.

Top-2m XOR Thompson sampling

- Adaptive.
- Simple.
- Bayesian (remember: confidence).
- Asymptotic consistency.
- Based on Thompson sampling and another layer of randomization.

Top-2m XOR Thompson sampling

- Adaptive.
- Simple.
- Bayesian (remember: confidence).
- Asymptotic consistency.
- Based on Thompson sampling and another layer of randomization.
- Idea: Always sample two distinct candidates.

Top-2m XOR Thompson sampling

- Adaptive.
- Simple.
- Bayesian (remember: confidence).
- Asymptotic consistency.
- Based on Thompson sampling and another layer of randomization.
- Idea: Always sample two distinct candidates.
- Goal: Convergence towards optimal fixed allocation.

TXTS: Algorithm

Algorithm 1 TXTS: Given a posterior Π_{n-1} in step n.

$\hat{\theta} \sim \Pi_{n-1}$
$S_1 = \text{top-}m(\hat{\theta})$
repeat
  $\hat{\theta} \sim \Pi_{n-1}$
  $S_2 = \text{top-}m(\hat{\theta})$
until $S_1 \neq S_2$
$I_n \sim \mathcal{U}(S_1 \oplus S_2)$
Play I_n, observe reward and update posterior.

TXTS: Exemplary candidates

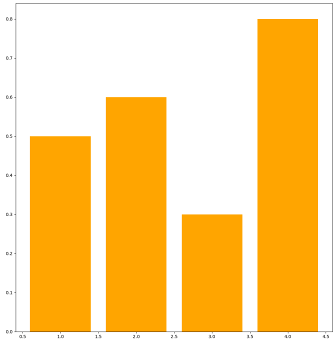

Figure: Arms and their true means θ^*. $k = 4$.

TXTS: Exemplary candidates

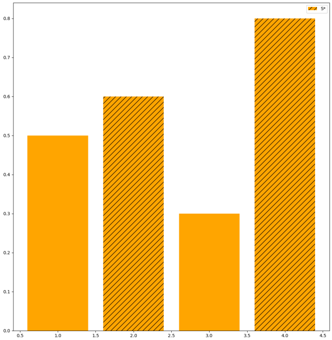

Figure: Arms, their true means θ^* and S_1. $k = 4$, $m = 2$.

TXTS: Exemplary candidates

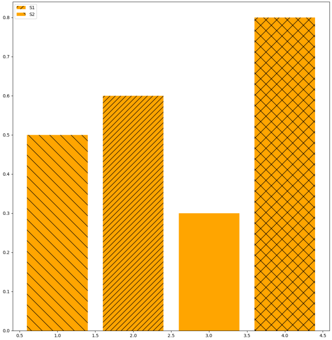

Figure: Arms, their true means θ^*, S_1 and S_2. $k = 4$, $m = 2$.

TXTS: Empirical results

- Bernoulli rewards.

TXTS: Empirical results

- Bernoulli rewards.
- Beta posteriors, $\alpha = \beta = 1$.

TXTS: Empirical results

- Bernoulli rewards.
- Beta posteriors, $\alpha = \beta = 1$.
- Top-4.

TXTS: Empirical results

- Bernoulli rewards.
- Beta posteriors, $\alpha = \beta = 1$.
- Top-4.
- $\theta^* = [.1,\ .2,\ .3,\ .4,\ .5,\ \underbrace{.6,\ .7,\ .8,\ .9}_{S^*}]$

TXTS: Evidence collection

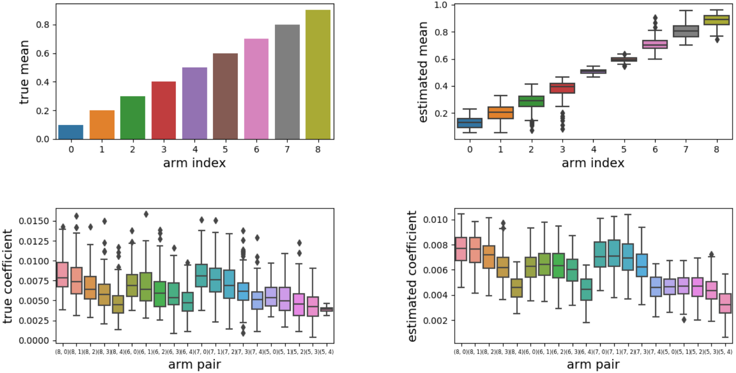

Figure: Top row: true and estimated θ. Bottom row: true and estimated coefficients $C_{j,i}$. 2000 steps, 150 seeds.

TXTS: Empricial average allocation comparison

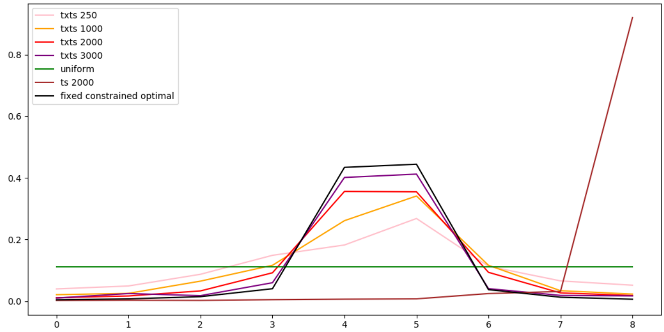

Figure: Comparison of allocations for different methods and numbers of samples. 50 seeds.

TXTS: Theoretical status quo

- Allocation ψ hard to explicitly express in closed form.

TXTS: Theoretical status quo

- Allocation ψ hard to explicitly express in closed form.
- Established bounds on ψ.

TXTS: Theoretical status quo

- Allocation ψ hard to explicitly express in closed form.
- Established bounds on ψ.
- Proven $\Pr[S^*$ is top-$m|\mathcal{D}_n] \to 1$.

TXTS: Theoretical status quo

- Allocation ψ hard to explicitly express in closed form.
- Established bounds on ψ.
- Proven $\Pr[S^*$ is top-$m|\mathcal{D}_n] \to 1$.
- Proven $\sum_{j \in S^*} \psi_{n,S^*} \to \frac{1}{2}$.

TXTS: Theoretical status quo

- Allocation ψ hard to explicitly express in closed form.
- Established bounds on ψ.
- Proven $\Pr[S^*$ is top-$m|\mathcal{D}_n] \to 1$.
- Proven $\sum_{j \in S^*} \psi_{n,S^*} \to \frac{1}{2}$.
- Proven that if converged on S^*, then convergence on S^{*c} and vice versa.

TXTS: Theoretical status quo

- Allocation ψ hard to explicitly express in closed form.
- Established bounds on ψ.
- Proven $\Pr[S^*$ is top-$m|\mathcal{D}_n] \to 1$.
- Proven $\sum_{j\in S^*} \psi_{n,S^*} \to \frac{1}{2}$.
- Proven that if converged on S^*, then convergence on S^{*c} and vice versa.
- Proven some more technical statements.

TXTS: Theoretical outlook

Prove overall convergence, i.e. $\frac{1}{N}\sum_{n \in [N]} \psi_n \rightarrow \psi^*$.

Thanks!

Appendix.

1-d exponential family

- They have a scalar parameter θ.
- They come with the definition of functions $T(x)$, $\nu(\theta)$, $h(x)$ and $A(\theta)$. T corresponds to the sufficient statistic and A to the log-partition function.
- The probability density function is defined by:

$$f_X(x|\theta) = h(x)\exp(\nu(\theta)T(x) - A(\theta))$$

- They have conjugate priors.
- Bernoulli, binomial with known number of trials, Poisson, exponential, Pareto with known minimal value, chi-squared, normal distribution with known variance and more.

Logarithmic equivalence

Convergence is reasoned about in a $\doteq$ sense, with:

$$a_n \doteq b_n \Leftrightarrow \lim_{n\to\infty} \frac{1}{n} \log \frac{a_n}{b_n} \to 0 \tag{5}$$

Computing evidence

The Bernoulli assumption allows us to solve the minimization over x analytically.

$$d(\theta_l||x) = \theta_l \log(\frac{\theta_l}{x}) + (1-\theta_l)\log(\frac{1-\theta_l}{1-x}) \tag{6}$$

$$x_0 = \frac{\psi_j\theta_j + \psi_i\theta_i}{\psi_j + \psi_i} \tag{7}$$

TXTS: Increase in confidence

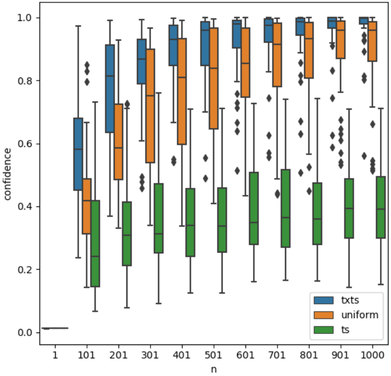

Figure: Confidence per steps for TXTS, uniform and Thompson sampling. 50 seeds.

TXTS: Empirical average allocation

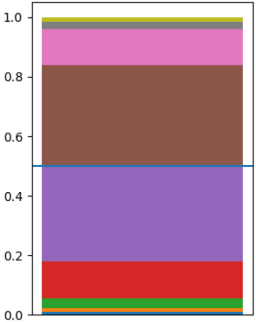

Figure: TXTS empirical average allocation after 1000 steps. Arms ordered by true mean. 50 seeds.

Theorem (Sufficient condition for optimality)

If

$$\forall l \in [k], \delta > 0 : \sum_{n \in \mathbb{N}} \psi_{n,l} \mathbb{I}[\bar{\psi}_{n,l} \geq \psi_l^* + \delta] < \infty \quad (8)$$

then $\bar{\psi}_n \to \psi^*$.