Bandit Social Learning under Myopic Behavior

Kiarash Banihashem, MohammadTaghi Hajiaghayi, Suho Shin (all: UMD), Aleksandrs Slivkins (MSR-NYC)

Takeaways

Exploration needed in bandits? **YES!**
Does Greedy algorithm fail? **Always (pretty much)**
First strong/general result on failure of Greedy!

Social Learning perspective: each round controlled by myopic agent

Failures beyond Greedy: for wide range of "myopic" behaviors that are consistent with constant-prob confidence intervals

Model: Bandit Social Learning (BSL)

T rounds, K actions ("arms")
$\text{Reward}(\text{arm } a) \sim \text{Bernoulli}(\mu_a),\ \mu_a \in (c, 1-c)$ for small constant c

In each round t:

- new agent arrives, observes *shared history* of (arm, reward) for all previous agents and N_0 "initial samples" per arm
- chooses an arm, collects reward

Focus: ***η*-confident agents.** ∀ arm a, maps history to "index"

$$\text{Ind}_{a,t} \in (\hat{\mu}_{a,t} - \sqrt{\eta/n_{a,t}}, \hat{\mu}_{a,t} + \sqrt{\eta/n_{a,t}})$$

consistent with const-prob confidence bounds.

Here, $\hat{\mu}_{a,t}$: empirical average, $n_{a,t}$: #draws

Special cases

Greedy: $\text{Ind}_{a,t} = \hat{\mu}_{a,t}$
Optimism: $\text{Ind}_{a,t}$ = Upper CB
Pessimism: $\text{Ind}_{a,t}$ = Lower CB

Other behaviors/biases
Correlation (across arms or time)

Main results (clean: 2 arms)

Learning failure: optimal arm never chosen

Thm: FailureProb > $p_\eta := \Omega\left(\sqrt{\frac{1+\eta}{N_0}}\right) e^{-O(\eta)}$ if all agents are η-confident

Cor: Linear regret: $\text{Regret}(T) \geq T\Delta\, p_\eta$, where $\Delta := |\mu_1 - \mu_2|$ is the gap

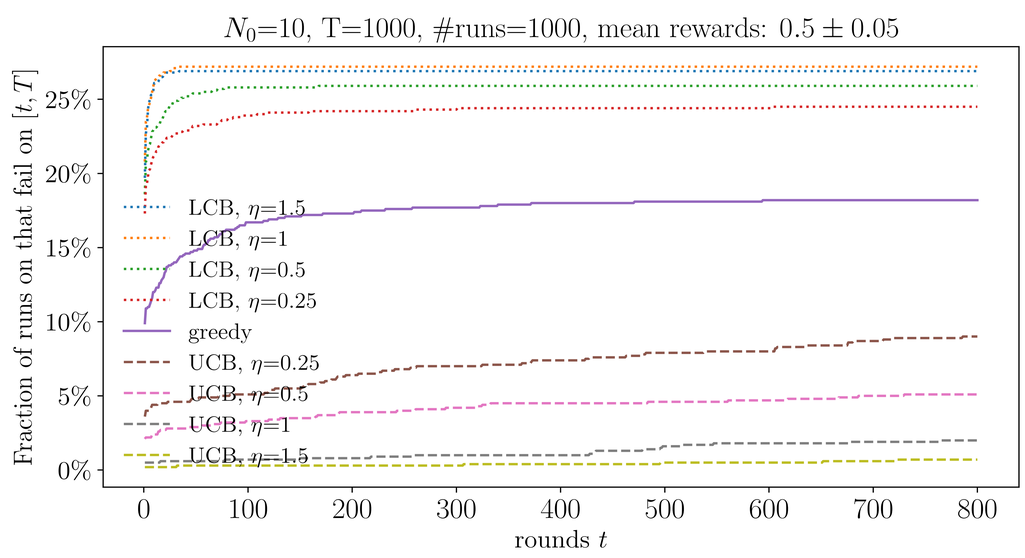
Assn: $\Omega(1+\eta) < N_0 < \Delta^{-2}$

Dependence on N_0: $1/\sqrt{N_0}$ vs. trivial failure, with FailProb exponential in N_0

Optimism → matching upper bound (in η)

Pessimism → worse failure: FailureProb > p_0 for any η.

Representative Simulation



Extensions

- Agents with Bayesian beliefs
- Stronger results for Greedy in Bayesian bandits
- All results extends to K arms (recent, full/arxiv version only)

arxiv:2302.07425 NeurIPS 2023, New Orleans — Poster 1827 {kiarash,hajiagha,suhoshin}@umd.edu; slivkins@microsoft.com