Fixed Confidence Best Arm Identification in Bayesian Settings

Kyoungseok Jang[1] Junpei Komiyama[2] Kazutoshi Yamazaki[3]

[1] Università Degli Studi di Milano [2] NYU [3] University of Queensland

November 15, 2024

Our problem setup: FC-BAI, Bayesian setting

Best-Arm Identification in a **Bayesian Setting**

Step 0: True mean values $\boldsymbol{\mu} := (\mu_i)_{i=1}^K$ are **drawn** from a known prior distribution $\boldsymbol{H}$ before the game starts.

Step 1: While stopping condition is False

- Sample arm A_t and observe its reward r_t

Step 2: Recommend arm J based on learner's history.

Fixed-Confidence setting

- Objective: Minimize $\boxed{\mathbf{E}_{\boldsymbol{\mu} \sim \boldsymbol{H}}[\tau]}$ (τ: Stopping time)
- Constraint: $\boxed{\mathbf{P}_{\boldsymbol{\mu} \sim \boldsymbol{H}}} \left(J \neq \max_{i \in [K]} \mu_i \right) \leq \delta$

Traditional problem setup: FC-BAI, Frequentist setting

Best-Arm Identification in a **Frequentist Setting**

Step 0: True mean values $\boldsymbol{\mu} := (\mu_i)_{i=1}^K$ are **predetermined** before the game starts.

Step 1: While stopping condition is False

- Sample arm A_t and observe its reward r_t

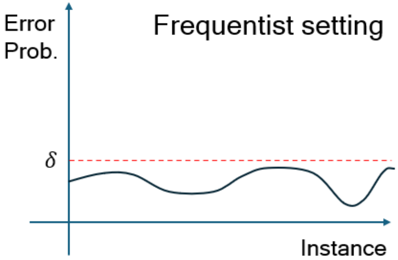
Step 2: Recommend arm J based on learner's history.

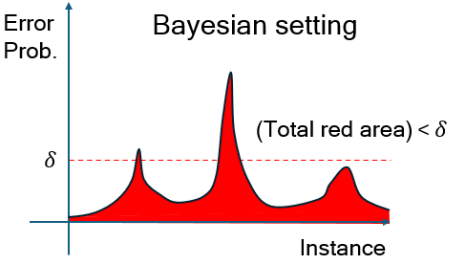
Fixed-Confidence setting

- Objective: Minimize $\boxed{\mathbf{E}_{\boldsymbol{\mu}}[\tau]}$
- Constraint: $\boxed{\mathbf{P}_{\boldsymbol{\mu}}}\left(J \neq \max_{i\in[K]} \mu_i\right) \leq \delta$

What's the difference?

- **Frequentist δ-correct algorithm:** For all possible $\boldsymbol{\mu}$, error probability should be smaller than δ.
- **Bayesian δ-correct algorithm:** It is okay to have a larger error probability in some $\boldsymbol{\mu}$, but 'the expectation of the error probability' should be smaller than δ.



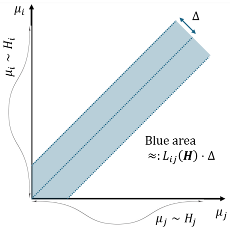


Main Contributions

This is **the first** FC-BAI in Bayesian settings, and

1. We found out that **no Frequentist δ-correct algorithm** can have the finite expected stopping time.
2. Lower bound of the stopping time of $\Omega\left(\frac{L(\boldsymbol{H})^2}{\delta}\right)$
3. Devised an elimination algorithm with stopping time upper bound of $O\left(\frac{L(\boldsymbol{H})^2}{\delta}\log\frac{L(\boldsymbol{H})}{\delta}\right)$

where $L(\boldsymbol{H})$ is our novel 'prior-dependent sample complexity.'



Why δ-correct algorithms fail in Bayesian settings

Informal explanation: $K = 2$

- By (Garivier et al., 2016), for each instance $\boldsymbol{\mu} = (\mu_1, \mu_2)$, the expected stopping time is lower bounded by $\frac{\log \delta^{-1}}{(\mu_1 - \mu_2)^2}$
- We proved that $\mathbf{P}_{\boldsymbol{\mu} \sim \boldsymbol{H}}(|\mu_1 - \mu_2| \leq \epsilon) \approx L(\boldsymbol{H})\epsilon$
- With probability $L(\boldsymbol{H})\epsilon$, the frequentist δ-correct algorithm spends at least $\frac{1}{\epsilon^2}$.
- Expected stopping time is at least $\frac{1}{\epsilon}$ for an arbitrary small ϵ!

Problem: The learner spends too many samples for the case when the suboptimality gap is small!

Algorithm: Successive Elimination with Early-Stopping

Algorithm 1 Successive Elimination with Early-Stopping

Input: Confidence level δ, prior $\boldsymbol{H}$
$\Delta_0 := \frac{\delta}{4L(\boldsymbol{H})}$
Draw each arm once.
Initialize the candidate of best arms $\mathcal{A}(1) = [K]$.
$t = 1$
while True **do**
  Draw each arm in $\mathcal{A}(t)$ once.
  $t \leftarrow t + |\mathcal{A}(t)|\ \hat{\Delta}^{\text{safe}}(t)$.
  for $i \in \mathcal{A}(t)$ **do**
    Calculate $\text{UCB}(i,t)$ and $\text{LCB}(i,t)$ from (4).
    if $\text{UCB}(i,t) \le \max_j \text{LCB}(j,t)$ **then**
      $\mathcal{A}(t) \leftarrow \mathcal{A}(t) \setminus \{i\}$.
    end if
  end for
  if $|\mathcal{A}(t)| = 1$ **then**
    Return arm J in $\mathcal{A}(t)$.
  end if
  Calculate safe empirical gap
  if $\hat{\Delta}^{\text{safe}}(t) \le \Delta_0$ **then**
    Return arm J which is uniformly sampled from $\mathcal{A}(t)$.
  end if
end while

- Successive elimination-based
- Main difference:
  - Early-stopping: When the gap is small enough, stop additional sampling!
  - It is possible since the 'narrow gap event' rarely happens in Bayesian setups.

Conclusions

First FC-BAI in the bayesian setting

- Contribution 1: Frequentist δ-correct algorithms usually fail in bayesian settings.
- Contribution 2: Lower bound $\Omega\left(\frac{L(\boldsymbol{H})^2}{\delta}\right)$
- Contribution 3: Upper bound $O\left(\frac{L(\boldsymbol{H})^2}{\delta}\log\frac{L(\boldsymbol{H})}{\delta}\right)$

Future work:

- Close the gap between UB and LB
- More distributions (beside gaussians)
- Adapt other popular Frequentist FC-BAI algorithms (Track-and-Stop, Top-two type) to Bayesian FC-BAI problem.

Thank you!

Please come to my poster!

Contact: ksajks@gmail.com

Full details are available in NeurIPS 2024 (QR below)