Fixed-confidence guarantees for Bayesian best-arm identification

X. Shang[1], R. de Heide[2], P. Ménard[1], E. Kaufmann[3,1], M. Valko[4,1,3]

[1] Inria Lille, SequeL team
[2] Universiteit Leiden & CWI
[3] CNRS, CRIStAL
[4] DeepMind

AISTATS 2020

August 2020

What — BAI for finitely-armed bandits

- sampling rule;
  - selects an arm I at each round

What — BAI for finitely-armed bandits

- sampling rule;
- recommendation rule;
  - outputs a guess of the best arm J when the algorithm stops

What — BAI for finitely-armed bandits

- sampling rule;
- recommendation rule;
- stopping rule.
  - Fixed-budget: stops when reach the budget n
  - Fixed-confidence: stops when the probability of recommending a wrong arm is less than δ

What — BAI for finitely-armed bandits

- sampling rule;
- recommendation rule;
- stopping rule.

We are interested in `TTTS` (Top-Two Thompson Sampling)

Why

- Beyond fixed-budget and fixed-confidence: anytime BAI framework [Jun and Nowak 2016];
- A Bayesian competitor for BAI as Thompson sampling to `UCB` for regret minimizing?

How - Contributions of this paper

- New theoretical insights on `TTTS`;
- Computational improvement.

What we know about TTTS...

Input: β
for $n = 1, 2, \ldots$ **do**
  $\forall i \in \mathcal{A}, \ \theta_i \sim \Pi_n$
  $I^{(1)} = \arg\max_{i=0,\ldots,m} \theta_i$
  if $U(\sim \mathcal{U}([0,1])) > \beta$ **then**
    while $I^{(2)} \neq I^{(1)}$ **do**
      $\forall i \in \mathcal{A}, \ \theta'_i \sim \Pi_n$
      $I^{(2)} \leftarrow \arg\max_{i=0,\ldots,m} \theta'_i$
    end while
    $I^{(1)} \leftarrow I^{(2)}$
  end if
  `evaluate arm` $I^{(1)}$
  `update` Π_n
end for

What we know about TTTS... (Posterior convergence)

Theorem (Russo 2016)

*Under TTTS, for **bounded** 1-dimensional exponential family priors, it holds almost surely that*

$$\lim_{n\to\infty} -\frac{1}{n}\log(1-\alpha_{n,I^\star}) = \Gamma^\star_\beta,$$

where

$$\alpha_{n,i} \triangleq \Pi_n(\theta_i > \max_{j\neq i}\theta_j).$$

What we know about TTTS... (Complexity)

Definition

Let $\Sigma_K = \{\boldsymbol{\omega} : \sum_{k=1}^{K} \omega_k = 1, \omega_k \geq 0\}$ and define for all $i \neq I^\star$

$$C_i(\omega, \omega') \triangleq \min_{x \in \mathcal{I}} \ \omega d(\mu_{I^\star}; x) + \omega' d(\mu_i; x),$$

where $d(\mu, \mu')$ is the KL-divergence. We define

$$\Gamma^\star_\beta \quad \triangleq \quad \max_{\substack{\boldsymbol{\omega} \in \Sigma_K \\ \omega_{I^\star} = \beta}} \min_{i \neq I^\star} C_i(\omega_{I^\star}, \omega_i).$$

What we know about TTTS... (Complexity)

Definition

Let $\Sigma_K = \{\boldsymbol{\omega} : \sum_{k=1}^{K} \omega_k = 1, \omega_k \geq 0\}$ and define for all $i \neq I^\star$

$$C_i(\omega, \omega') \triangleq \min_{x \in \mathcal{I}} \; \omega d(\mu_{I^\star}; x) + \omega' d(\mu_i; x),$$

where $d(\mu, \mu')$ is the KL-divergence. We define

$$\Gamma^\star_\beta \quad \triangleq \quad \max_{\substack{\boldsymbol{\omega} \in \Sigma_K \\ \omega_{I^\star} = \beta}} \min_{i \neq I^\star} C_i(\omega_{I^\star}, \omega_i).$$

In particular, for Gaussian bandits...

$$\Gamma^\star_\beta = \max_{\boldsymbol{\omega} : \omega_{I^\star} = \beta} \min_{i \neq I^\star} \frac{(\mu_{I^\star} - \mu_i)^2}{2\sigma^2(1/\omega_i + 1/\beta)}.$$

What we want to know about TTTS

- Can we remove the *boundedness* assumption of the prior?

Note however that...

Boundedness is not quite precise, the exact term should be INECCSI = Interior-Non-Empty Closure (is) Compact Subset (of) Interior.

What we want to know about TTTS

- Can we remove the *boundedness* assumption of the prior?
- What can we say about the sample complexity in the fixed-confidence setting?

What we want to know about TTTS

- Can we remove the *boundedness* assumption of the prior?
- What can we say about the sample complexity in the fixed-confidence setting?

Lower bound

Under any δ-correct strategy satisfying $T_{n,I^\star}/n \to \beta$,

$$\liminf_{\delta \to 0} \frac{\mathbb{E}\left[\tau_\delta\right]}{\ln(1/\delta)} \geq \frac{1}{\Gamma^\star_\beta}.$$

What we want to know about TTTS

- Can we remove the *boundedness* assumption of the prior?
- What can we say about the sample complexity in the fixed-confidence setting?
- Can we have finite-time guarantees?

Main result — Posterior convergence

Theorem

Under TTTS, for Gaussian bandits with improper Gaussian priors, it holds almost surely that

$$\lim_{n\to\infty} -\frac{1}{n}\log(1-\alpha_{n,I^\star}) = \Gamma_\beta^\star.$$

Theorem

Under TTTS, for Bernoulli bandits and uniform priors, it holds almost surely that

$$\lim_{n\to\infty} -\frac{1}{n}\log(1-\alpha_{n,I^\star}) = \Gamma_\beta^\star.$$

Main result — Sample complexity

Theorem

The TTTS sampling rule coupled with the Chernoff stopping rule form a δ-correct BAI strategy. Moreover, if all the arms means are distinct, it satisfies

$$\limsup_{\delta \to 0} \frac{\mathbb{E}\left[\tau_\delta\right]}{\log(1/\delta)} \leq \frac{1}{\Gamma_\beta^\star}.$$

Main result — Sample complexity

Theorem

The TTTS sampling rule coupled with the Chernoff stopping rule form a δ-correct BAI strategy. Moreover, if all the arms means are distinct, it satisfies

$$\limsup_{\delta \to 0} \frac{\mathbb{E}\left[\tau_\delta\right]}{\log(1/\delta)} \leq \frac{1}{\Gamma_\beta^\star}.$$

Recall (Lower bound)

$$\liminf_{\delta \to 0} \frac{\mathbb{E}\left[\tau_\delta\right]}{\ln(1/\delta)} \geq \frac{1}{\Gamma_\beta^\star}.$$

Sample complexity sketch — δ-correctness

Stopping rule

$$\tau_\delta^{\text{Ch.}} \triangleq \inf\left\{n \in \mathbb{N} : \max_{i\in\mathcal{A}} \min_{j\in\mathcal{A}\setminus\{i\}} W_n(i,j) > d_{n,\delta}\right\}. \tag{1}$$

Sample complexity sketch — δ-correctness

Stopping rule

$$\tau_\delta^{\mathrm{Ch.}} \triangleq \inf\left\{ n \in \mathbb{N} : \max_{i\in\mathcal{A}} \min_{j\in\mathcal{A}\setminus\{i\}} W_n(i,j) > d_{n,\delta} \right\}. \tag{1}$$

Transportation cost

Let $\mu_{n,i,j} \triangleq (T_{n,i}\mu_{n,i} + T_{n,j}\mu_{n,j})/(T_{n,i} + T_{n,j})$, then we define

$$W_n(i,j) \triangleq \begin{cases} 0 & \text{if } \mu_{n,j} \geq \mu_{n,i}, \\ W_{n,i,j} + W_{n,j,i} & \text{otherwise,} \end{cases} \tag{2}$$

where $W_{n,i,j} \triangleq T_{n,i} d\left(\mu_{n,i}, \mu_{n,i,j}\right)$ for any i, j.

Sample complexity sketch — δ-correctness

Stopping rule

$$\tau_\delta^{\text{Ch.}} \triangleq \inf\left\{ n \in \mathbb{N} : \max_{i\in\mathcal{A}} \min_{j\in\mathcal{A}\setminus\{i\}} W_n(i,j) > d_{n,\delta} \right\}. \tag{1}$$

In particular, for Gaussian bandits...

$$W_n(i,j) = \frac{(\mu_{n,i} - \mu_{n,j})^2}{2\sigma^2(1/T_{n,i} + 1/T_{n,j})} \mathbb{1}\{\mu_{n,j} < \mu_{n,i}\}.$$

Sample complexity sketch — δ-correctness

Stopping rule

$$\tau_\delta^{\text{Ch.}} \triangleq \inf \left\{ n \in \mathbb{N} : \max_{i \in \mathcal{A}} \min_{j \in \mathcal{A} \setminus \{i\}} W_n(i,j) > d_{n,\delta} \right\}. \tag{1}$$

Theorem

The TTTS sampling rule coupled with the Chernoff stopping rule (1) *with a threshold* $d_{n,\delta}$ *derived from Kaufmann and Koolen 2018 and the recommendation rule* $J_t = \arg\max_i \mu_{n,i}$, *form a* δ*-correct BAI strategy.*

Sample complexity sketch — Sufficient condition for β-optimality

Lemma

Let $\delta, \beta \in (0,1)$. For any sampling rule which satisfies $\mathbb{E}\left[T_\beta^\varepsilon\right] < \infty$ for all $\varepsilon > 0$, we have

$$\limsup_{\delta \to 0} \frac{\mathbb{E}\left[\tau_\delta\right]}{\log(1/\delta)} \le \frac{1}{\Gamma_\beta^\star},$$

if the sampling rule is coupled with stopping rule (1).

Sample complexity sketch — Sufficient condition for β-optimality

Lemma

Let $\delta, \beta \in (0,1)$. For any sampling rule which satisfies $\mathbb{E}\left[T_\beta^\varepsilon\right] < \infty$ for all $\varepsilon > 0$, we have

$$\limsup_{\delta \to 0} \frac{\mathbb{E}\left[\tau_\delta\right]}{\log(1/\delta)} \leq \frac{1}{\Gamma_\beta^\star},$$

if the sampling rule is coupled with stopping rule (1).

$$T_\beta^\varepsilon \triangleq \inf\left\{N \in \mathbb{N} : \max_{i \in \mathcal{A}} |T_{n,i}/n - \omega_i^\beta| \leq \varepsilon, \forall n \geq N\right\}$$

Sample complexity sketch — Core theorem

Theorem

Under TTTS, $\mathbb{E}\left[T_\beta^\varepsilon\right] < +\infty$.

The proof is inspired by Qin et al. (2017), but some technical novelties are introduced. In particular, our proof is much more intricate due to the randomized nature of the two candidate arms...

Alleviate the computational burden?

Input: β
for $n = 1, 2, \ldots$ **do**
    $\forall i \in \mathcal{A},\ \theta_i \sim \Pi_n$
    $I^{(1)} = \arg\max_{i=0,\ldots,m} \theta_i$
    if $U(\sim \mathcal{U}([0,1])) > \beta$ **then**
        while $I^{(2)} \neq I^{(1)}$ **do**
            $\forall i \in \mathcal{A},\ \theta'_i \sim \Pi_n$
            $I^{(2)} \leftarrow \arg\max_{i=0,\ldots,m} \theta'_i$ {Re-sampling phase}
        end while
        $I^{(1)} \leftarrow I^{(2)}$
    end if
    `evaluate arm` $I^{(1)}$
    `update` Π_n
end for

Alleviate the computational burden?

Input: β
for $n = 1, 2, \ldots$ **do**
  $\forall i \in \mathcal{A}$, $\theta_i \sim \Pi_n$
  $I^{(1)} = \arg\max_{i=0,\ldots,m} \theta_i$
  if $U(\sim \mathcal{U}([0,1])) > \beta$ **then**
    $I^{(2)} \leftarrow \arg\min_{i \neq I^{(1)}} W_n(I^{(1)}, i)$ {T3C}
    $I^{(1)} \leftarrow I^{(2)}$
  end if
  evaluate arm $I^{(1)}$
  update Π_n
end for

Main result — Sample complexity T3C

Theorem

The T3C sampling rule coupled with the Chernoff stopping rule form a δ-correct BAI strategy. Moreover, if all the arms means are distinct, it satisfies

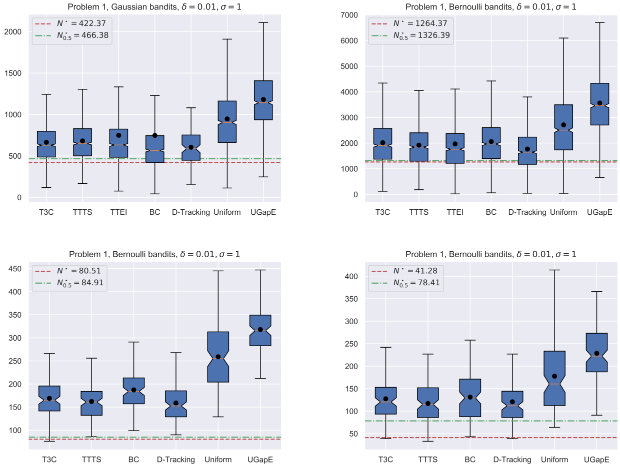
$$\limsup_{\delta \to 0} \frac{\mathbb{E}\left[\tau_\delta\right]}{\log(1/\delta)} \leq \frac{1}{\Gamma_\beta^\star}.$$

Some illustrations — Time consumption

Sampling rule	T3C	TTTS	Uniform
Execution time (s)	1.6×10^{-5}	2.3×10^{-4}	6×10^{-6}

Table: average execution time in seconds for different sampling rules.

Some illustrations — Average stopping time



Still far from the Holy Grail...

- Finite-time analysis (fixed-budget setting?)

Conclusion

More details on `TTTS` and `T3C`

Check out [Shang et al. 2020].

References

Thank you!

Emilie Kaufmann and Wouter Koolen. "Mixture martingales revisited with applications to sequential tests and confidence intervals". In: 2018.

Daniel Russo. "Simple Bayesian algorithms for best arm identification". In: *29th CoLT*. 2016.

Xuedong Shang, Rianne de Heide, Pierre Ménard, Emilie Kaufmann, and Michal Valko. "Fixed-confidence guarantees for Bayesian best-arm identification". In: *23rd AIStats*. 2020.