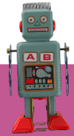
Instance-optimal algorithms for A/B testing

Wouter M. Koolen

May 25, 2022

Senior Researcher
Machine Learning Group
Centrum Wiskunde & Informatica

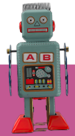


Main message

- A/B Testing aims to support decision making
- A/B Testing tools constrain decision making
- Flexible testing ⇔ creative decisions

Today:

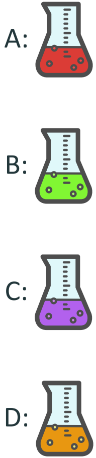
- How difficult is a given testing problem?
- How to solve a given testing problem?

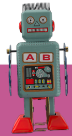


Setting and Problem

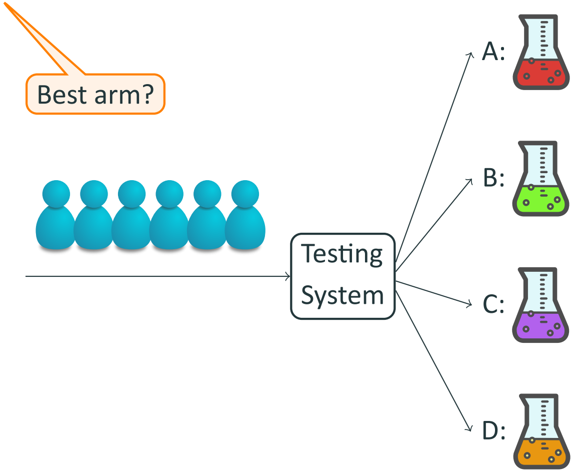
Bare-bones sequential testing setup

Best arm?

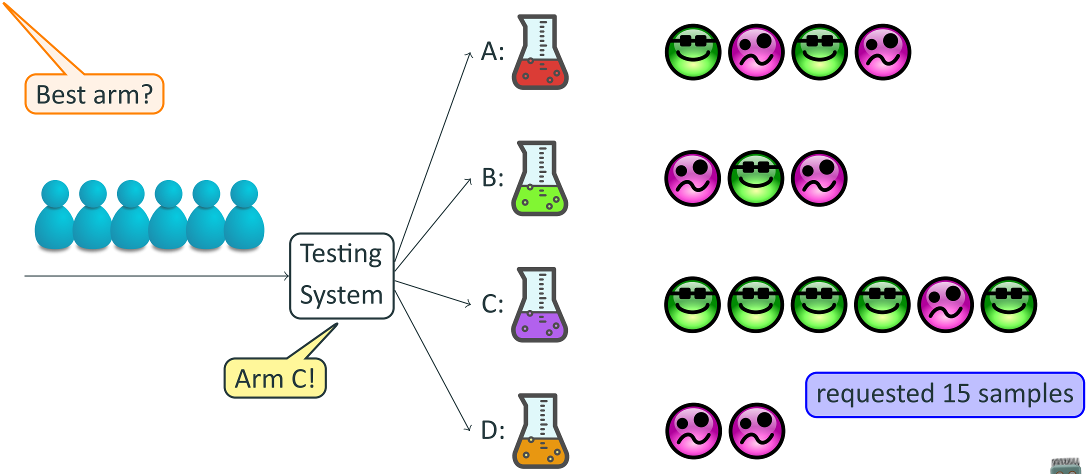




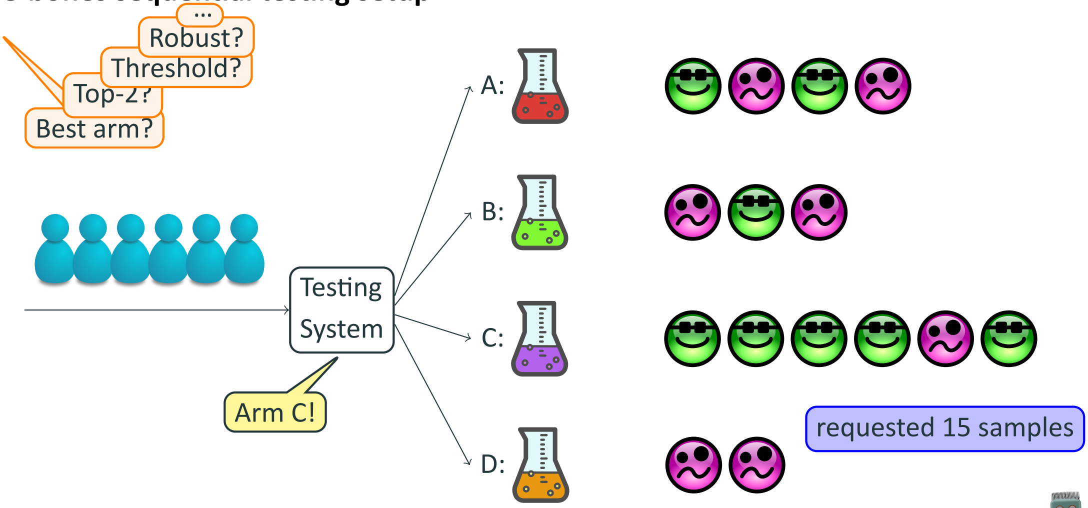
Bare-bones sequential testing setup



Bare-bones sequential testing setup

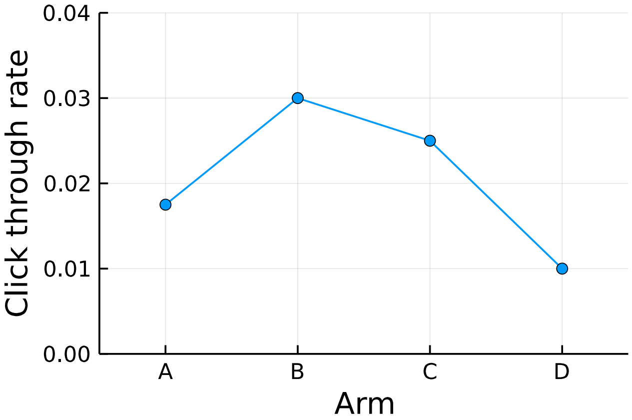


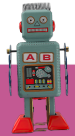
Bare-bones sequential testing setup



Model for the Environment

The unknown true bandit instance $\boldsymbol{\mu} = (\mu_A, \mu_B, \mu_C, \mu_D)$





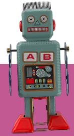
Algorithms for fixed-confidence testing $\delta = 0.05$

Specified by:

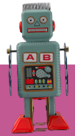
- Sampling rule
- Stopping rule
- Recommendation rule

Reliable Must be δ-correct for *any* bandit

Efficient Minimise # samples



Characteristic Time and Oracle Weights



Characteristic Time and Oracle Weights

Answering correctly for μ requires data to reject all bandits where that answer is wrong.

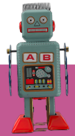
Theorem (Garivier and Kaufmann, 2016)

Any δ-correct testing algorithm must, for any bandit instance μ, take samples at least

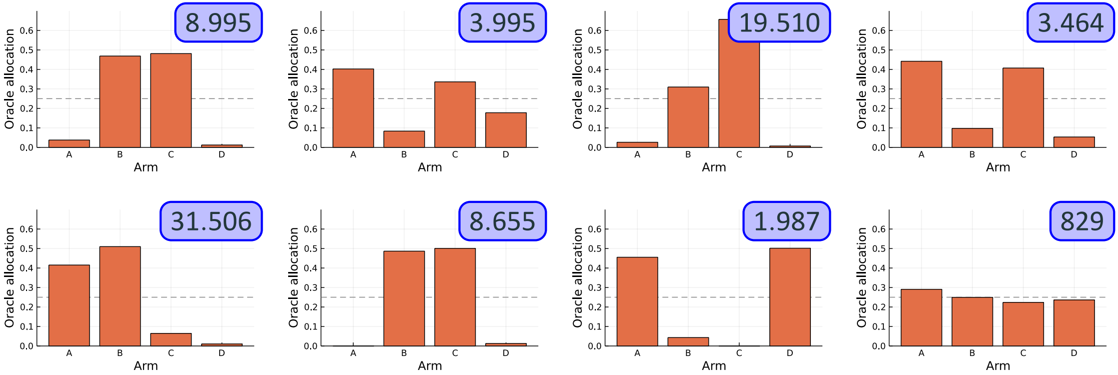
$$\text{samples}(\mu) \geq \ln \frac{1}{\delta} \cdot \frac{1}{\displaystyle \max_{\text{arm proportions } \boldsymbol{w}} \; \min_{\substack{\text{bandit } \boldsymbol{\lambda} \text{ with answer} \\ \text{different from that of } \boldsymbol{\mu}}} \; \sum_{\text{arm } a} w_a \operatorname{KL}(\mu_a, \lambda_a)}$$

Why should we care?

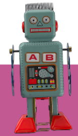
- Characterises* complexity of each problem instance μ
- Optimal testing algorithm must sample with proportions $\arg\max_{\boldsymbol{w}}$



Examples: variations of Best Arm question



- Sample complexities vastly different between questions
- Optimal allocation depends strongly on the specific question being asked



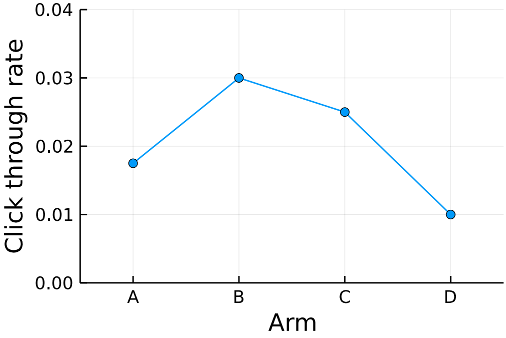
CWI Machine Learning

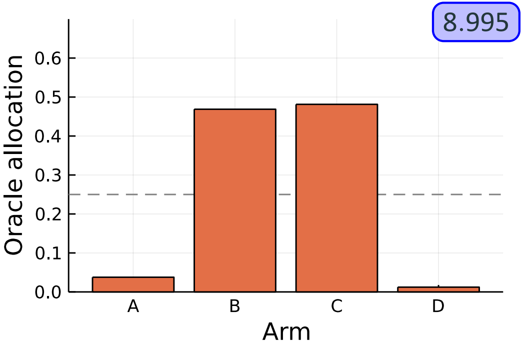
A/B TESTING

Best Arm Identification (BAI)

$$\arg\max_{a \in \mathcal{A}} \mu_a$$

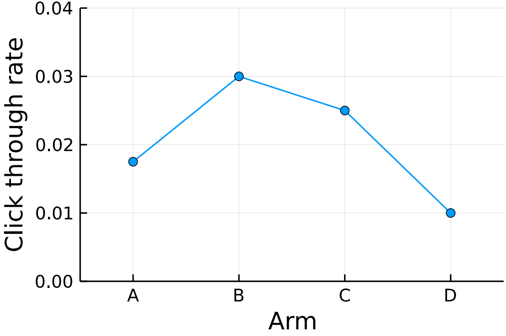
where $\mathcal{A} = \{A, B, C, D\}$

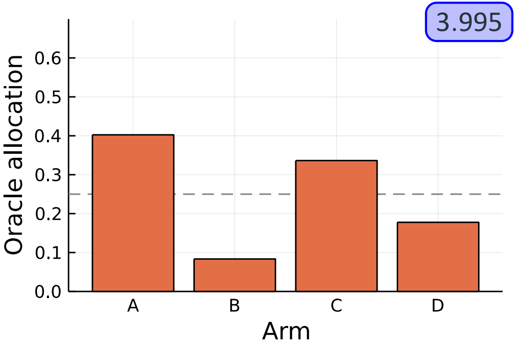




All-Better-than-the-Control (ABC)

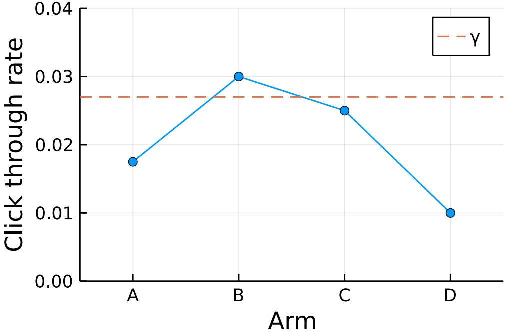
$$\{a \in \{B, C, D\} \mid \mu_a \geq \mu_A\}$$

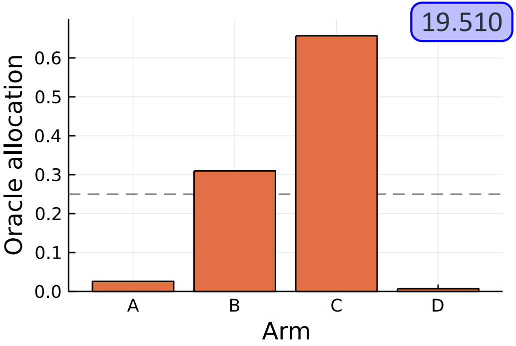




All-Better-than-Threshold

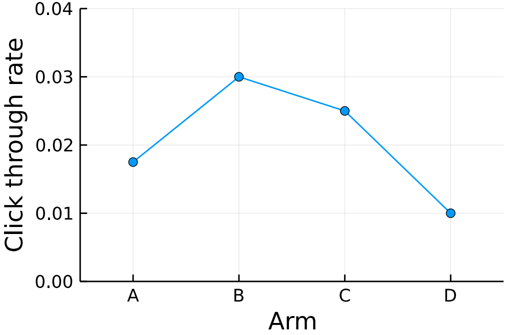
$$\{a \in \mathcal{A} \mid \mu_a \geq \gamma\}$$

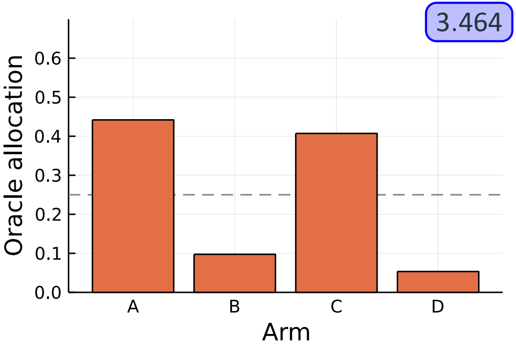




Top-2

$$\{a \in \mathcal{A} \mid \mu_a \geq \mu_{(2)}\}$$ where $\mu_{(1)} \geq \mu_{(2)} \geq \dots$

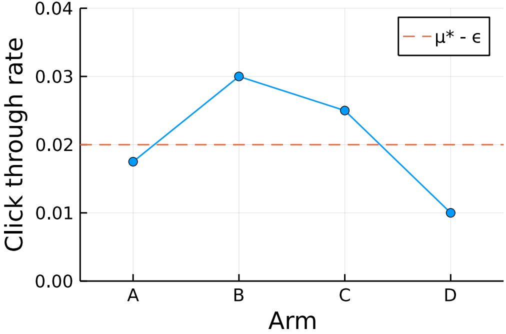


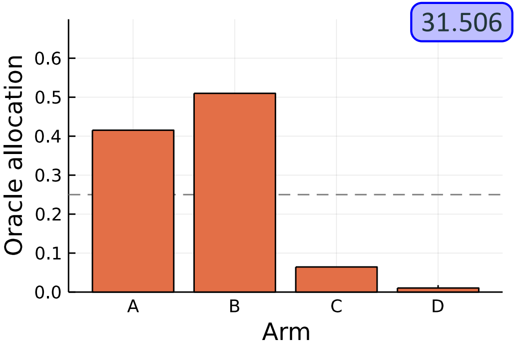


Near-optimal arms

$$\{a \in \mathcal{A} \mid \mu_a \geq \mu^* - \epsilon\}$$

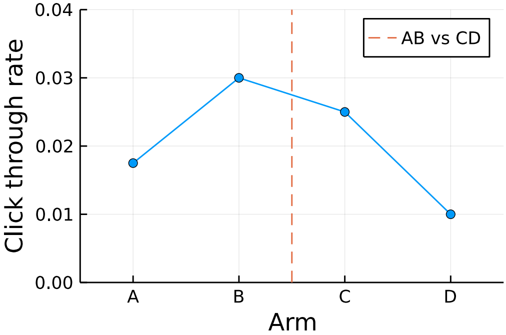
where $\mu^* = \max_{a \in \mathcal{A}} \mu_a$

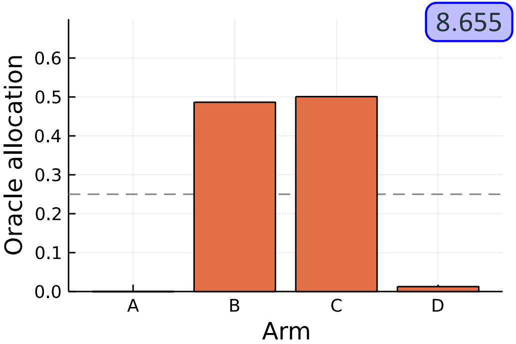




Winning Side

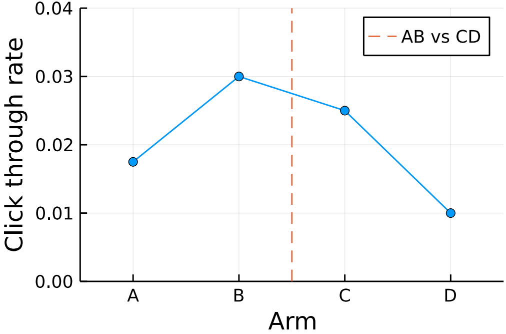
$$\arg\max\{\max\{\mu_A, \mu_B\}, \max\{\mu_C, \mu_D\}\}$$

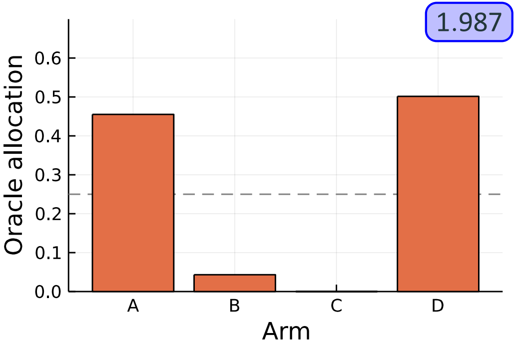




Robust best arm

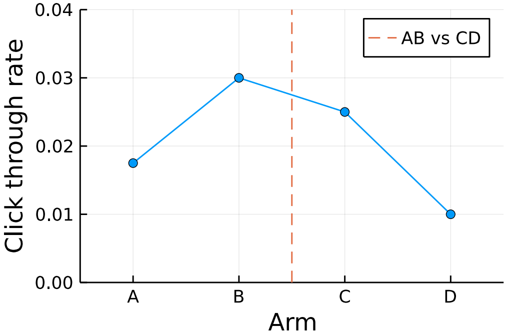
$$\arg\max\left\{\min\left\{\mu_A,\mu_B\right\},\min\left\{\mu_C,\mu_D\right\}\right\}$$

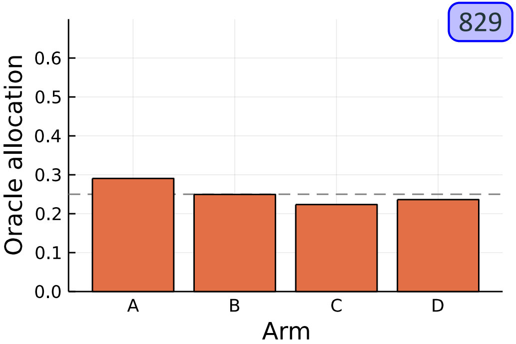




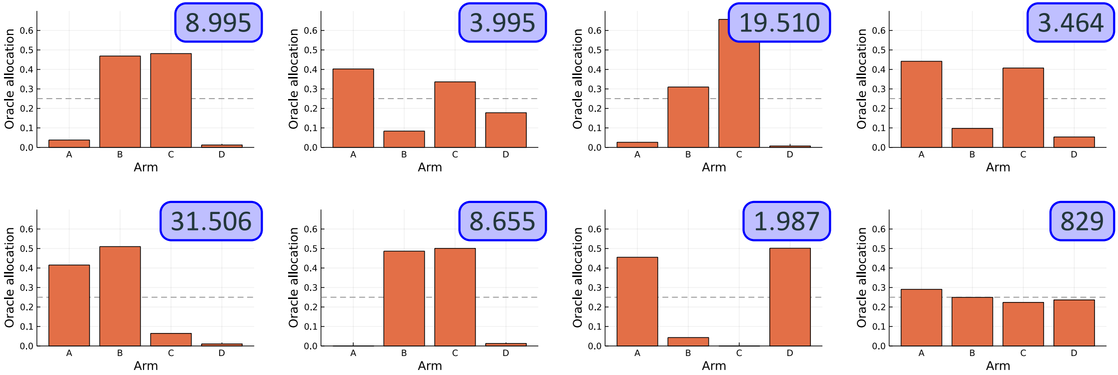
Largest Profit

$$\arg\max \{\mu_A - \mu_B, \mu_C - \mu_D\}$$

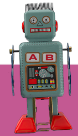




Overview of Optimal Sampling Allocations



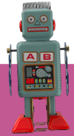
- Sample complexities vastly different between questions
- Optimal allocation depends strongly on the specific question being asked



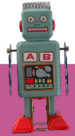
Where this brings us

- Specific question posed matters
- Optimise it for the eventual decision of interest

But how?



Canonical Path to Optimal Algorithms

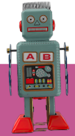


Instance-Optimal Algorithms

Sample complexity lower bound at $\boldsymbol{\mu}$ governed by:

$$\max_{\text{arm proportions } \boldsymbol{w}} \; \min_{\substack{\text{bandit } \boldsymbol{\lambda} \text{ with answer} \\ \text{different from that of } \boldsymbol{\mu}}} \; \sum_{\text{arm } a} w_a \, \mathrm{KL}(\mu_a, \lambda_a)$$

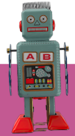
Main challenge: sampling like $\arg\max_{\boldsymbol{w}}$ without knowing $\boldsymbol{\mu}$.



Saddle Point Approach

CWI specialty

Approx. solve saddle point problem iteratively: $w_1, w_2, \ldots \to w^*(\mu)$



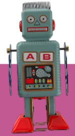
Saddle Point Approach

CWI specialty

Approx. solve saddle point problem iteratively: $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots \to \boldsymbol{w}^*(\boldsymbol{\mu})$

Main pipeline (Degenne, Koolen, and Ménard, 2019):

- Pick arm $A_t \sim \boldsymbol{w}_t$
- Plug-in estimate $\hat{\boldsymbol{\mu}}_t$ (so problem is shifting).
- Advance the saddle point solver one iteration per bandit interaction.
- Add optimism to gradients to induce exploration ($\hat{\boldsymbol{\mu}}_t \to \boldsymbol{\mu}$).
- Regret bounds + concentration + optimism $\Rightarrow$ finite-confidence guarantee:



Saddle Point Approach

CWI specialty

Approx. solve saddle point problem iteratively: $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots \to \boldsymbol{w}^*(\boldsymbol{\mu})$

Main pipeline (Degenne, Koolen, and Ménard, 2019):

- Pick arm $A_t \sim \boldsymbol{w}_t$
- Plug-in estimate $\hat{\boldsymbol{\mu}}_t$ (so problem is shifting).
- Advance the saddle point solver one iteration per bandit interaction.
- Add optimism to gradients to induce exploration ($\hat{\boldsymbol{\mu}}_t \to \boldsymbol{\mu}$).
- Regret bounds + concentration + optimism $\Rightarrow$ finite-confidence guarantee:

Theorem (Instance-Optimality)

For every $\delta \in (0,1)$ and bandit $\boldsymbol{\mu}$, the above scheme takes samples bounded by

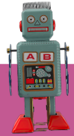
$$\textit{samples}(\boldsymbol{\mu}) \le \boxed{\textit{samples}(\boldsymbol{\mu})} \ln \frac{1}{\delta} + o\left(\ln \frac{1}{\delta}\right)$$

Conclusion

Conclusion

- Every sequential testing problem has associated
  - characteristic time: quantifying sample complexity, and
  - oracle allocation: encoding desired optimal behaviour
- Both are highly sensitive to the precise question posed
- So: a lot to gain by fine-tuning the testing effort to the "why"
- Once the question is crisp, optimal algorithms are quickly becoming technology.
  - State-of-art performance in many applications

Thanks!



References

Degenne, R., W. M. Koolen, and P. Ménard (Dec. 2019). “Non-Asymptotic Pure Exploration by Solving Games”. In: *Advances in Neural Information Processing Systems (NeurIPS) 32*. Ed. by H. Wallach, H. Larochelle, A. Beygelzimer, F. d’Alché-Buc, E. Fox, and R. Garnett. Curran Associates, Inc., pp. 14492–14501.

Garivier, A. and E. Kaufmann (2016). “Optimal Best arm Identification with Fixed Confidence”. In: *Proceedings of the 29th Conference On Learning Theory (COLT)*.