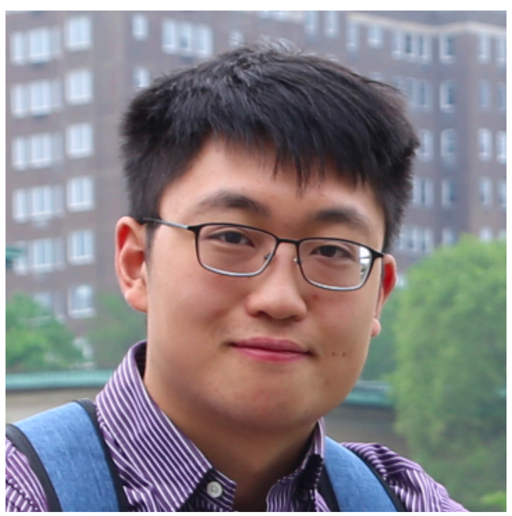
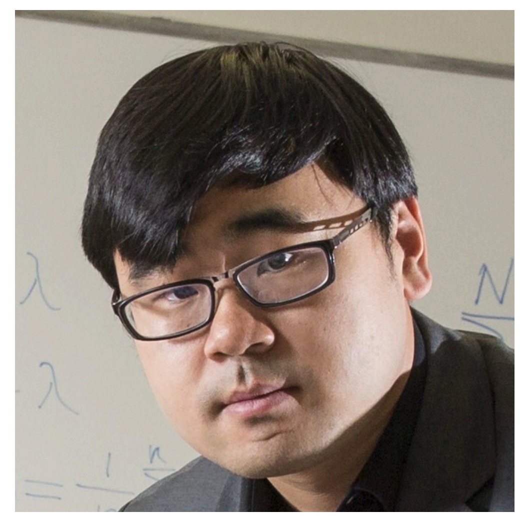
A unified framework for bandit multiple testing

NeurIPS 2021

Joint work with:



Ziyu Xu (Neil)
(CMU)



Ruodu Wang
(Waterloo)

Aaditya Ramdas
(CMU)

UNIVERSITY OF WATERLOO

Carnegie Mellon University

Subjects

In a sequential clinical trial, adaptively allocate subjects to *minimize* the number of trials required to:

1. Find most of the drug candidates with positive effect (true discoveries)

2. Make sure we don't erroneously believe ineffective drug candidates are effective (make false discoveries).

Drug candidates

The multiple testing problem

K hypotheses we wish to test $H_1, H_2, \ldots, H_K$ e.g. "this drug candidate has no positive effect."

The set of null hypotheses is $\mathrm{N} \subseteq \{1,\ldots,K\}$

e.g. drugs candidates that actually have no positive effect

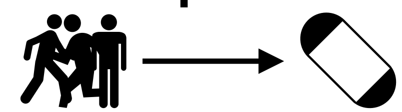
Procedure outputs a discovery set $S \subseteq \{1,\ldots,K\}$

We want to control the frequency of making "false discoveries".

$$\mathrm{FDP} := \frac{|\mathrm{N} \cap \mathrm{S}|}{|\mathrm{S}| \vee 1} \qquad \mathrm{FDR} := \mathbb{E}[\mathrm{FDP}]$$

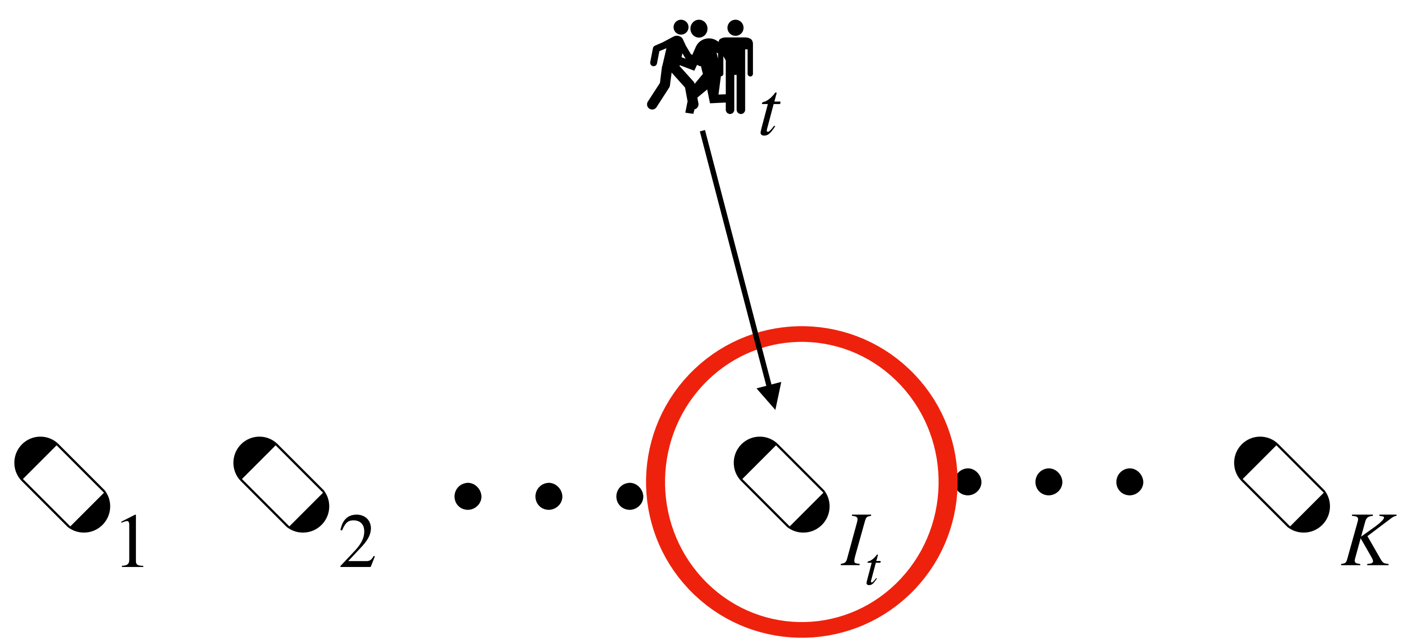
FDR is required to be controlled under a fixed constant $\delta \in (0,1)$

The bandit approach to multiple testing

The multi-armed bandit models the adaptive allocation of new subjects to one of the treatments. 

k-th arm is associated with the hypothesis (drug) H_k

At each time step t select a single arm I_t and sample $X_{I_t,t}$



When we stop sampling, produce S that satisfies: $\mathrm{FDR} \leq \delta$

Prior work: p-values (+sequential analog: p-processes)

Uses rewards sampled from arm k by time t to construct:

a p-process $(p_{k,t})_{t\in\mathbb{N}}$ for each hypothesis k

$\mathbf{Pr}(\exists t \in \mathbb{N} : p_{k,t} \leq s) \leq s$ for all $s \in (0,1)$ when $k \in \mathrm{N}$

No restrictions on $(p_{k,t})_{t\in\mathbb{N}}$ when $k \notin \mathrm{N}$

(BH procedure at level α) Output the largest set satisfying:

$$\max_{k\in \mathrm{S}} p_{k,t} \leq \frac{\alpha|\mathrm{S}|}{K}$$

Theorem (Benjamini and Hochberg 1995, Benjamini and Yekutieli 2001)

$\mathrm{FDR} \leq \alpha$ — $p_{1,t},\ldots,p_{K,t}$ are independent

$\mathrm{FDR} \leq \alpha\log(1/\alpha)$ — $p_{1,t},\ldots,p_{K,t}$ are dependent only through sampling

$\mathrm{FDR} \lesssim \alpha\log(K)$ — $p_{1,t},\ldots,p_{K,t}$ are arbitrarily dependent

(Jamieson and Jain 2018, Su 2018)

Limitations of p-processes

Adaptivity in the sampling algorithm ($\log(1/\alpha)$ blow up)

Using p-processes requires correction by an extra $\log K$ factor when there is arbitrary dependence between $p_{1,t}, p_{2,t}, \ldots, p_{K,t}$

Dependence can arise from:

- Arbitrary dependence among rewards $X_{1,t}, X_{2,t}, \ldots, X_{K,t}$ (combinatorial bandits)
- Hypotheses that test a property of multiple arms (e.g. is the covariance of the rewards of two arms 0?)
- Using previous data that may have come from dependent sources.

E-processes: an alternative to p-processes

Let $e_1, e_2, \ldots, e_K$ be k e-values corresponding to $H_1, H_2, \ldots, H_K$

e_k is an e-value iff $\mathbb{E}[e_k] \leq 1$ and e_k is nonnegative when $k \in \mathrm{N}$

E-processes are the sequential analog of e-values.

$(e_{k,t})_{t\in\mathbb{N}}$ is an e-process iff $e_{k,\tau}$ is an e-value for all stopping times τ

(random time that is a function of the already observed rewards)

Recently developed alternative to p-values and p-processes, and are fundamentally connected to martingales. (Grünwald et al. 2020, Shafer 2020, Ramdas et al. 2020)

E-values for multiple testing: e-BH

We make no assumptions on the e-values - they may be arbitrarily dependent

self-consistency at level α property: $\min_{k \in S} e_k \geq \frac{K}{\alpha |S|}$

Theorem (Wang and Ramdas 2020)

$\mathrm{FDR} \leq \alpha$ for any self-consistent procedure on e-values.

Fewer assumptions and applies to more procedures than BH

e-BH: outputs largest self-consistent set.

$1/e_k$ is a p-value. Thus, e-BH is identical to applying BH on inverse of e-values.

E-processes in the bandit setting

Uses rewards sampled from arm k by time t to construct an e-process $(e_{k,t})_{t\in\mathbb{N}}$ for each arm/hypothesis k

When the algorithm stops sampling (at stopping time τ)

(e-BH procedure at level δ) Output the largest set satisfying:

$$\min_{k\in S} e_k \geq \frac{k}{\delta |S|}$$

$e_{1,\tau}, e_{2,\tau}, \ldots, e_{K,\tau}$ are e-values.

Output of e-BH guarantees $\mathrm{FDR} \leq \delta$

regardless of dependence structure (e.g. adaptive sampling algorithm, dependence among rewards etc.) among the e-values

Power and sample complexity

$$\text{TPR} := \mathbb{E}\left[\frac{|\,\text{A} \cap \text{S}\,|}{|\,\text{A}\,|}\right] \qquad \text{A} := \{1,\ldots,K\} \backslash \text{N}$$

When we stop sampling, produce S that satisfies:

$$\text{FDR} \le \delta, \qquad \text{TPR} \ge 1 - \delta$$

power constraint

Sample complexity: # of samples required to output S with the above guarantees.

Theorem:

Under the same assumptions (i.e. independent and bounded rewards, single arm, etc.) and sampling strategy as the p-process algorithm (Jamieson and Jain 2018),

E-processes with e-BH achieve matching sample complexity bounds w/ p-processes and BH (up to a constant).

Conclusion: a unified framework for FDR control

Any algorithm that outputs discoveries through e-BH has valid FDR control regardless of the underlying data distributions or hypotheses being tested.

Thus, e-processes and e-BH provide a framework for designing algorithms with valid FDR control in any situation, including:

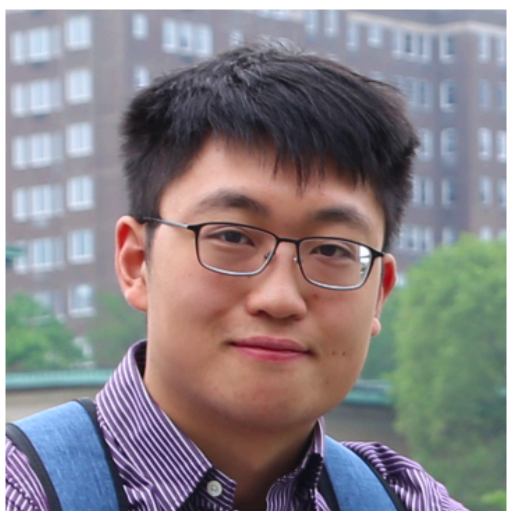
- Dependent reward distributions
- Hypotheses involving multiple arms
- Multi-agent scenarios where agents want to combine collected data
- Structural constraints on the discovery set

Provably matches performance of best algorithm for basic single arm bandit case (Jamieson and Jain 2018)

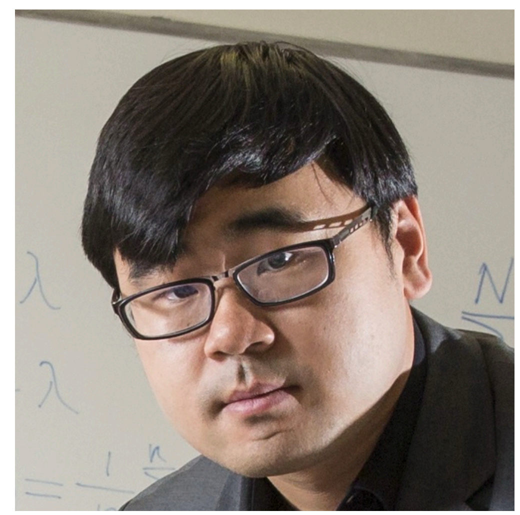
A unified framework for bandit multiple testing

2nd WSSAF Conference

Joint work with:



Ziyu Xu (Neil)
(CMU)



Ruodu Wang
(Waterloo)

Aaditya Ramdas
(CMU)

Link to paper: https://arxiv.org/abs/2107.07322

UNIVERSITY OF WATERLOO

Carnegie Mellon University