A framework for Multi-A(rmed)/B(andit) testing with online FDR control

Fanny Yang, UC Berkeley

Joint work with Aaditya Ramdas, Kevin Jamieson and Martin Wainwright

MCP Conference, June 2017
A company has an on-going sale, and it’s going alright…

…a new marketing student suggests different layout → much better!

Source: https://blog.optimizely.com
What and how to optimize the choices?
Outline

• Status quo (A/B testing)
• Unaddressed practical requirements
• Review of known sequential procedures
  Multi-Armed Bandit algorithms
  Online FDR procedures
• Combining all three frameworks and guarantees
Role of A/B Testing platform

A/B Testing (e.g. Optimizely) only recommends new version if evidence is “significant”!
Allows continuous monitoring given same number of samples
A/B Testing model

• Distributions $P_A$, $P_B$ for A (control), B with means $\mu_{\text{control}}$, $\mu_B$

\[ H_0: \mu_{\text{control}} > \mu_B \quad \text{vs.} \quad H_1: \mu_{\text{control}} < \mu_B \]

• Data are i.i.d. samples from $P_A$, $P_B$

• Compute test statistic $T'$ given data

• Compute probability of $T$ being more extreme: $p = P_{H_0}(T > T')$

• Recommend new arm (reject null) if p-value $p < \text{desired significance } \alpha$ (often 0.05)
A/B testing has been done...

(Commercial Platform: e.g. Optimizely, papers by Johari et al.)

Practical desiderata still unaddressed

• Many arms, limited budget
• Many tests throughout the year

Still be able to continuously monitor
Desideratum 1: many variants (arms)

Problem: Uniform sampling scales linearly with # arms
Approach: sample adaptively to need less traffic/time
Desideratum 2: p-value peeking

Valid p-value (for each t) satisfies: \( P_{H_0}(p_t < \alpha) = \alpha \)

Number of users (samples)

Source: https://blog.optimizely.com
Desideratum 2: p-value peeking

Confidence

\[ 1 - p_t \]

\[ p_t < 0.05 ! \]

\[ p_t < 0.05 ! \]

Number of users (samples)

Problem: \( \Pr(p_t < 0.05 \text{ for some } t) \gg \Pr(p_t < 0.05) = 0.05 \)

Approach: construct p-values that are valid when repeatedly queried for non-uniform sampling

Source: https://blog.optimizely.com
Desideratum 3: many tests over time

Time

Jan
- Exp. 1 (font)
  - A
  - B
- Keep!

April
- Exp. 2 (color)
  - A
  - B
- Change!

May
- Exp. 3 (ads)
  - A
  - B
- Change!

What kind of error to control and how?
Desideratum 3: many tests over time

Company’s interests

• implementing change system-wide has base cost
  → wants to be “sure”, i.e. not too many wrong rejections

• detect better treatments if they give higher revenue (high power)
  → FWER & Bonferroni not great 😒

Compromise
Control the expected ratio \( \frac{\text{# false rejections}}{\text{# rejections}} \) (FDR)

Problem: Neither Vanilla testing nor batch FDR at \( \alpha \) is sufficient
Approach: online control of FDR at level \( \alpha \)
Summary of desiderata

- Many variants (arms)
  \[ \rightarrow \text{Adaptive Sampling (Best-arm Multi-Armed Bandit)} \]
- Allow p-value peeking
  \[ \rightarrow \text{Construct always valid p-values} \]
- Many tests over time
  \[ \rightarrow \text{Online FDR control procedures} \]

\[ \rightarrow \text{Doubly sequential procedure!} \]

Combining two known adaptive procedures while preserving sample efficiency and statistical guarantees
Reviewing known procedures...

Multi Armed-Bandits

Online FDR procedures
Recap: Best-arm MAB schematic

- Compute
  - empirical means \( \hat{\mu}_1(t), \ldots, \hat{\mu}_K(t) \)
  - Confidence bound relative to level \( \delta \)
  - \( \hat{\mu}_{\text{best}}(t) = \max \hat{\mu}_i(t) \)

- Draw sample
- Stop criterion:
  - \( \hat{\mu}_{\text{best}}(t) \) certain, all other \( \hat{\mu}_i(t) \) either certain & little worse, uncertain & much worse

*Stop criterion: \( \hat{\mu}_{\text{best}} \) certain, all other \( \hat{\mu}_i \) either certain & little worse, uncertain & much worse
Recap: Best-arm MAB guarantees

Known results for K arms, confidence parameter $\delta$, i.e. $P(\text{MAB finds best arm}) \geq 1 - \delta$

- Sample complexity guarantees e.g. for the LUCB algorithm (Kalyanakrishnan ‘14) depending on gaps $\Delta_i = \mu_{\text{best}} - \mu_i$

$$\sum_{i \neq \text{best}} \Delta_i^{-2} \log(1/\delta) \quad \text{vs.} \quad K \cdot \max_{i \neq \text{best}} \Delta_i^{-2} \log(1/\delta)$$

(LUCB) (Uniform sampling)

- Matches lower bounds (Garivier et al. ‘16, Simchowitz et al. ’17)
Recap: Online FDR control

At a given time step $j$ …

An online FDR procedure (e.g. LORD, Javanmard, Montanari ‘16)

- $R_{j-1} = 1$
- $\alpha_{j-1}$
- Wealth
- $\gamma_j$

FDR control level $\alpha$
Open: How to combine everything

Meta framework

Multi Armed-Bandits

Online FDR procedures

Continuous monitoring
Conceptual and theoretical challenges

For embedding an MAB algorithm in a testing setup

• What is the right **null hypothesis**? How to incorporate asymmetry in algorithm?

• How to get *always valid* p–values for *non-uniform* and *dependent* samples?

For using MAB in online FDR framework

• What **interaction** between MAB and FDR **preserves best of both worlds** (FDR control, low sample complexity)?
Our contribution 1: Embedding MAB

Null hypothesis:

\[ H_0: \mu_{\text{control}} > \mu_i - \varepsilon \quad \forall i=1\ldots K \quad \text{vs.} \quad H_1: \mu_{\text{control}} + \varepsilon < \mu_i \quad \exists i \]

Prop. Modified MAB finds \( \varepsilon \)-better arm with confidence

Always valid p-value \( p_t \), i.e. \( P(p_t < \alpha \text{ for some } t) \leq \alpha \):

- Law of Iterated Logarithm (LIL) \( \rightarrow \) Always validity
  \[ P(\exists \ t : \mu_i \in [\text{LCB}_i(t, \gamma), \text{UCB}_i(t, \gamma)]) \leq \gamma \]

- For each arm compute
  \[ P_{i,t} = \sup\{\gamma \in [0,1]: \text{LCB}_i(t, \gamma) < \text{UCB}_0(t, \gamma)\} \]
  Final p-value: \( p_t = \min_{i=1,\ldots,K} P_{i,t} \)

Prop. Can compute always valid p-values by sampling with MAB!
Contribution 2: MAB-FDR and guarantees

MAB-FDR meta algorithm

Online FDR procedure

desired FDR level $\alpha$

$\alpha_j \rightarrow R_j (\alpha_j) \rightarrow \text{Exp } j \rightarrow \text{MAB} \rightarrow p_j (\alpha_j) \rightarrow \text{Test } p_j < \alpha_j$

$\alpha_{j+1} \rightarrow R_{j+1} (\alpha_{j+1}) \rightarrow \text{Exp } j+1 \rightarrow \text{MAB} \rightarrow p_{j+1} (\alpha_{j+1}) \rightarrow \text{Test } p_{j+1} < \alpha_{j+1}$

Arm

DEPENDENCE!
Contribution 2: MAB-FDR and guarantees

desired FDR level $\alpha$

Online FDR procedure

MAB-FDR meta algorithm

Theorem “FDR” is at most $\alpha$ at any time.

If the algorithm is not terminated early, then “power” is at least $(1- \alpha)$.
Summary

Introduced a doubly sequential procedure, which simultaneously

- Yields good sample complexity
- Allows continuous monitoring
- Controls FDR in an online fashion
"A framework for Multi-A(rmed)/B(andit) testing with online FDR control”