Outline

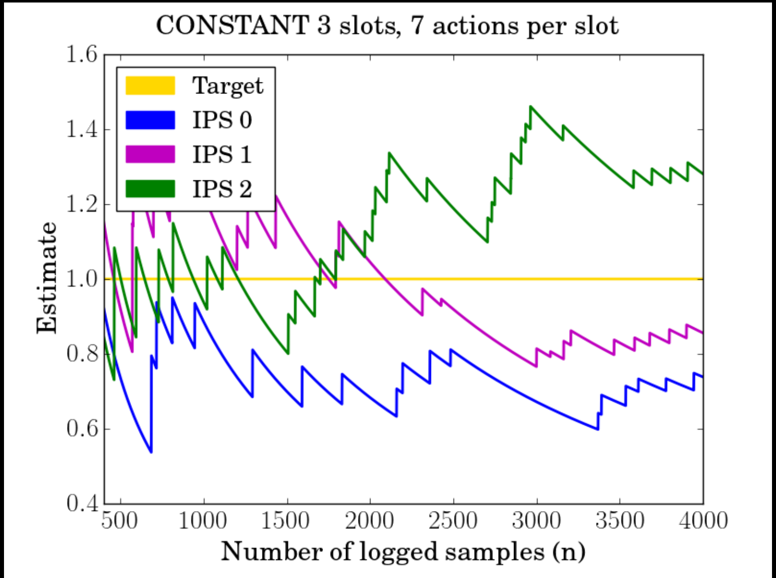
- Offline Evaluation of Online Metrics
- Counterfactual Estimation
- Advanced Estimators
  1. Self-Normalized Estimator
  2. Doubly Robust Estimator
  3. Slates Estimator
- Case Studies & Demo
- Summary

IPS: Issues

$$\hat{U}_{\text{ips}}(\pi) = \frac{1}{n}\sum_i \frac{\pi(y_i|x_i)}{\pi_0(y_i|x_i)}\delta_i$$

Suppose $\delta \equiv 1$

$\hat{E}[\mathbf{Constant}] \neq \mathbf{Constant}$



Two variables in IPS

- δ
- $\pi(y|x)/\pi_0(y|x)$

Fix: Control Variates

Use correlated quantities to control $\pi(y|x)/\pi_0(y|x)$ variability

$E[s_i] = \theta$ known

Multiplicative

$$\frac{\theta}{E[s]} E\left[\frac{\pi(y|x)}{\pi_0(y|x)} \delta\right]$$

e.g. Self-Normalization

Additive

$$E\left[\frac{\pi(y|x)}{\pi_0(y|x)} \delta - s\right] + \theta$$

e.g. Doubly Robust Estimator

Self-Normalized Estimator

Use expected sample size as multiplicative control variate

$$\hat{s}(\pi) = \frac{1}{n}\sum_i \frac{\pi(y_i|x_i)}{\pi_0(y_i|x_i)}$$

$$E[\hat{s}(\pi)] = 1$$

$$\widehat{U}_{\mathrm{SNips}}(\pi) = \sum_i \frac{\pi(y_i|x_i)}{\pi_0(y_i|x_i)}\delta_i \Bigg/ \sum_i \frac{\pi(y_i|x_i)}{\pi_0(y_i|x_i)}$$

[Trotter & Tukey, 1952] [Hesterberg, 1983] [Swaminathan & Joachims, 2016]

Self-Normalization: Properties

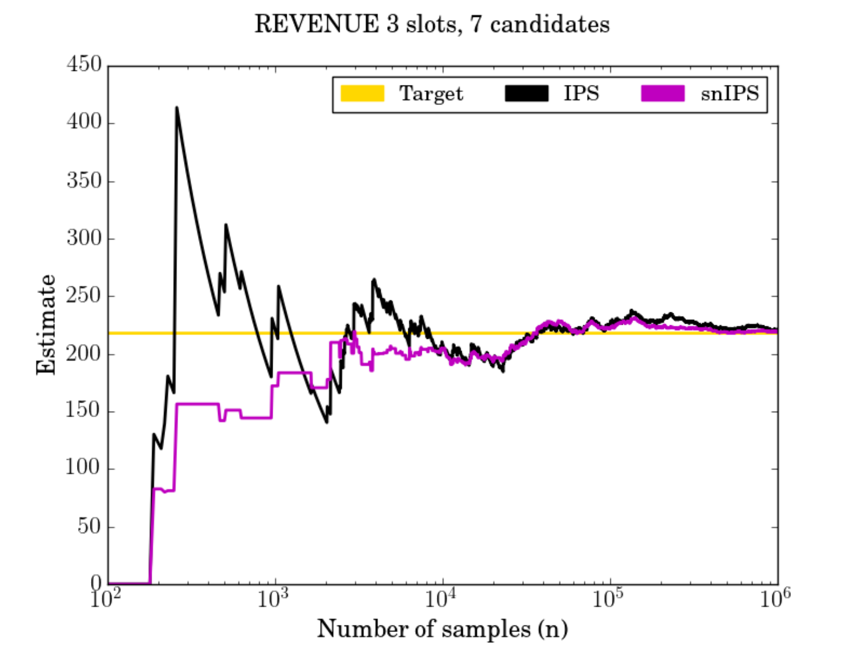
$$\hat{E}[\mathbf{Constant}] = \mathbf{Constant}$$

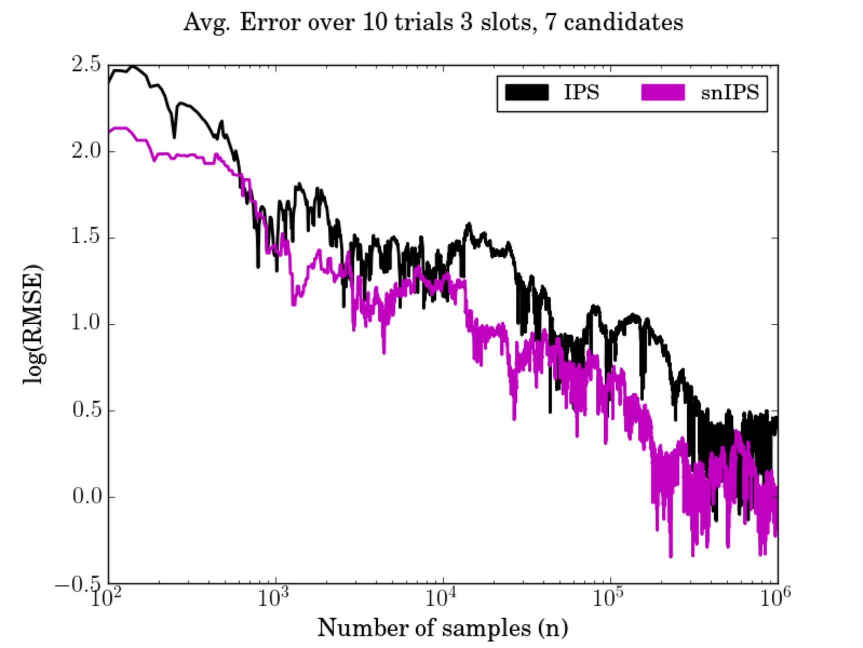
Equivariant

Asymptotically consistent $\Pr\left(\lim_{n\to\infty} \hat{U}_{\mathrm{SNips}}(\pi) = U(\pi)\right) = 1$

Small bias which decays $O(\frac{1}{n})$ while variance decays $O(\frac{1}{\sqrt{n}})$

News Recommender: Results





snIPS often achieves a better bias-variance trade-off

Doubly Robust Estimator

Reward Prediction

$$\hat{U}_{\mathrm{rp}}(\pi) = \frac{1}{n}\sum_i E_{y\sim\pi|x_i}[\hat{\delta}(x_i, y)]$$

Low variance, High bias

IPS

$$\hat{U}_{\mathrm{ips}}(\pi) = \frac{1}{n}\sum_i \frac{\pi(y_i|x_i)}{\pi_0(y_i|x_i)}\delta_i$$

High variance, No bias

$$\hat{U}_{\mathrm{dr}}(\pi) = \frac{1}{n}\sum_i \frac{\pi(y_i|x_i)}{\pi_0(y_i|x_i)}\left(\delta_i - \hat{\delta}(x_i, y_i)\right) + E_{y\sim\pi|x_i}[\hat{\delta}(x_i, y)]$$

[Langford, Li, Dudik, 2011] [Jiang & Li, 2016]

Doubly Robust: Properties

Useful when using estimated propensities $\hat{p}_i \approx \pi_0(y_i|x_i)$

Unbiased if, either

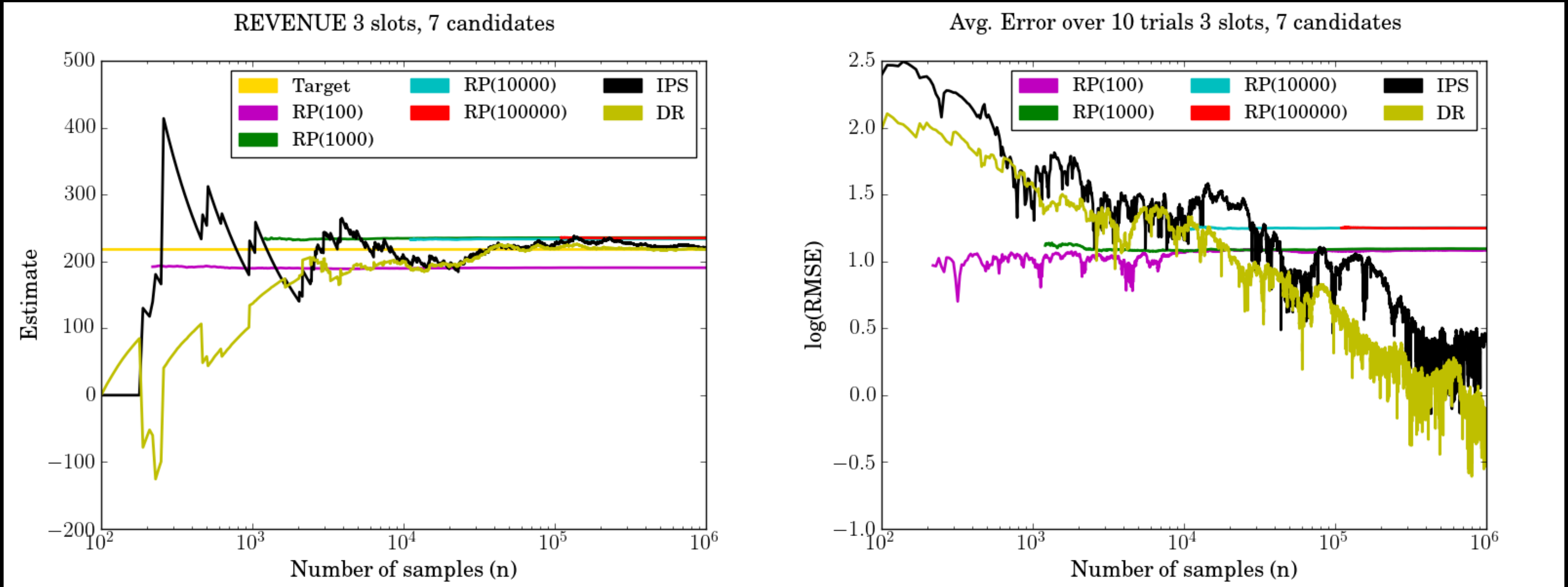
$$\hat{\delta}(x, y) = \delta(x, y)$$

Or,

$$\hat{p}_i = \pi_0(y_i|x_i)$$

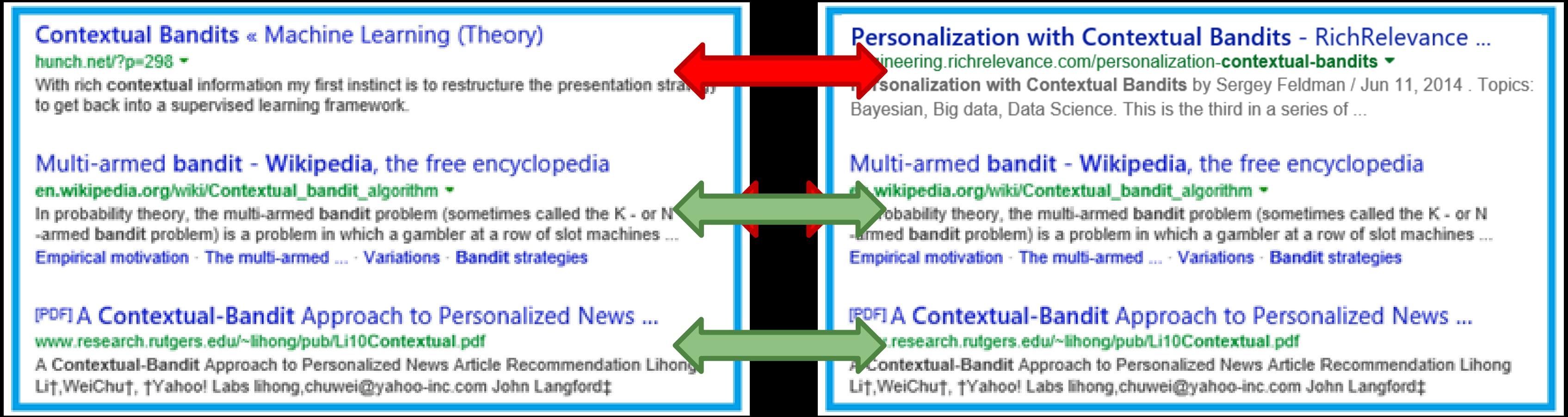
Default in Vowpal Wabbit http://hunch.net/~vw/

News Recommender: Results



DR dominates IPS even with a noisy $\hat{\delta}(x, y)$

Evaluating rankings (slates)



y_i $\pi(x_i)$

Exact match of composite actions in logs unlikely

Idea: Count per-slot matches

Slates Estimator

If π_0 samples l documents from a multinomial $\mu(d|x)$, with replacement

$$\hat{U}_{\text{slates}}(\pi) = \frac{1}{n}\sum_{i}\left(1 - l + \sum_{j=1}^{l}\frac{\mathbb{I}\{y_i[j] = \pi(x_i)[j]\}}{\mu(y_i[j]|x_i)}\right)\delta_i$$

For general π_0, need to record $\pi_0(y[j] = d, y[k] = d'|x)$

Slates Estimator: General π_0

Define:

$$\Gamma_{\pi_0(x)}[d,j\,;\,d',k] = \pi_0(y[j]=d, y[k]=d'|x)$$

$$\mathbb{1}_y[d,j] = \mathbb{I}\{y[j]=d\}$$

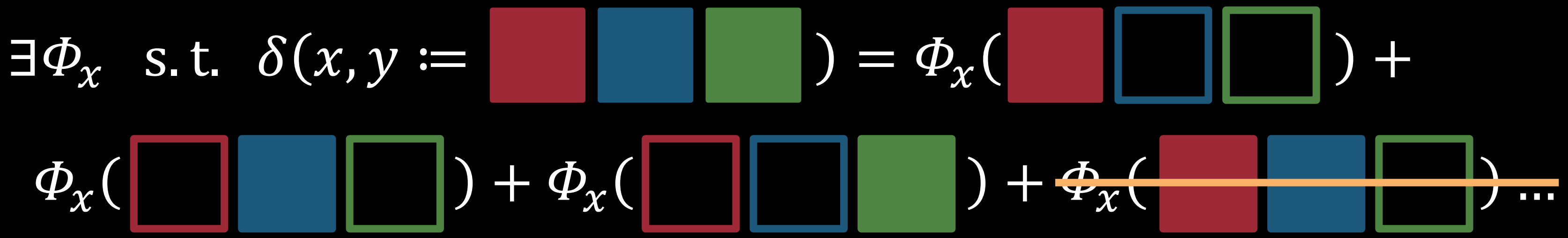
$$\hat{U}_{\text{slates}}(\pi) = \frac{1}{n}\sum_i E_{y\sim\pi|x_i}\left[\mathbb{1}_y^T\right]\Gamma_{\pi_0(x_i)}^{\dagger}\mathbb{1}_{y_i}\delta_i$$

Can also develop self-normalized/doubly robust variants

[Swaminathan et al, 2016]

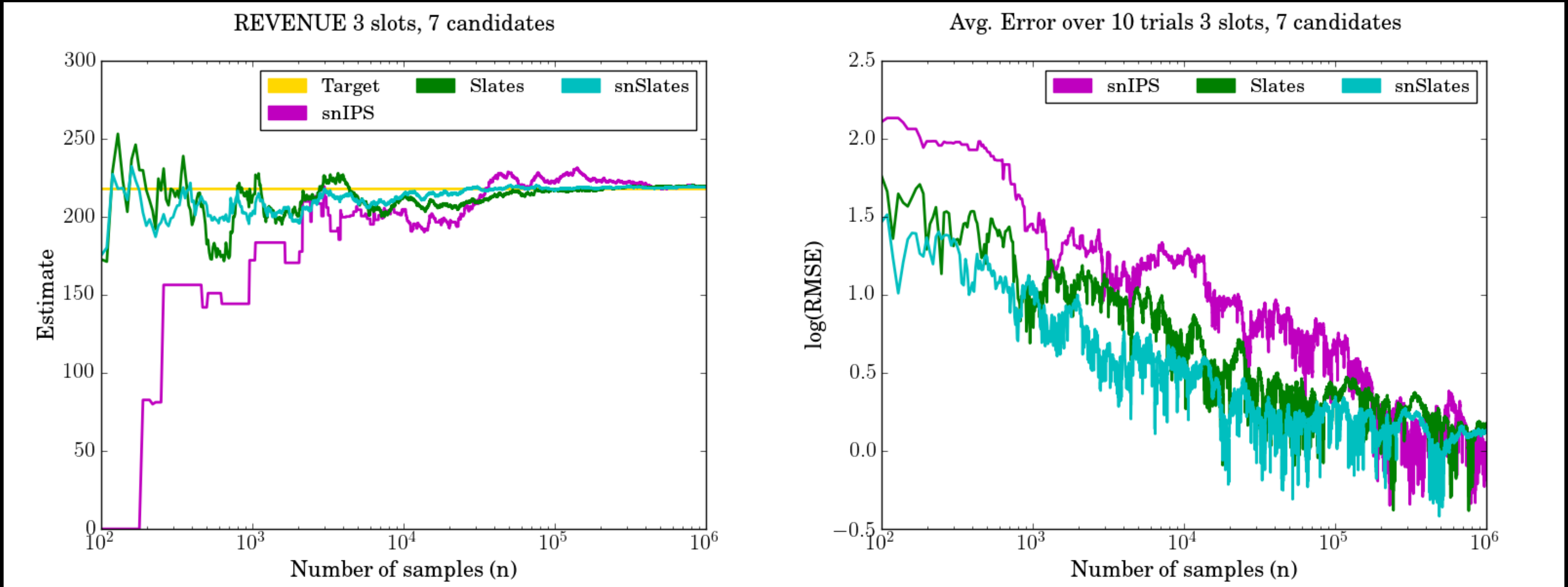
Slates Estimator: Properties

- Typically, exponentially better sample complexity than IPS
- Unbiased if reward decomposes per-slot



Can capture higher-order interactions with suitable $\Gamma_{\pi_0(x)}$

News Recommender: Results



(sn)Slates better than IPS but can have asymptotic bias

Outline

- Offline Evaluation of Online Metrics
- Counterfactual Estimation
- Advanced Estimators
- Case Studies & Demo
  1. Yahoo Front Page [Li,Chu,Langford,Wang; 2011]
  2. Bing Speller [Li,Chen,Kleban,Gupta; 2014]
  3. Search Ranking [Swaminathan et al; 2016]
- Summary

Yahoo Front Page

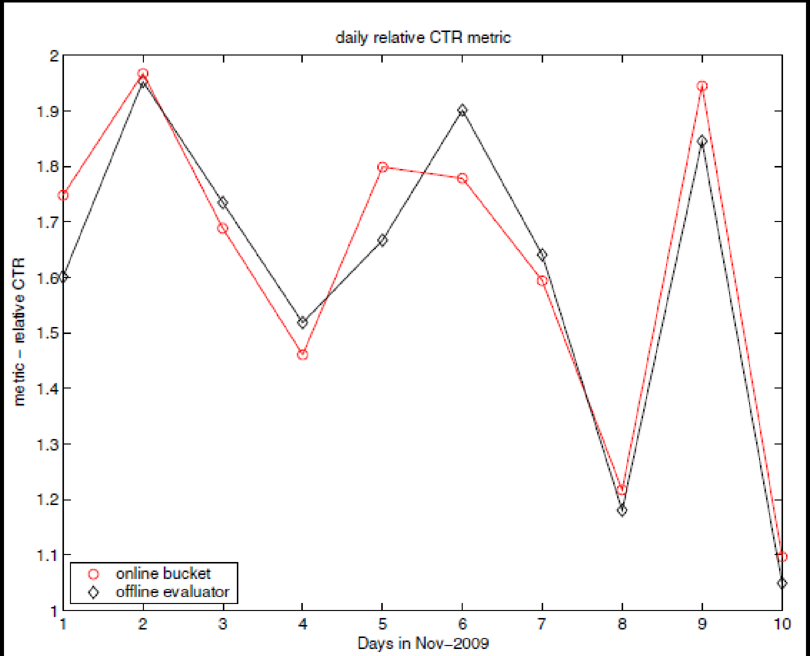
$y \sim \pi(x)$ Pick $\text{STORY} \in \{\text{F1}, \dots \text{F20}\}$ to highlight for different users

Metric δ: CTR

Logging π_0: Uniform random

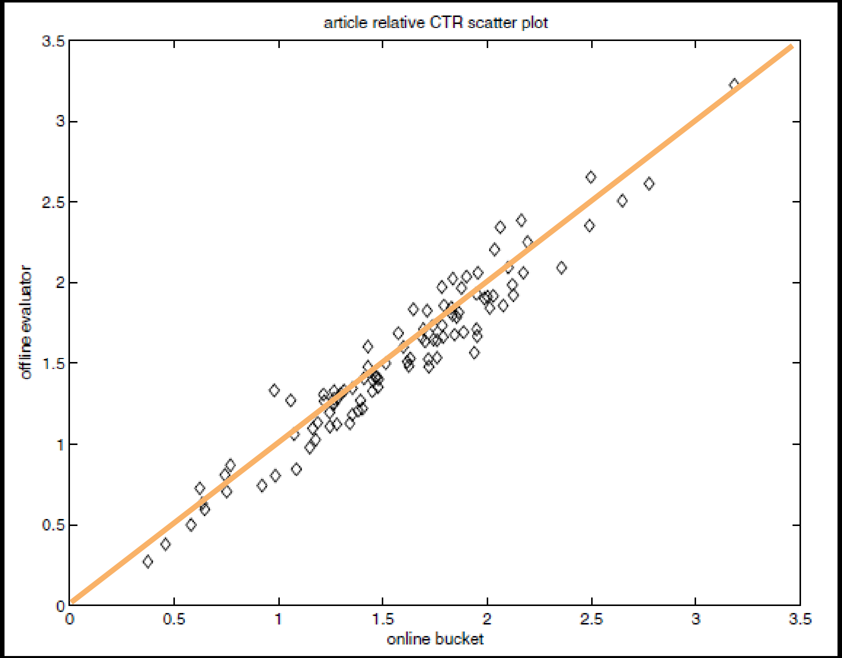
Setup: Deploy policies, check online $\leftrightarrow$ IPS correlation

Yahoo Front Page: Case Study



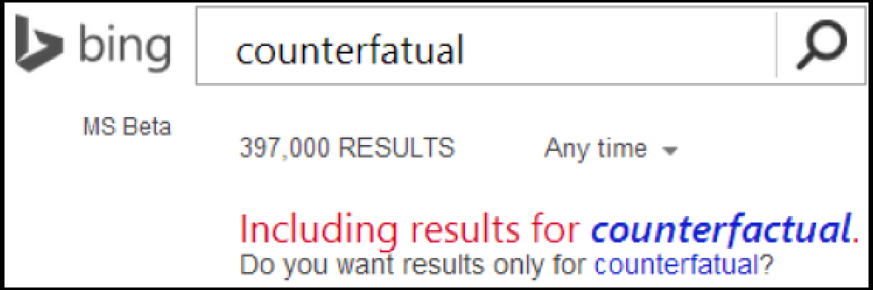
IPS is quite accurate for several (spatio-temporal) policies

Yahoo Front Page: Case Study



IPS indeed gives unbiased CTR estimates for different articles

Bing Speller



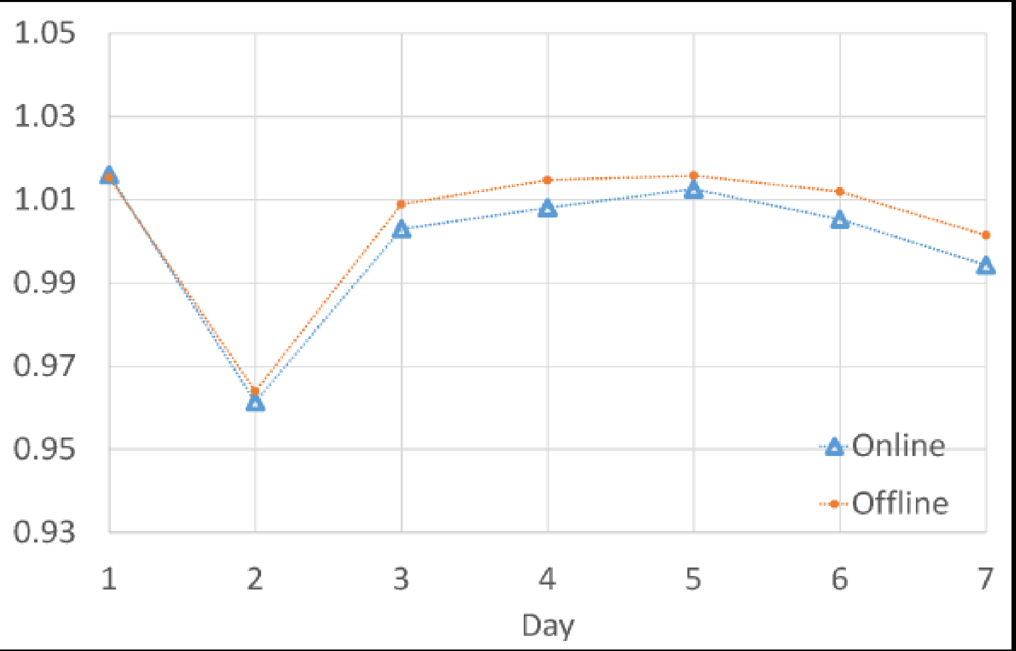
Pick (possibly many) reformulation candidates

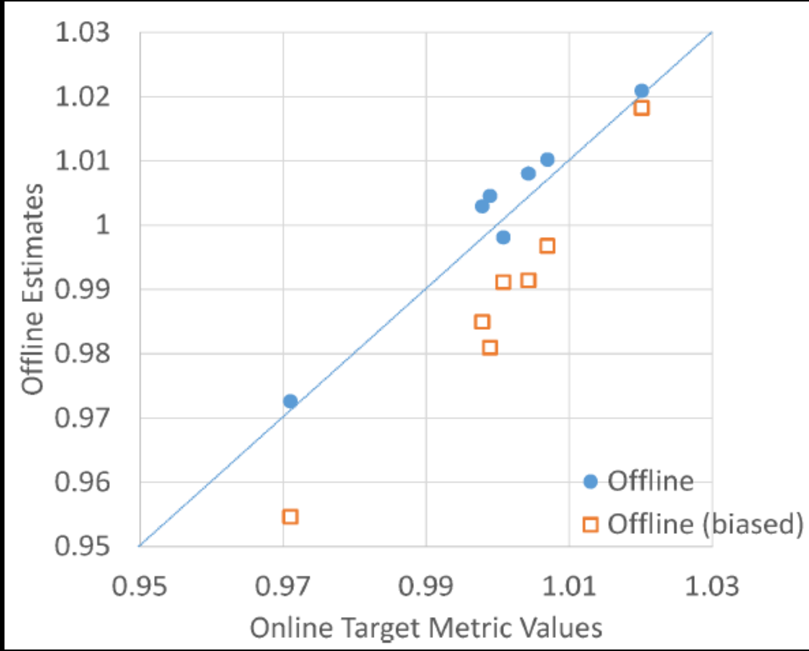
Logging π_0: Independent Bernoulli per rank

$$\Pr\left(\text{Pick } d_j\right) = 1/\left(1 + \alpha \exp\left[\beta\left\{\text{score}\left(d_j\right) - \text{score}\left(d_1\right)\right\}\right]\right)$$

Setup: Deploy policies, check online ↔ IPS correlation

Bing Speller: Case Study

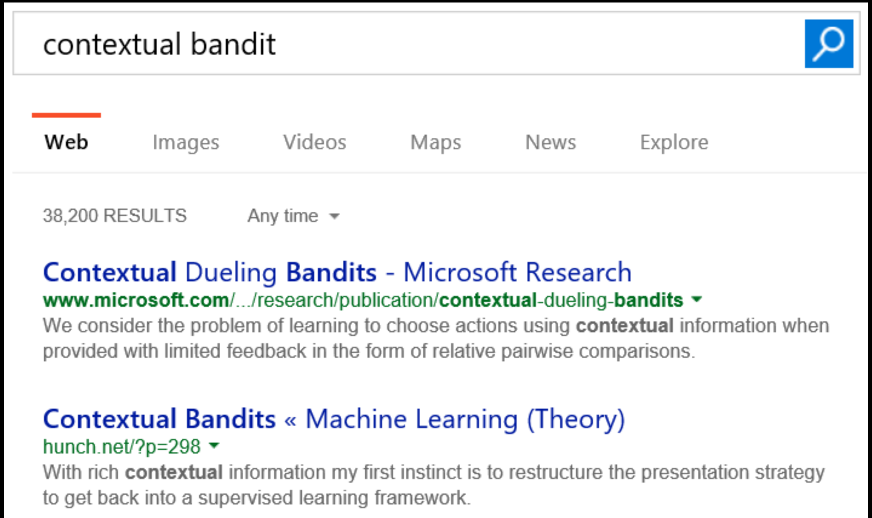




IPS reliably estimates CTR, etc. despite non-uniform logging

Search Ranking

Re-rank 5 out of 8 candidates



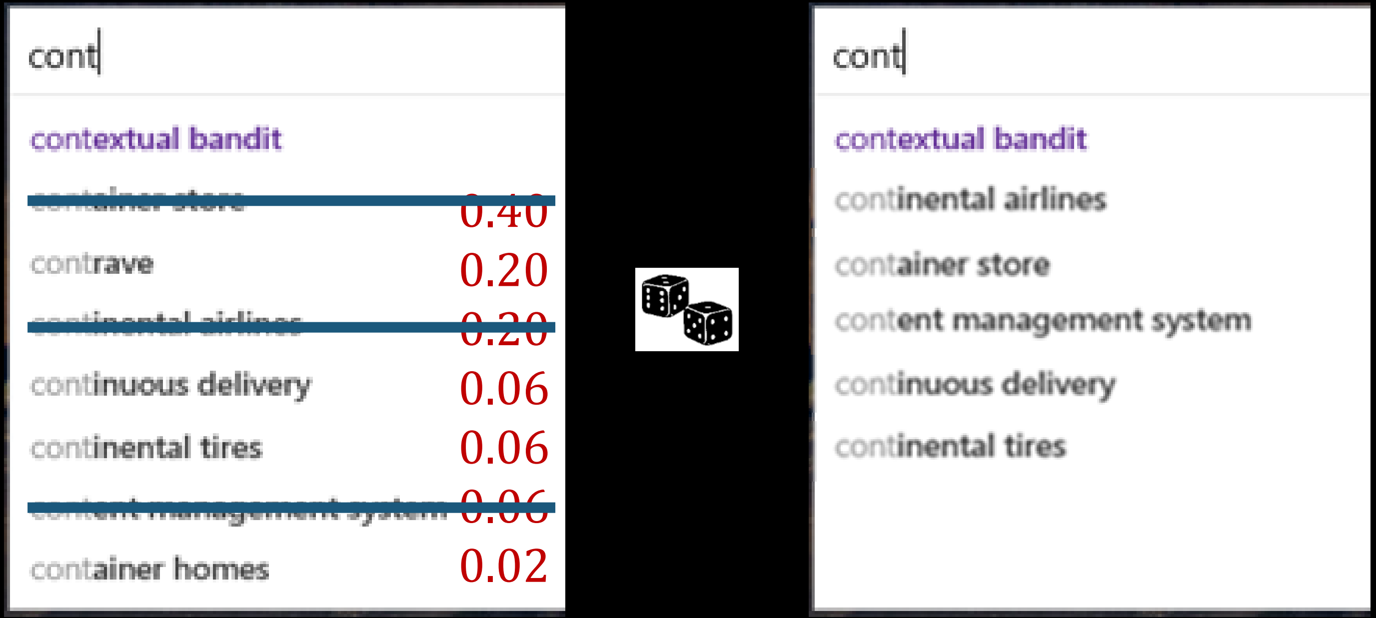
Metric δ: Time-to-success
Utility rate

Logging π_0: Bootstrap from Uniform/Plackett-Luce

Setup: Report RMSE vs. bootstrap sample size

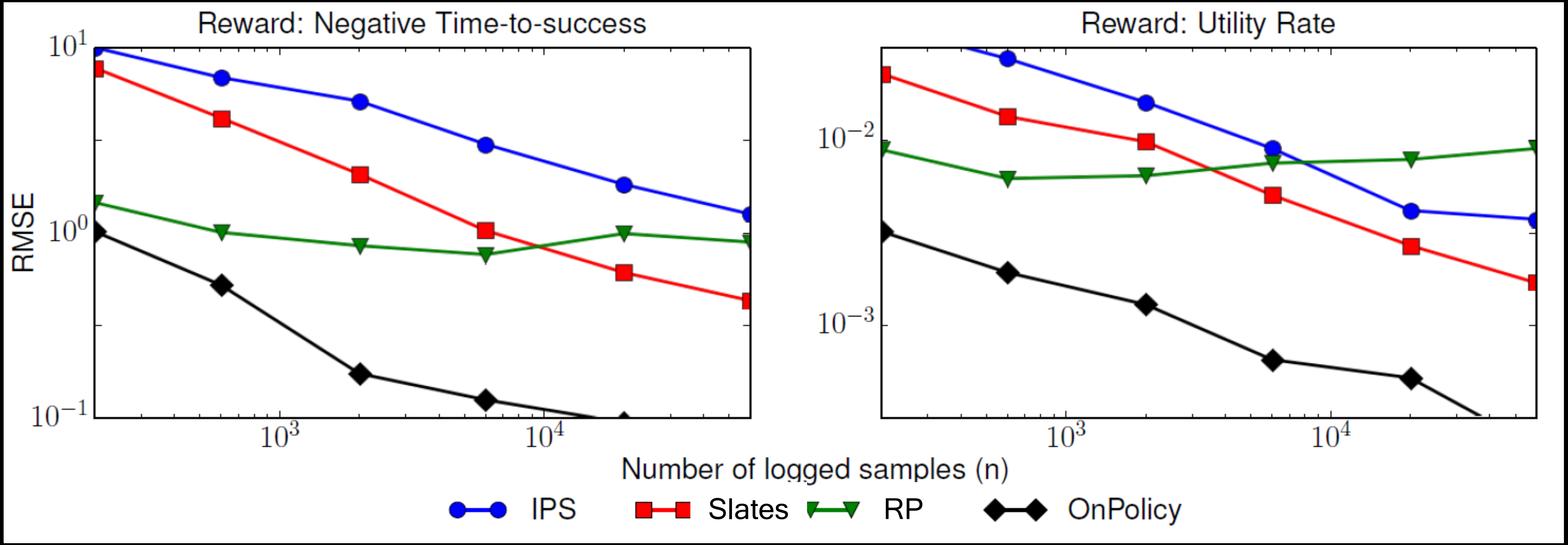
Plackett-Luce for Slates

SoftMax/Multinomial without replacement



(renormalize probabilities after each draw)

Search Ranking: Case Study



Slates estimator dominates IPS

Outline

- Offline Evaluation of Online Metrics
- Counterfactual Estimation
- Advanced Estimators
- Case Studies & Demo
- Summary

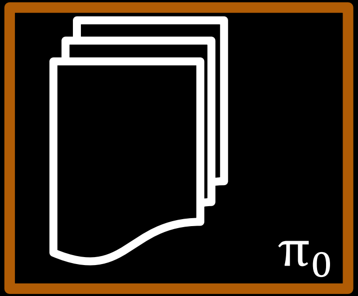
Evaluation: Key Questions

What system π_0 should we deploy?

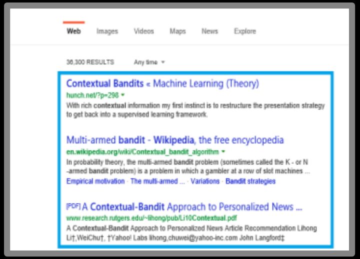
What should we record in our logs?

x

$\delta(x, y)$



$y \sim \pi_0(x)$



How can we estimate $U(\pi)$ using logged data from π_0?

What system π_0 to deploy?

To enable reliable counterfactual estimation

If possible, stochastic with logged randomization

If not, log enough to estimate propensities $\hat{p}_i$

How to explore?

Uniform? typically, bad

"Around current system"? "less risk" & "better targeted"

…

[Also see "Online contextual Bandit Algorithms"]

What should we record?

Log EVERYTHING! $\langle x_i, y_i, \delta_i, p_i \rangle$

To reliably "replay" π on logged data,

- Candidate set of actions $\{Y_i\}$
- Features for each candidate $\{f(x_i, y)\}$
- Action at the point of randomization y_i

[Bottou et al, 2013] [Li,Chen,Kleban,Gupta, 2015]

How can we evaluate $U(\pi)$?

- Offline Evaluation of Online Metrics
  - Related: Test collections for offline metrics
    [Carterette et al, 2010] [Aslam et al, 2009] [Schnabel et al, 2016b]…
- "Model the world" **Reward Prediction**
  - Related: Click models; Collaborative filtering
    [Chuklin et al, 2015] [Schnabel et al, 2016a]
- "Model the bias in data" **Off-policy estimator**
  - Randomization is essential

Summary: Off-policy Estimators

- IPS Estimator

  Simple, effective fix for non-uniform (biased) data

- Self-Normalized Estimator

- Doubly Robust Estimator

- Slates Estimator

Demo: Code Samples

Visit	http://www.cs.cornell.edu/~adith/CfactSIGIR2016/
Download	Code_Data.zip
(Recommend)	Install Anaconda-Python3; joblib

- Run experiment:

  python EvalExperiment.py

- Plot results:

  python plot_sigir.py --mode [estimate/error] --path ../../Logs/ssynth…

QUESTIONS?