DeepMind

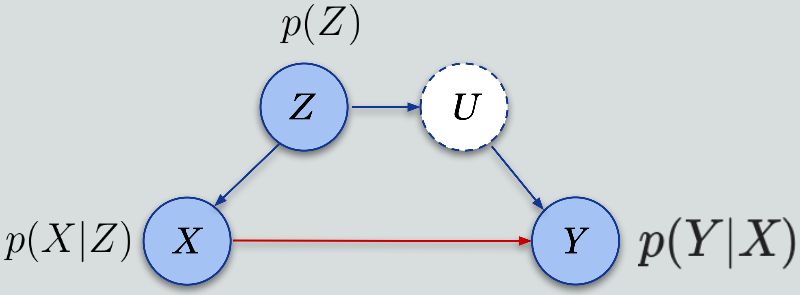
Asymptotically Best Causal Effect Identification with Multi-Armed Bandits

Alan Malek and Silvia Chiappa

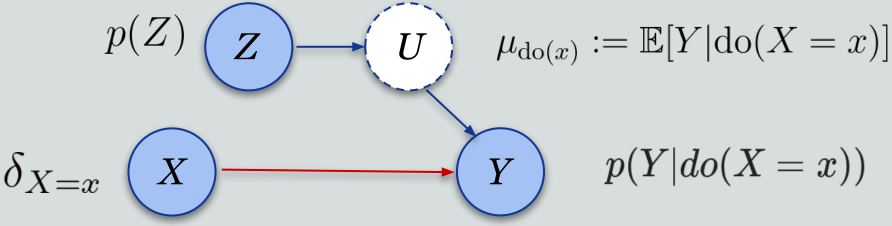
NeurIPS 2021

Identifying Causal Effects from Observation Data

- Causal graph: paths indicate causation
- Have observation data $p(X, Y, Z)$

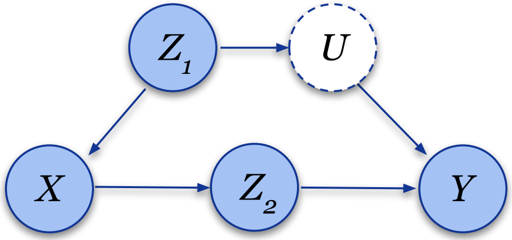


- Want the *Causal Effect* of *X* on *Y*
- Not equal to $\mathbb{E}[Y|X=1] - \mathbb{E}[Y|X=0]$
  - Confounding b.t. *X* and *Y* through *U*
- Instead, want $\tau := \mu_{do(1)} - \mu_{do(0)}$ from



Do-calculus: transforms do-expressions

- Transform into expressions learnable from observational data
- Multiple formulas may exist



Formula 1: Adjustment Criterion using Z_1

$$\mu_{do(x)} = \mathbb{E}[\mathbb{E}[Y|X=x, Z_1]]$$

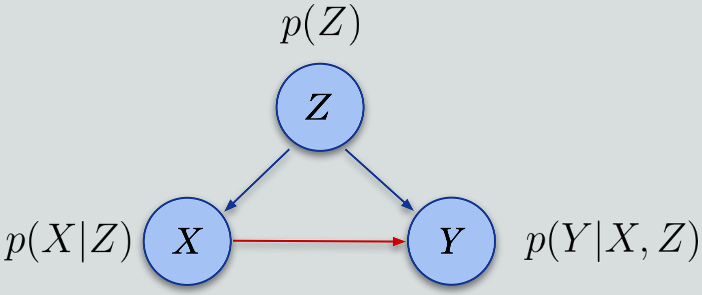
Formula 2: Frontdoor Criterion using Z_2

$$\mu_{\mathrm{do}(x)} = \sum_{z_2} p(Z_2 = z_2|X=x) \sum_{x'} p(X=x')\mu_{x'}(Z_2 = z_2)$$

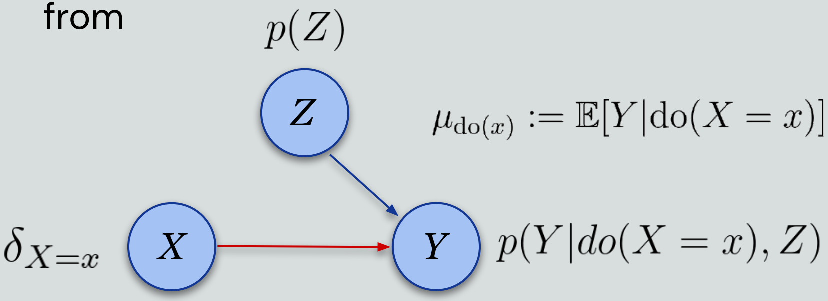
Requires nuisance function estimation!

Identifying Causal Effects from Observation Data

- Have observation data $p(X, Y, Z)$



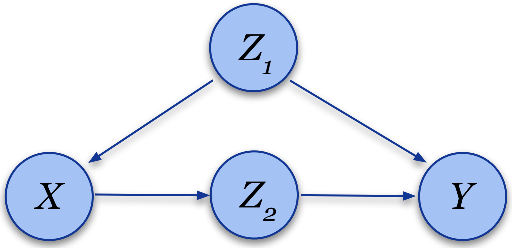
- Want the *Causal Effect* of *X* on *Y*
- Not equal to $\mathbb{E}[Y|X=1] - \mathbb{E}[Y|X=0]$
    - Correlations b.t. *Z* and *X*, *Z* and *Y*
- Instead, want $\tau := \mu_{do(1)} - \mu_{do(0)}$ from



$$\mu_{\mathrm{do}(x)} := \mathbb{E}[Y|\mathrm{do}(X=x)]$$

Do-calculus: transforms do-expressions

- Transform into expressions learnable from observational data
- Multiple formulas may exist



Formula 1: Adjustment Criterion using Z_1

$$\mu_{do(x)} = \mathbb{E}[\mathbb{E}[Y|X=x, Z_1]]$$

Formula 2: Frontdoor Criterion using Z_2

$$\mu_{\mathrm{do}(x)} = \sum_{z_2} p(Z_2 = z_2|X=x) \sum_{x'} p(X=x')\mu_{x'}(Z_2 = z_2)$$

Requires nuisance function estimation!

Estimation with nuisance functions and the AIPW estimator

For estimators using the back-door, the Augmented Inverse Propensity Weight estimator is optimal

$$\hat{\tau}(\mathcal{D}) = \mathbb{E}_n\left[\left(\frac{X}{\hat{e}(\mathcal{Z})}(Y - \hat{\mu}_1(\mathcal{Z})) + \hat{\mu}_1(\mathcal{Z})\right) - \left(\frac{1-X}{1-\hat{e}(\mathcal{Z})}(Y - \hat{\mu}_0(\mathcal{Z})) + \hat{\mu}_0(\mathcal{Z})\right)\right]$$

- Need to estimate *nuisance functions* $\eta = (e(\cdot), \mu_x(\cdot))$
  - The propensity score $e(z) = P(X = 1 | Z = z)$
  - The conditional response $\mu_x(z) = \mathbb{E}[Y | X = x, Z = z]$
- Generally estimate $\hat{\eta}$ on the first fold of data, then estimate $\hat{\tau} \in \{\tau : \mathbb{E}_n[\phi(W, \hat{\eta}, \tau)] = 0\}$
- Worry: a slow $O\left(n^{-1/4}\right)$ estimation rate of $\hat{\eta}$ forces a “slow” rate for $\hat{\tau}$
- Double/Debiased machine learning [Chernozhukov, et al. 2018] shows that we can recover fast $O\left(n^{-1/2}\right)$ rates for $\hat{\tau}$ under Neyman orthogonality

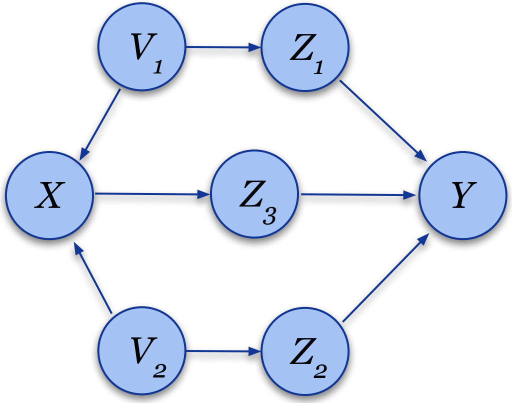
Selecting an Estimator

- Over-identified: have estimators $\hat{\tau}_k \in \{\hat{\tau}_1, \ldots, \hat{\tau}_K\}$; each estimator
  - is asymptotically linear: $\sqrt{n}(\hat{\tau}_k(\mathcal{D}_n) - \tau_0) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \underbrace{\phi_k(W_i, \eta_{k,0}, \tau_0)}_{=:IF} + o_p(1)$
  - needs covariates $\mathcal{Z}_k$ with cost of collection c_k
  - has influence function ϕ_k and nuisance function η_k
- Asymptotic variance $\sigma_k^2 = \mathbb{E}[\phi_k^2(W, \eta_{k,0}, \tau_0)]$
  - Such estimators exist for *any* identification formula [Jung, Tian, Bareinboim 2021]
- Can assume an *uncentered influence function*, $\phi_k(W, \eta_{k,0}, \tau_0) = \psi_k(W, \eta_{k,0}) - \tau_0$
- Goal: identify the best estimator $k^* = \arg\min_k c_k \sigma_k^2$ to use for a large, observational study
  - e.g. choose lab tests/sensors/survey questions
  - Trade-off cost with statistical efficiency

Sequential Decision Problem

- Investigator allowed to dynamically sample data
  - Can update what covariates are observed
- Model as a best-arm-identification bandit problem
  - Each estimator is an arm
  - Target: asymptotic variance, not mean

Bandit Model



Experimental Protocol

Given: Estimators $\hat{\tau}_1, \ldots, \hat{\tau}_K$, costs $c_1, \ldots c_K$, $\epsilon > 0$, $\delta > 0$

for $t = 1, 2, \ldots$: **do**

  Choose $k_t \in [K]$

  Obtain observation $w_{k_t}^t$

  Choose whether to stop sampling

end

Return: (ϵ, δ)-PAC index $\hat{k}$ s.t.

$$\mathbb{P}\left(c_{\hat{k}}\sigma_{\hat{k}}^2 \geq \min_k c_k \sigma_k^2 + \epsilon\right) \leq \delta$$

Estimating the Asymptotic Variance

- Our goal: estimate $\sigma_k^2 = \mathbb{E}[(\psi_k(W,\eta_k) - \tau)^2]$, derive a *finite-sample* confidence set
- Our estimator $\hat{\sigma}_k^2$ for data $\mathcal{D}_n$ (inspired by [Chernozhukov et al., 2016])
  - Randomly split $\mathcal{D}_n$ into two folds, $\mathcal{D}_n^\eta, \mathcal{D}_n^\sigma$
  - Fit the nuisance function $\hat{\eta}_k(\mathcal{D}_n^\eta)$
  - Fit $\hat{\sigma}_k^2$ with an empirical variance: $\hat{\sigma}_k^2(\mathcal{D}_n) = \mathrm{Var}_{\mathcal{D}_n^\sigma}\left[\psi_k(W, \hat{\eta}_k(\mathcal{D}_k^\eta))\right]$
  - Do not have Neyman orthogonality

Confidence Sequence for the Asymptotic Variance

Theorem 1. *Let $\alpha > 0$. Assume that ψ is L-Lipschitz, and let $\tilde{\tau}$ be an upper bound on $|\tau|$. Let $\mathcal{D}_1 \subseteq \mathcal{D}_2 \subseteq \ldots$ be a sequence of datasets with $\mathcal{D}_n = \mathcal{D}_n^\eta \cup \mathcal{D}_n^\sigma$. Assume:*

1. $P\left(\forall n \geq 1 : \left|\mathbb{E}_{\mathcal{D}_n^\sigma}[\psi(W,\eta_0)] - \mathbb{E}[\psi(W,\eta_0)]\right| \leq u_n^1\right) \geq 1 - \alpha/3,$ ← $O\left(n^{-1/2}\right)$
2. $P\left(\forall n \geq 1 : \left|\mathbb{E}_{\mathcal{D}_n^\sigma}[\psi(W,\eta_0)]^2 - \mathbb{E}[\psi(W,\eta_0)]^2\right| \leq u_n^2\right) \geq 1 - \alpha/3$, *and* ← $O\left(n^{-1/2}\right)$
3. $P\left(\forall n \geq 1 : \|\hat{\eta}(\mathcal{D}_n^\eta) - \eta_0\| \leq u_n^\eta\right) \geq 1 - \alpha/3,$ ← $O\left(n^{-1/4}\right)$

then $\mathbb{P}\left(\forall n \geq 1 : \left|\hat{\sigma}^2(\mathcal{D}_n) - \sigma^2\right| \leq 2L^2(u_n^\eta)^2 + \quad O\left(n^{-1/2}\right) \quad\right) \geq 1 - \alpha.$

$2L^2(u_n^\eta)^2$ ← $O\left(n^{-1/2}\right)$

Corollary 1. *Let $\alpha \in (0,1)$ and assume the same setting as Theorem 1, and additionally that $\psi(W,\eta)$ is λ sub-Gaussian. Then, with a certain choice of parameters, the confidence sequence of Lemma 4 guarantees that, for $\lambda' = \lambda \vee 8\lambda^2$, any $m > 0$ and for any $n \geq (91\lambda'(\log(\lambda' n/m) + \log(1/\alpha))) \vee (m/\lambda')$,*

$$\mathbb{P}\left(\exists n \geq 1 : \left|\hat{\sigma}^2(\mathcal{D}_n) - \sigma^2\right| \geq 2L^2(u_n^\eta)^2 + (3 + 6\tilde{\tau})\sqrt{\frac{\lambda'}{n}\left(\frac{1}{2}\log\left(\frac{\lambda' n}{m}\right) + \log\frac{2}{\alpha}\right)}\right) \leq \alpha.$$

Bandit Algorithm 1: LUCB

Algorithm 1 CS-LUCB

Input $\epsilon > 0$, $\delta > 0$, $\Delta_n > 1$,$\tilde{\tau} > 0$

$\{\psi_k, \hat{\eta}_k, u_k, c_k : k \in [K]\}$

for k=*1*,..., K **do**

  Obtain Δ_n new samples $\mathcal{D}$

  Add half of $\mathcal{D}$ to $\mathcal{D}_k^\eta$ and half to $\mathcal{D}_k^\sigma$

  $\hat{\sigma}_k^2, \beta_k \leftarrow \text{CSUpdate}(u_k, \hat{\eta}_k, \psi_k, L, \tilde{\tau}, \mathcal{D}_k^\eta, \mathcal{D}_k^\sigma)$

end

for $t = 1, 2, \ldots$ **do**

  $l_t \leftarrow \arg\min_{k \in [K]} c_k \hat{\sigma}_k^2$

  $u_t \leftarrow \arg\min_{k \neq l_t} c_k (\hat{\sigma}_k^2 - \beta_k)$

  if $c_{l_t} \left(\hat{\sigma}_{l_t}^2 + \beta_{l_t}\right) \leq c_{u_t} \left(\hat{\sigma}_{u_t}^2 - \beta_{u_t}\right) - \epsilon$ **then**

    Return $\hat{k} = l_t$

  end

  for $k \in u_t, l_t$ **do**

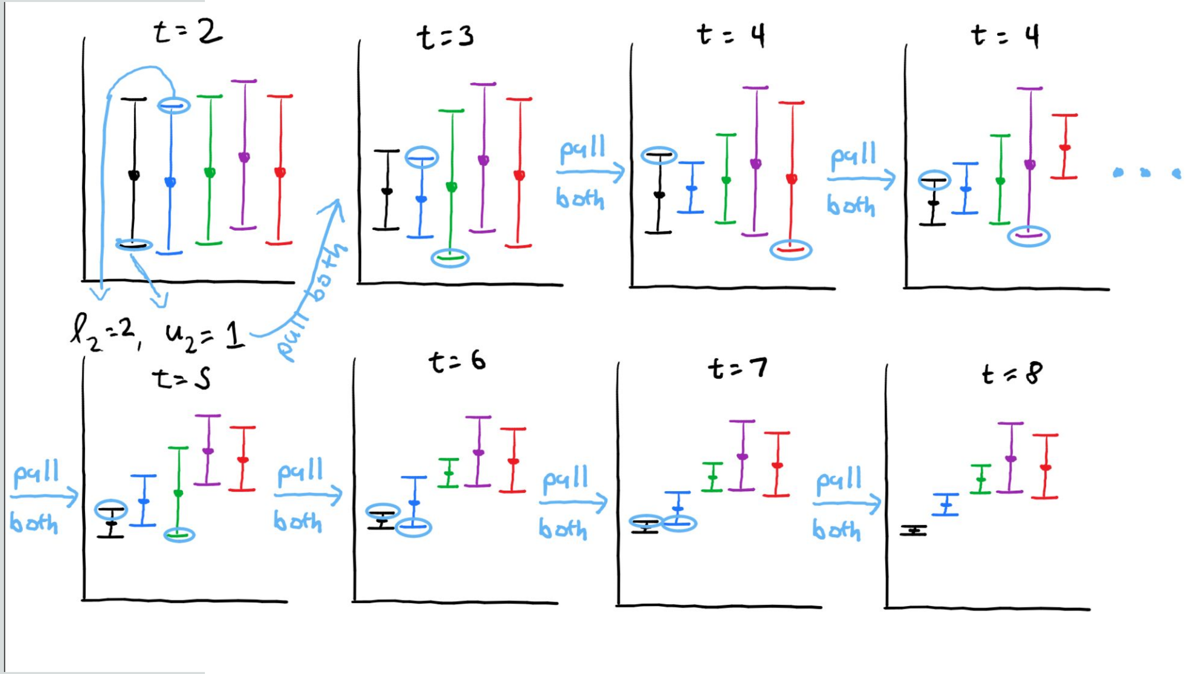
    Obtain Δ_n new samples $\mathcal{D}$

    Add half of $\mathcal{D}$ to $\mathcal{D}_k^\eta$ and half to $\mathcal{D}_k^\sigma$

    $\hat{\sigma}_k^2, \beta_k \leftarrow \text{CSUpdate}(u_k, \hat{\eta}_k, \psi_k, L, \tilde{\tau}, \mathcal{D}_k^\eta, \mathcal{D}_k^\sigma)$

  end

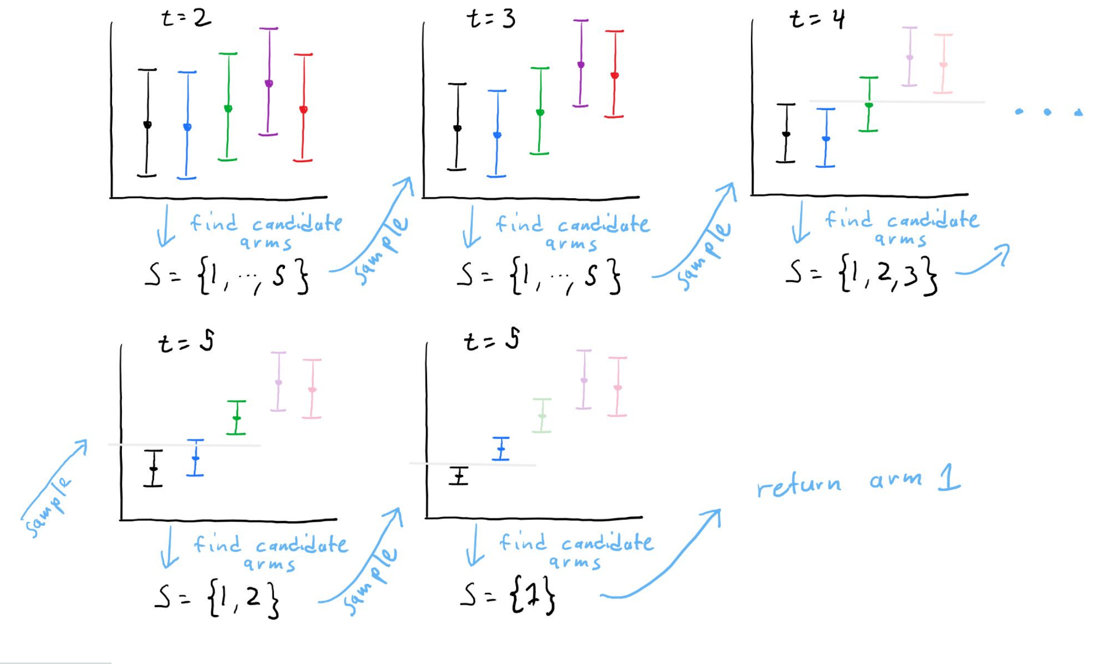
end



Bandit Algorithm 2: Successive Elimination

Algorithm 2 CS-SE

Input $\delta > 0$, $\epsilon > 0$, $\Delta_n > 1$, $\tilde{\tau} > 0$
$\{\psi_k, \hat{\eta}_k, u_{k,n}, c_k, : k \in [K]\}$
$S \leftarrow [K], \mathcal{D}_k^{\eta}, \mathcal{D}_k^{\sigma} \leftarrow \emptyset \; \forall k \in [K]$
while $|S| > 1$ **do**
  for $k \in S$ **do**
    Obtain Δ_n new samples $\mathcal{D}_k^{\Delta}$
    Add half of $\mathcal{D}_k^{\Delta}$ to $\mathcal{D}_k^{\eta}$ and half to $\mathcal{D}_k^{\sigma}$
    $\hat{\sigma}_k^2, \beta_k \leftarrow \text{CSUpdate}(u_k, \hat{\eta}_k, \psi_k, L, \tilde{\tau}, \mathcal{D}_k^{\eta}, \mathcal{D}_k^{\sigma})$
  end
  $k^* \leftarrow \arg\min_{k \in S} c_k \hat{\sigma}_k^2$
  $R \leftarrow \{k \in S : c_{k^*}(\hat{\sigma}_{k^*}^2 + \beta_{k^*}) \leq c_k(\hat{\sigma}_k^2 - \beta_k)\}$
  $S \leftarrow S \setminus R$
  if $\beta_k \leq \frac{\epsilon}{2}$ *for all* $k \in \S$ **then**
    $S \leftarrow \arg\min_{k \in S} c_k \hat{\sigma}_k^2$
  end
end
Return $\hat{k} = S$



Bandit Algorithms: Sample complexity Upper Bound

Theorem 3. *Assume that the conditions of Theorem 1 hold and that* $u^{\theta}_{k,n} \to 0$ *for* $\theta \in \{\eta, (\psi, 1), (\psi, 2)\}$. *Then both CS-LUCB and CS-SE return an* (ϵ, δ)*-PAC estimator.*

If we have a deterministic upper bound $B_k(n, \delta)$ *with, for all* $\delta > 0$, $P(\beta_k(n) \leq B_k(n, \delta)) \geq 1 - \delta$, *then the number of samples required for either algorithm to terminate in at most* $\sum_{k \in [K]} \min \left\{ n : B_k(n, \delta/K) \leq \frac{\Delta_k}{4} \vee \frac{\epsilon}{2} \right\}$ *samples.*

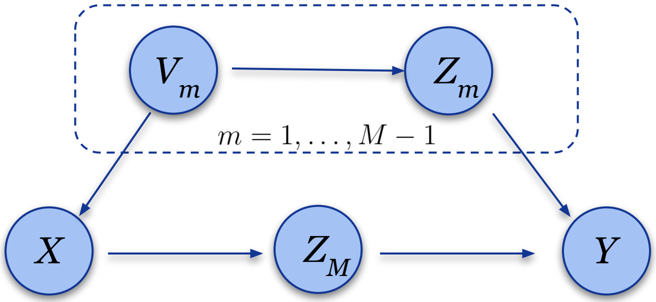
If, additionally, there exists constants ν_η, ν_1, *and* ν_2 *such that* $u^{\theta}_{k,n} \leq \mathcal{O}(n^{-\nu_\theta} \log(nK/\delta))$ *for all* $\theta \in \{\eta, (\psi, 1), (\psi, 2)\}$ *and all* $k \in [K]$, *the sample complexity is*

$$\mathcal{O}\left(\sum_{k=1}^{K} (\Delta_k \vee \epsilon)^{-1/\nu} \left(\log \frac{K}{\delta (\Delta_k \vee \epsilon)^{1/\nu}} \right)^{1/\nu} \right),$$

where $\nu = \min\{2\nu_\eta, \nu_{\psi,1}, \nu_{\psi,2}\}$ *with probability at least* $1 - \delta$. *In particular, we recover the sample complexity results (up to log factors) of [16, 4] under the mild condition of* $\nu_\eta \geq 1/4$.

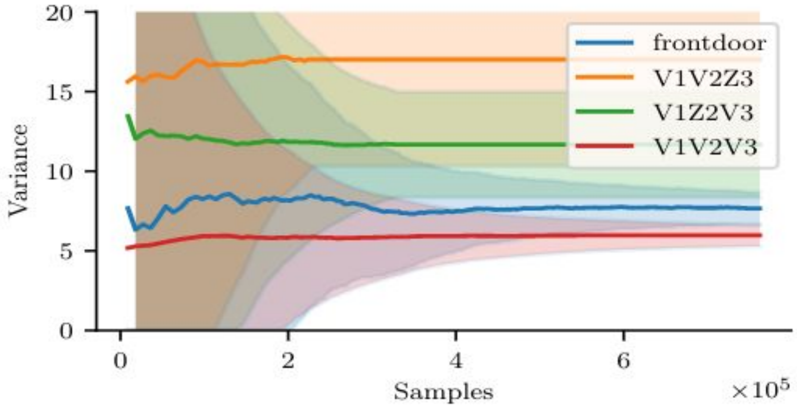
Experiments

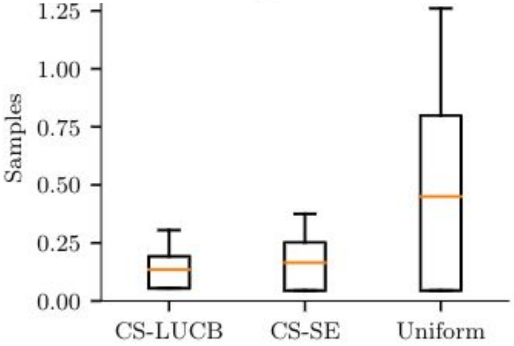
$$V_m \sim \mathcal{N}(0, I_2) \qquad Z_m = A_m V_m + \epsilon$$



$$\mu_x = \sigma\left(\sum_m C_m V_m\right) \qquad Y = \sum_m B_m Z_m + \epsilon$$

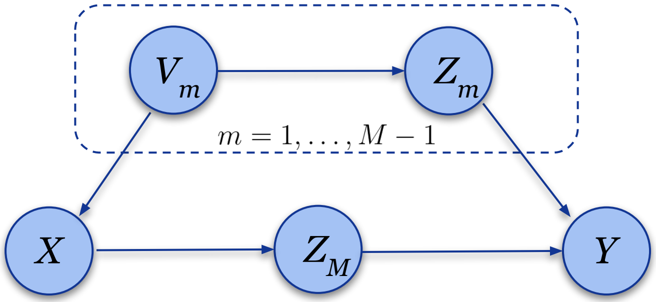
- There are 2^{M-1} estimators for adjustment criteria
- There is an estimator using the frontdoor criterion with Z_M





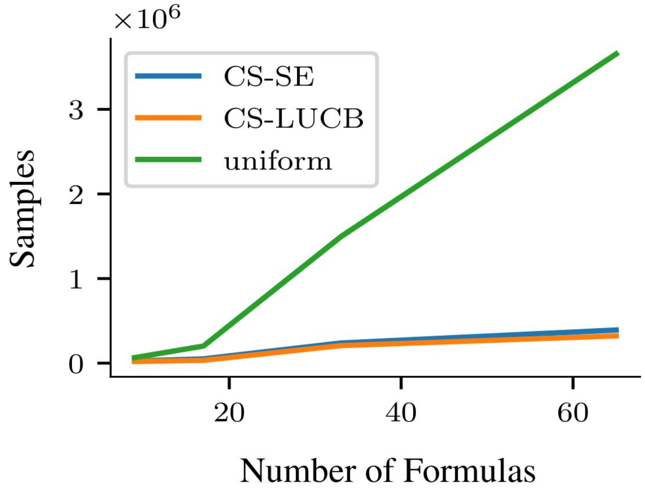
Non–Linear Experiments

$$V_m \sim \mathcal{N}(0, I_2) \qquad Z_m = f_m(V_m) + \epsilon$$



$$\mathbb{E}[X] = \sigma\left(\sum_m h_m(V_m)\right) \qquad Y = \sum_m g_m(Z_m) + \epsilon$$

- f_m, g_m, h_m are sampled from a Gaussian process prior
- Noise is Gaussian



Thank you!