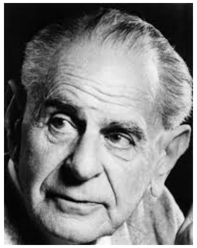
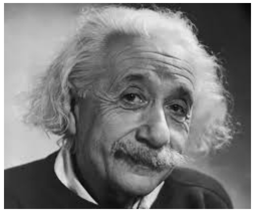
# Causality

*We call to mind the constant conjunction (of flame and heat) in all past instances. “Without any farther ceremony, we call the one cause and the other effect, and infer the existence of the one from that of the other.”* **David Hume 1739**

*“Beyond such discarded fundamentals as ‘matter’ and ‘force lies still another fetish amidst the inscrutable arcana of modern science, namely, the category of cause and effect.” … “The ultimate scientific statement of description of the relation between two things can always be thrown back upon a contingency table.”* **Karl Pearson 1911**



*“The belief in causality is metaphysical. It is nothing but a typical metaphysical hypostatization of a well-justified methodological rule- the scientist's decision never to abandon his search for laws.”* **Karl Popper 1934**

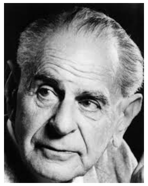


*“Development of Western science is based on two great achievements: the invention of the formal logical system (in Euclidean geometry) by the Greek philosophers, and the discovery of the possibility to find out causal relationships by systematic experiment.”* **Albert Einstein 1953**

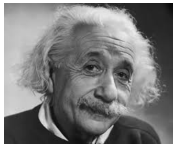
## Causality

Causality is an illusion arising from repeated temporal conjunction (Hume 1739)

Causality is outdated, correlation is all that is required (Pearson 1911)



Looking for scientific laws works. Causality is a metaphysical abstraction. (Popper 1934)



Learning causal relationships is central to science and we do it via systematic experiment (Einstein 1953)

❖ **A causal model is one that predicts the outcome of intervention in a system**

# Correlation is not causation

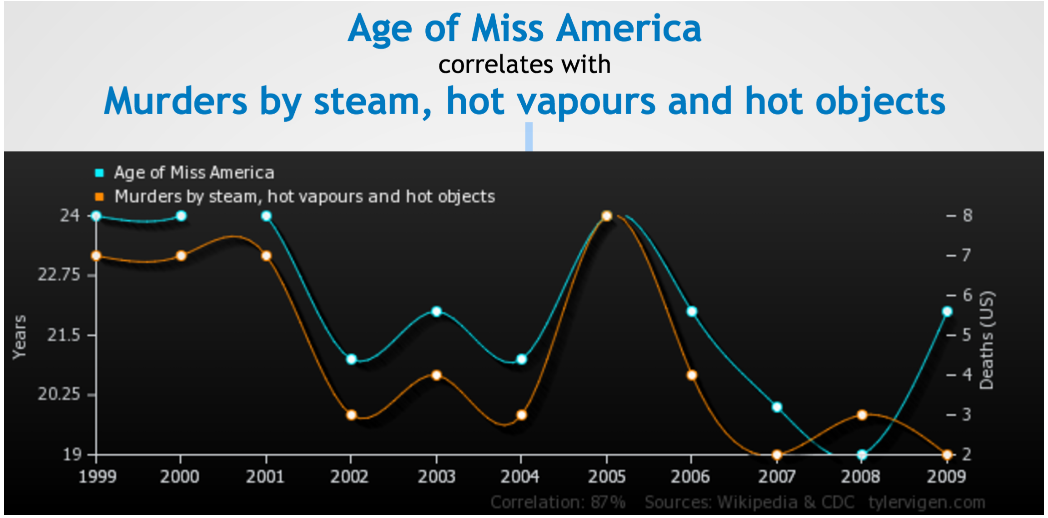
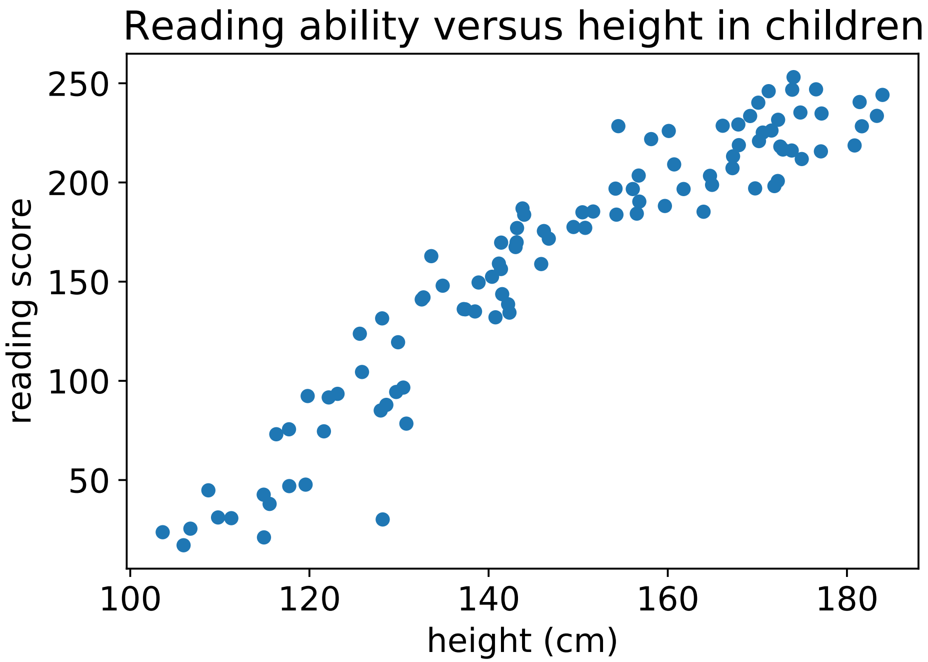



Image source: www.tylervigen.com/



# When do we care about causality?

- Forecasting the weather
- Image classification
- Predicting which patients are at risk of death from pneumonia

# When do we care about causality?

- Forecasting the weather
- Image classification
- Predicting which patients are at risk of death from pneumonia

**The real question is "when don't we care?"**

# Trajectory of a PhD

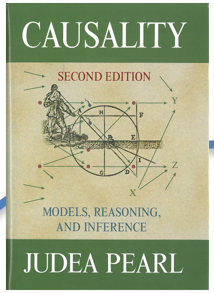

ambiata

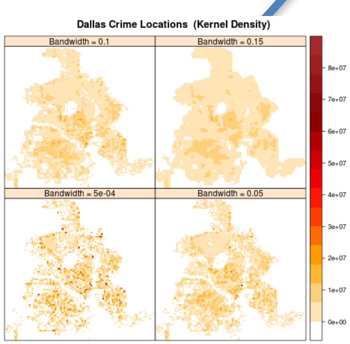
**Causal Bandits: Learning Good Interventions via Causal Inference**

Finnian Lattimore
Australian National University and Data61/NICTA
finn.lattimore@gmail.com

Tor Lattimore
Indiana University, Bloomington
tor.lattimore@gmail.com

Mark D. Reid
Australian National University and Data61/NICTA
mark.reid@anu.edu.au



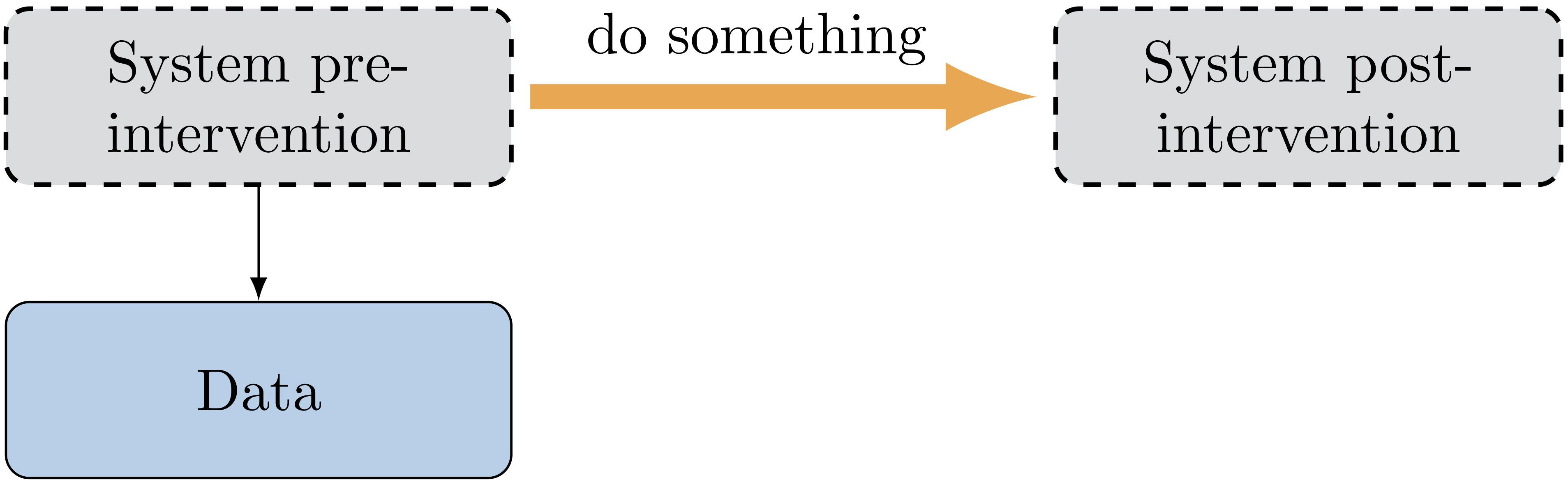

# Contributions

- A synthesis of the observational and interventional approaches to learning to intervene that combines causal graphical models with multi-armed bandits.
- Clarifying the causal assumptions underlying previous bandit formulations.
- Demonstrating the role a graphical framework for causality can play in Bayesian inference.

# Overview – Causal inference in ML

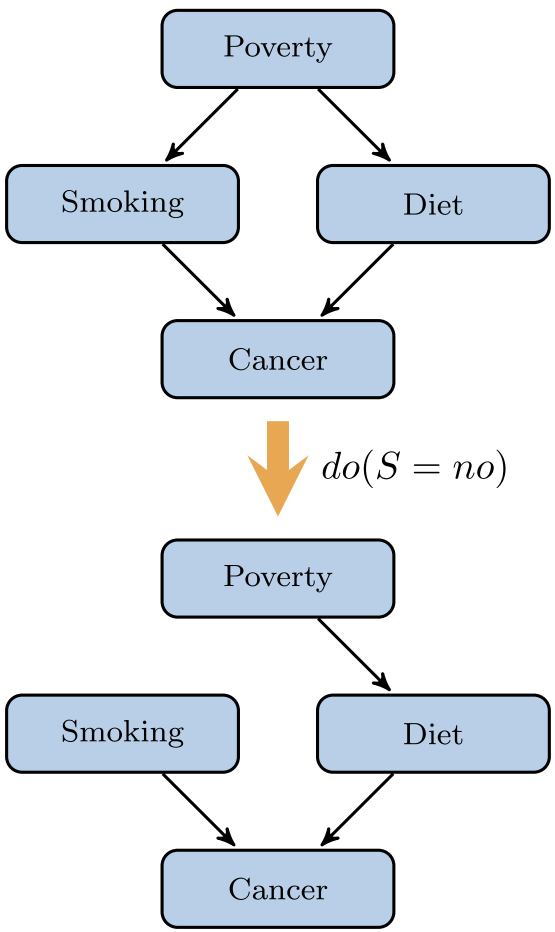
- What is causality?
  - Predicting the outcome of interventions
- Why should we care in machine learning?
  - Because we want to act in response to our models
- Observational causal inference
- Interventional causal inference
- A unified approach - Causal Bandits

# Observational causal inference



# Causal Bayesian Networks

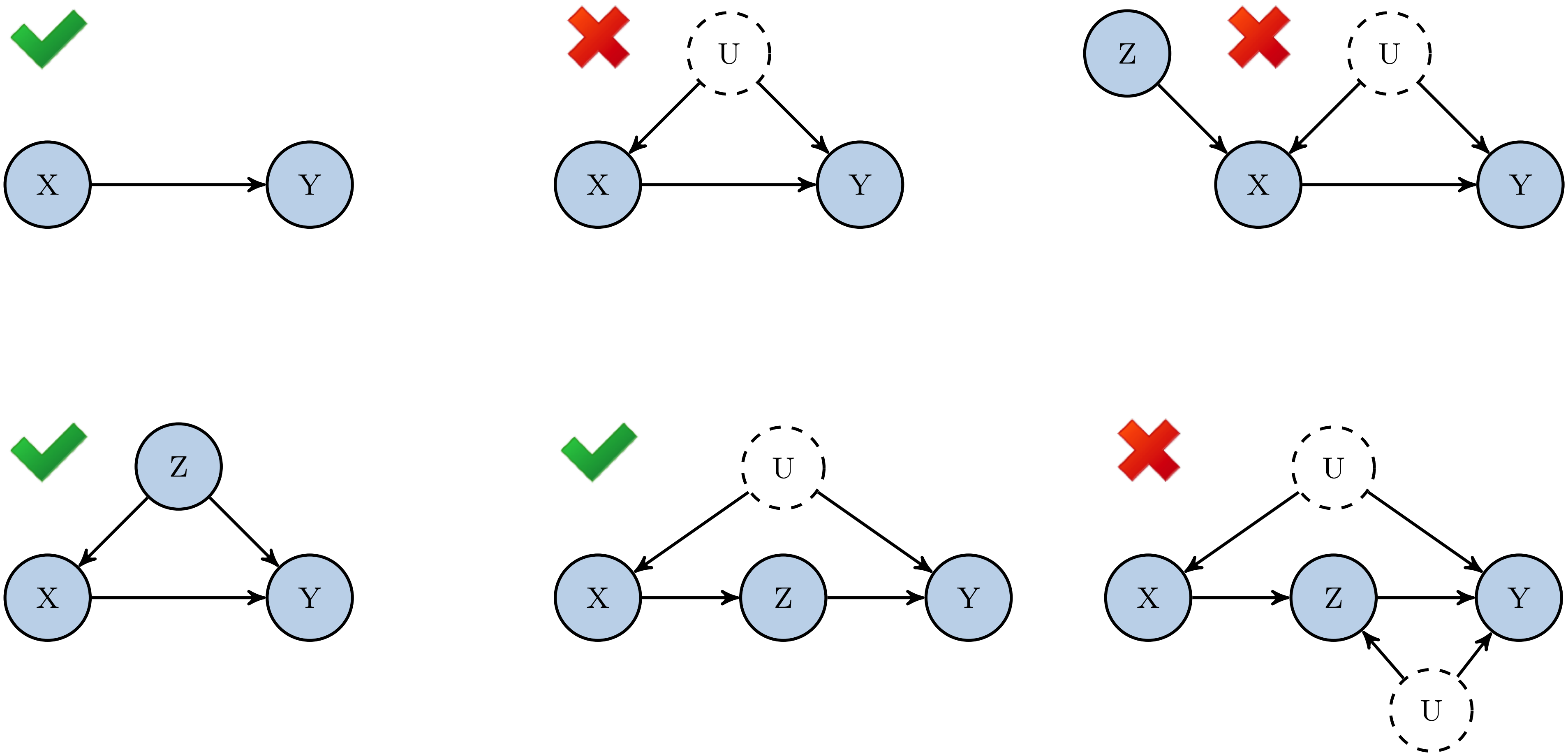
$$P(P, D, C, S) = P(P)P(S|P)P(D|P)P(C|S, D)$$



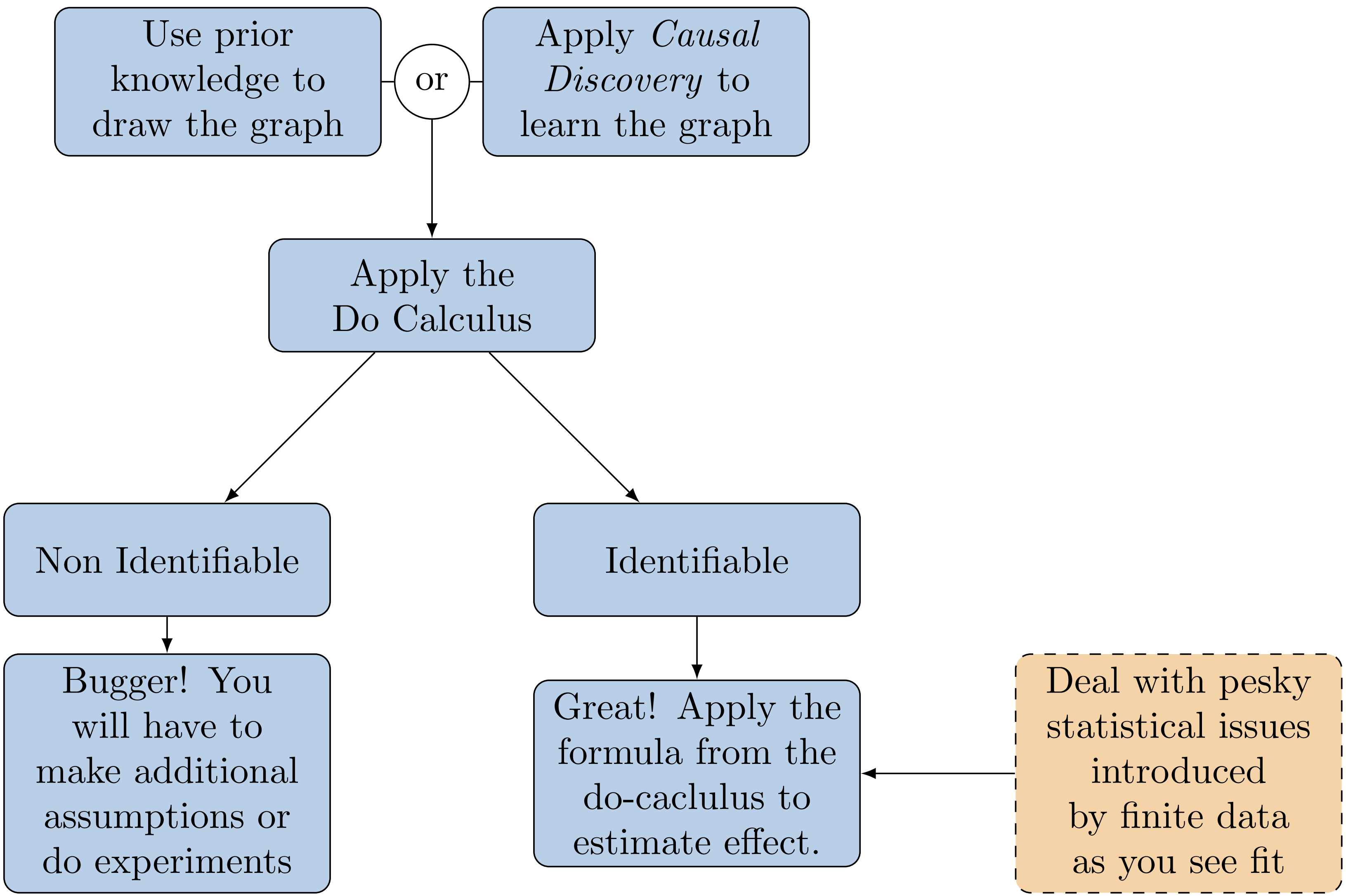

$$P(P, D, C|do(S = no)) = P(P)P(D|P)P(C|S, D)$$

- Truncated product formula: drop terms for intervened variables from the factorisation.
- A CBN represents the set of all possible interventional distributions over its variables

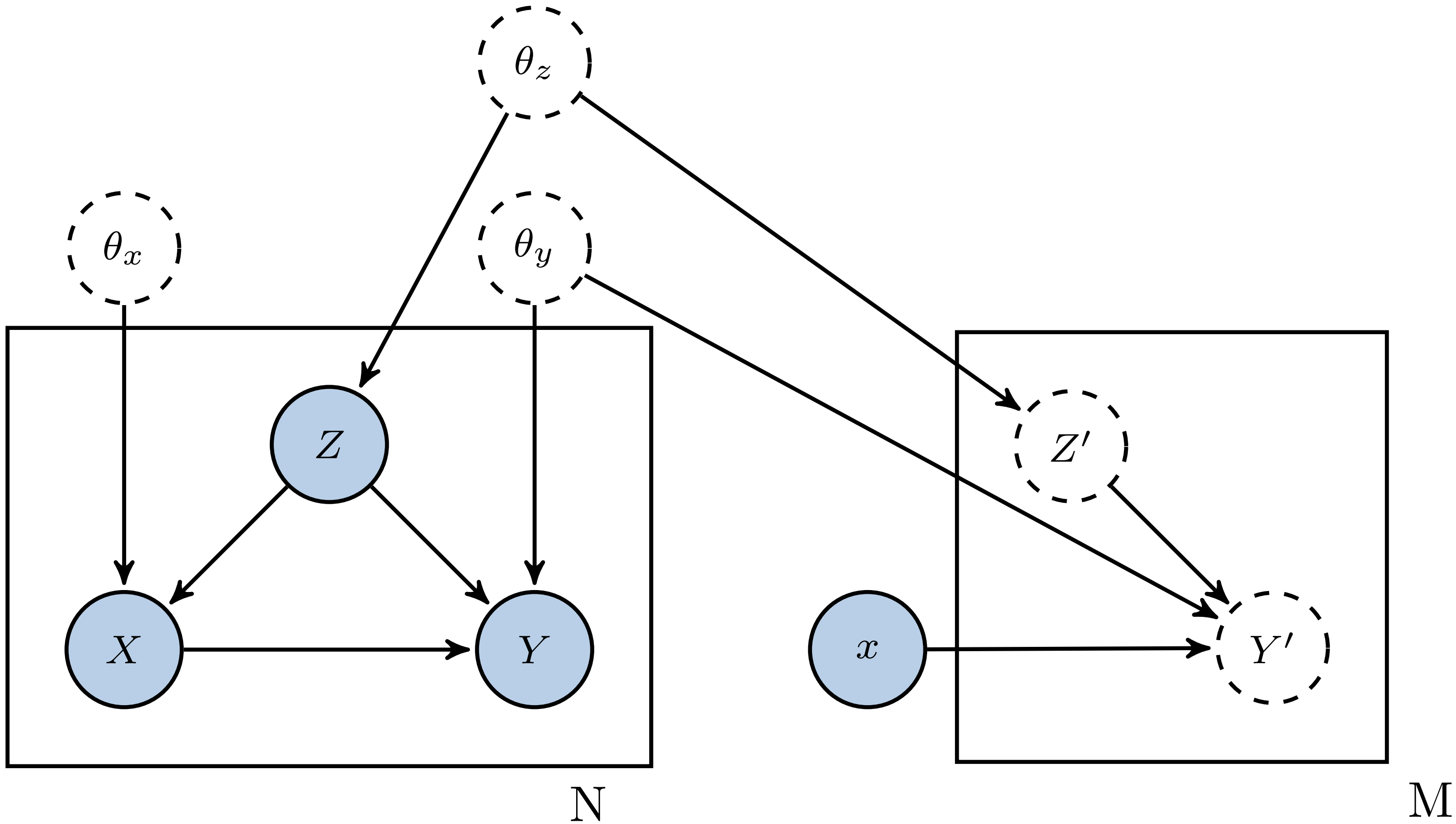
# Can the effect of X on Y be identified?



# Inference from observational data – aka Pearl

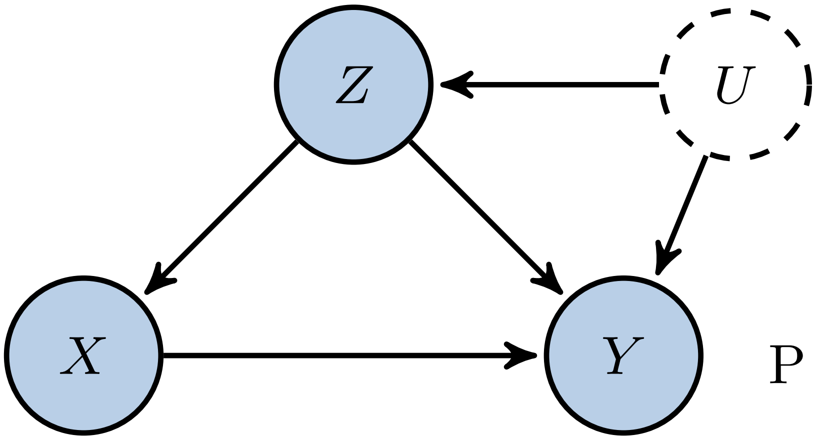


# Can’t we just be Bayesian?



# Careful with that prior

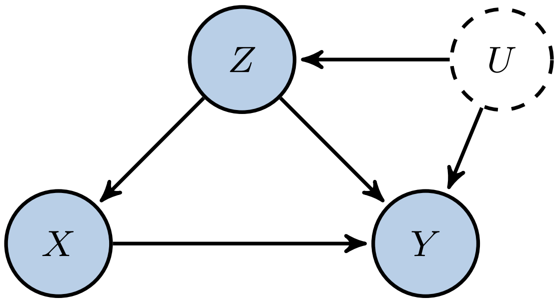
- The goal is to estimate the causal effect of X on Y
- It is indefinable via the do-calculus




$$\mathrm{P}\left(Y|U,Z,X\right) = N(w_{yu}U + w_{yz}Z + w_{yx}X, v_y)$$

$$\frac{d}{dx}\mathbb{E}\left[Y|do(X=x)\right] = w_{yx}$$

# Careful with that prior



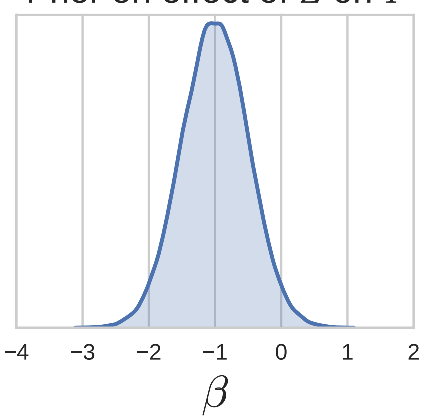

$$\mathrm{P}\left(Y|U,Z,X\right) = N(\underset{3}{w_{yu}}U + \underset{-1}{w_{yz}}Z + \underset{0.5}{w_{yx}}X, v_y)$$

$$\mathrm{P}\left(Y|Z,X\right) = N\left(\underbrace{\left(w_{yz} + \frac{w_{yu}w_{zu}}{w_{zu}^2 + \frac{v_z}{v_u}}\right)}_{\beta\,=\,0.47} Z + w_{yx}X, \varepsilon\right)$$

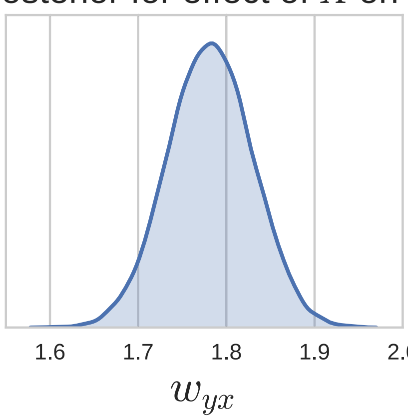
$$\frac{d}{dx}\mathbb{E}\left[Y|do(X=x)\right] = w_{yx}$$

## Prior centered on causal effect

Prior on effect of $Z$ on $Y$

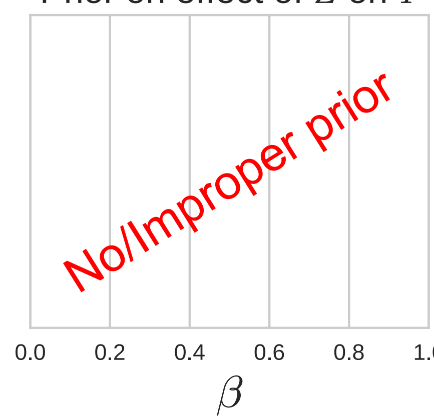



Posterior for effect of $X$ on $Y$

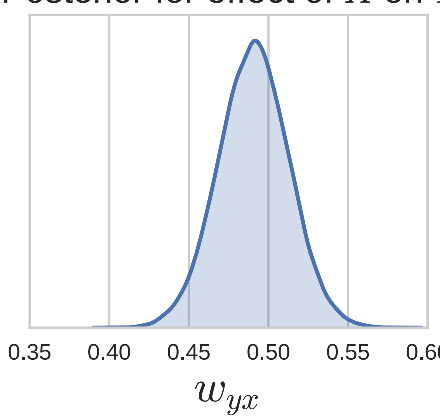



## Improper prior

Prior on effect of $Z$ on $Y$




Posterior for effect of $X$ on $Y$

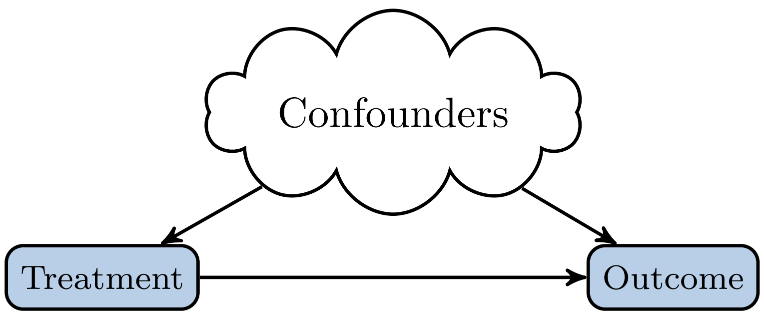


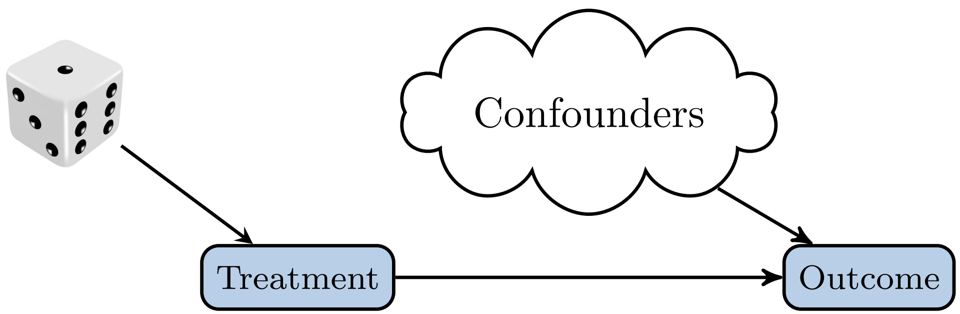
# Where are we up to?

- What is causality?
  - Predicting the outcome of interventions
- Why should we care in machine learning?
  - Because we want to act in response to our models
- Observational causal inference
  - Leveraging invariance assumptions to map properties from a system pre-intervention to the system post-intervention
- Interventional causal inference
- A unified approach - Causal Bandits

# The interventional approach to causality

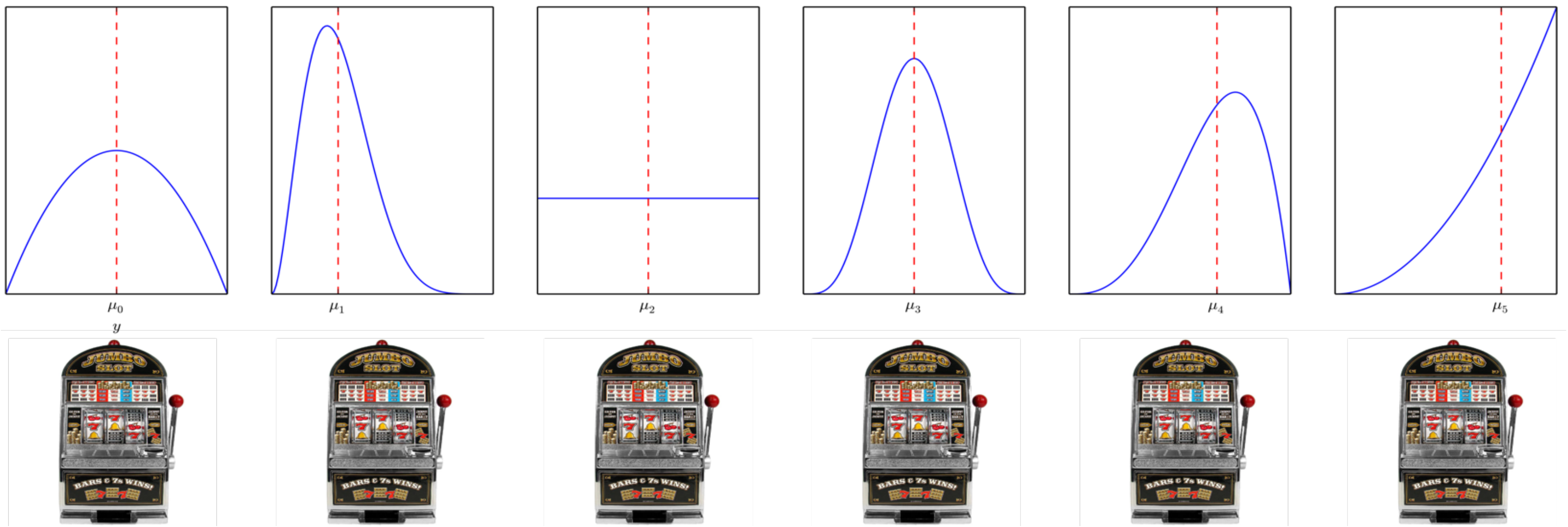
- Randomized trials are the traditional gold standard for determining causality




Bandits?

# Classic Multi-armed Bandits

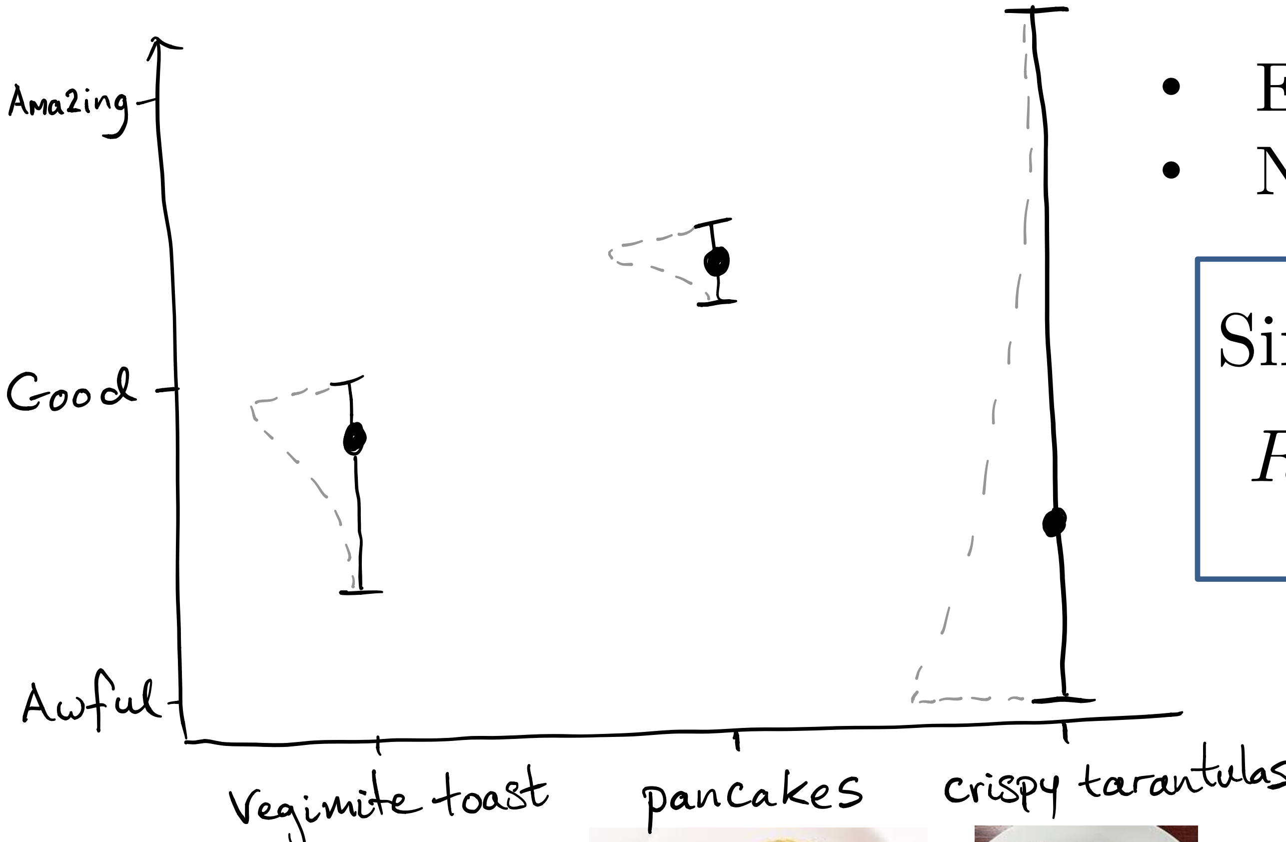


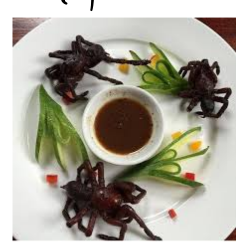

In each round $t \in [1, ...T]$,

1. the learner selects an action $a_t \in \{1, ..., k\}$
2. the world samples rewards for each action, $[Y_t^1, ..., Y_t^k] \sim \mathrm{P}\left(\boldsymbol{y}\right)$
3. the learner receives the reward for the selected action $Y_t^{a_t}$

The goal is to maximise the total reward

# What should we suggest for breakfast?



- Explore-exploit trade-off
- Non i.i.d data

**Simple Regret**

$$R_T = \mu_{i^*} - \mathbb{E}\left[\mu_{\hat{i}^*}\right]$$

# Contextual Bandits

$X$ = Australian

$X$ = Likes sweets

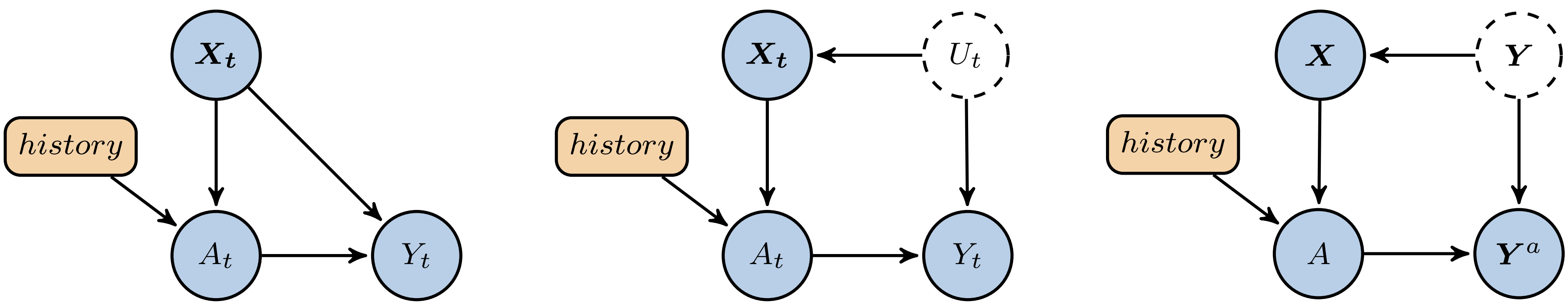
$X$ = Enjoys Novelty

$$\boldsymbol{X} \sim \mathrm{P}(x) \qquad Y \sim \mathrm{P}(y|X, a)$$

# Causal structure of bandit problems

- Contextual bandit algorithms do not require that the context is a cause of the outcome.

(a) $\boldsymbol{X}$ causes $\boldsymbol{Y}$ (b) $\boldsymbol{X}$ and $\boldsymbol{Y}$ are confounded (c) $\boldsymbol{Y}$ causes $\boldsymbol{X}$



# Where are we up to?

- What is causality?
  - Predicting the outcome of interventions
- Why should we care in machine learning?
  - Because we want to act in response to our models
- Observational causal inference
  - Leveraging invariance assumptions to map properties from a system pre-intervention to the system post-intervention
- Interventional causal inference
  - Learning by doing
- A unified approach - Causal Bandits

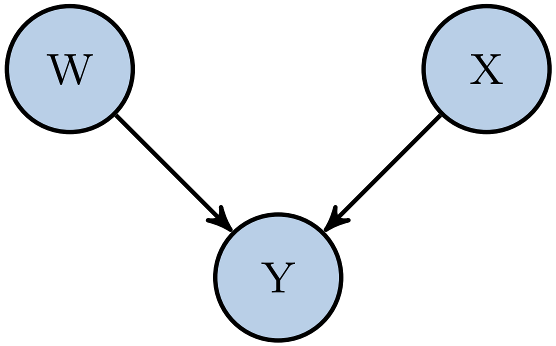
# Why unify causal graphs and bandits?

- Regret grows linearly with the number of actions
- There is a lot of data available for which we do not control/know the process selecting actions
- Bandits are a nice subset of general RL problems

# Causal Bandit Problems

- Every (allowable) assignment of variables to values is a bandit arm.
- Reward is value of a single specified node in the graph after the action.

Observe $\boldsymbol{X}_c$, $do(\boldsymbol{X}_{\boldsymbol{a}} = \boldsymbol{x})$, observe $\boldsymbol{X}_o$, obtain reward $Y$

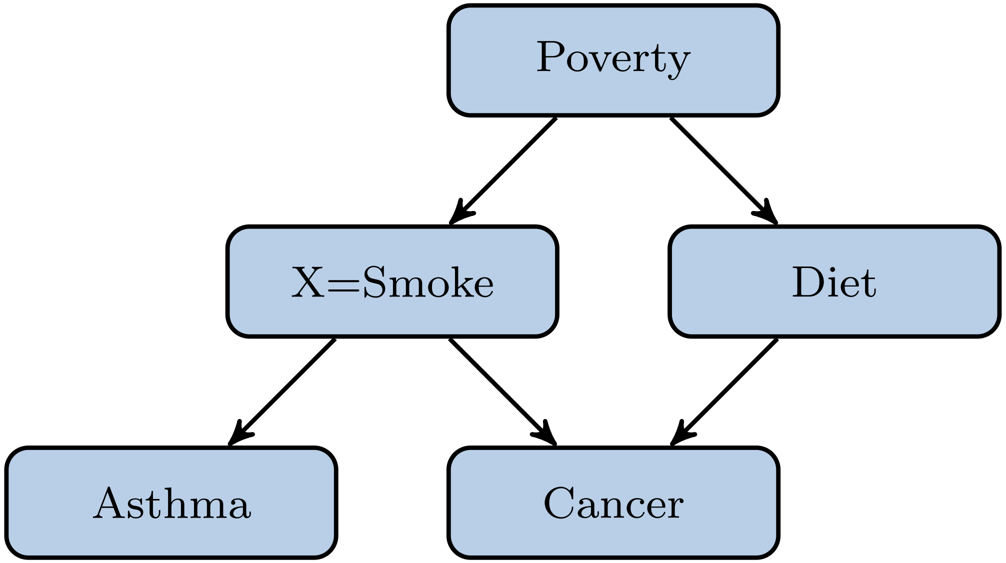



$$\mathcal{A} = \begin{array}{|c|}\hline do(W=0, Z=0) \\ do(W=0, Z=1) \\ do(W=1, Z=0) \\ do(W=1, Z=1) \\ \hline do(W=0) \\ do(W=1) \\ do(Z=0) \\ do(Z=1) \\ do() \\ \hline \end{array}$$

# How can causal structure be leveraged

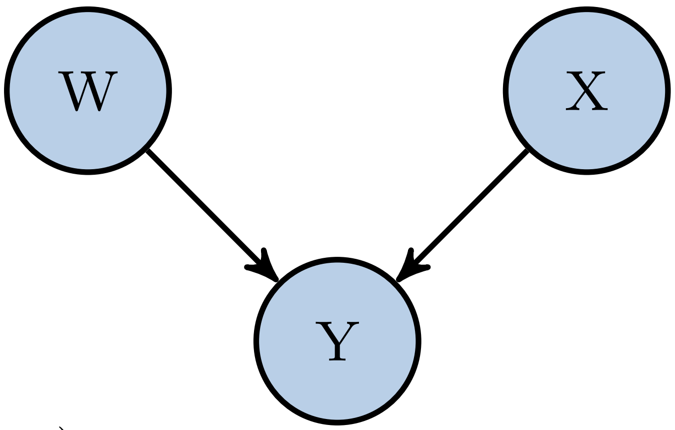
1. Prune actions before commencing

2. Obtain information about the reward for one action by selecting another.



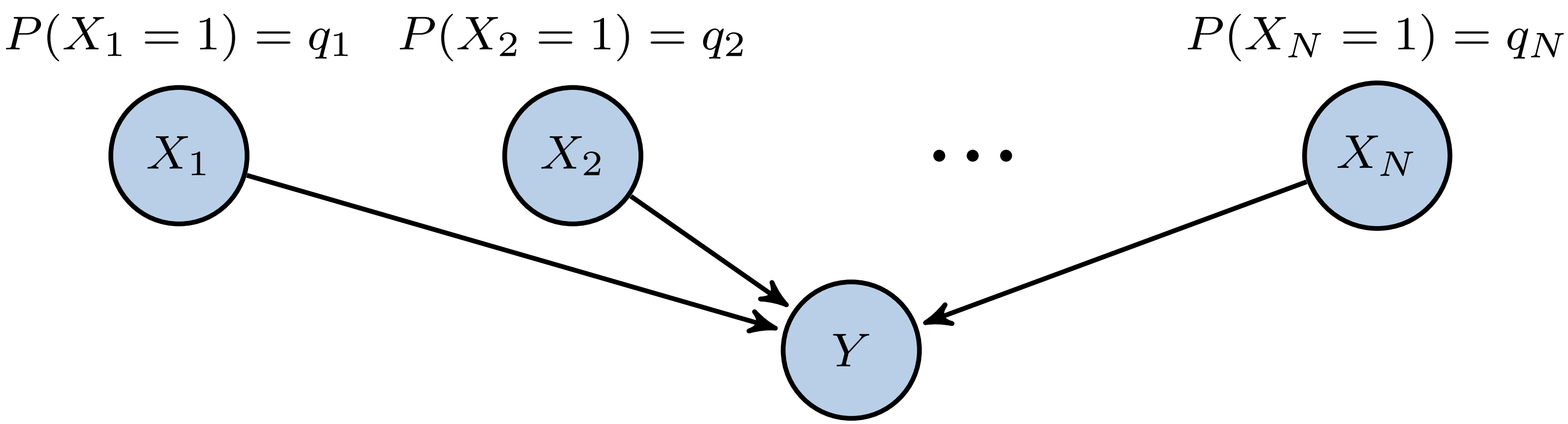
# C-Bandit problems with post-action feedback

- Problems for which additional feedback available only **after** action selected (no context).




$$P(Y|do(W=1)) = P(Y|W=1)$$
$$= P(Y|W=1, do(Z=1))P(Z=1) + P(Y|W=1, do(Z=0))P(Z=0)$$

# The parallel bandit problem

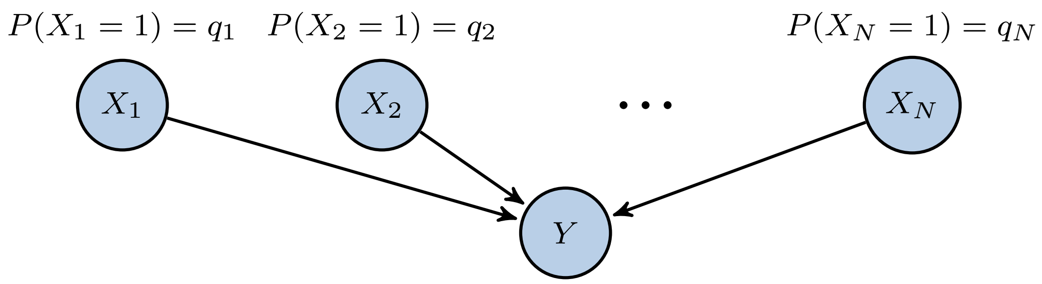



At each timestep $t \in 1, ..., T$:

1. The agent sets the value of zero or one variables, $do()$ or $do(X_i = j)$.
2. Non-intervened variables sampled from underlying distributions.
3. The agent receives reward $Y$ sampled from $P(Y|x_1, ..., x_N)$.
4. The values of all variables, $x_1, ..., x_N$, are revealed to the agent.

The variables are binary $\implies$ there are $2N + 1$ arms.
If the agent opts to observe, $do()$, it will obtain an estimate for half the arms

# The parallel bandit problem – algorithm

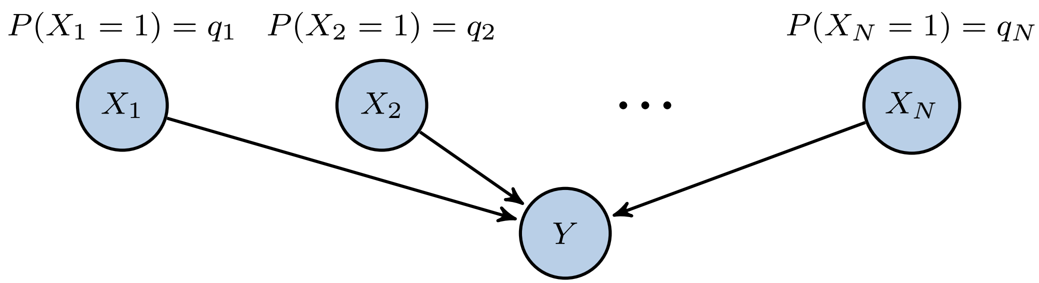



- Observe for first T/2 rounds to estimate reward for actions that occur frequently naturally.
- Uniformly explore infrequent actions in remaining rounds.
- But how do we define infrequent?

$$m(\boldsymbol{q}) = min\left\{m : q_m \geq \frac{1}{m}\right\} \qquad \text{infrequent} \equiv \left\{i : q_i < \frac{1}{m(\boldsymbol{q})}\right\}$$

$$\boldsymbol{q} = [0, 0.01, 0.32, 0.33, 0.33, 0.35, .37, 0.41, 0.45, 0.49, 0.49] \implies m(\boldsymbol{q}) = 4$$

# The parallel bandit problem – regret bounds




Parallel bandit $$\Omega\left(\sqrt{\frac{m(\boldsymbol{q})}{T}}\right) \geq R_T \leq \mathcal{O}\left(\sqrt{\frac{m(\boldsymbol{q})}{T}\log\left(\frac{NT}{m(\boldsymbol{q})}\right)}\right)$$

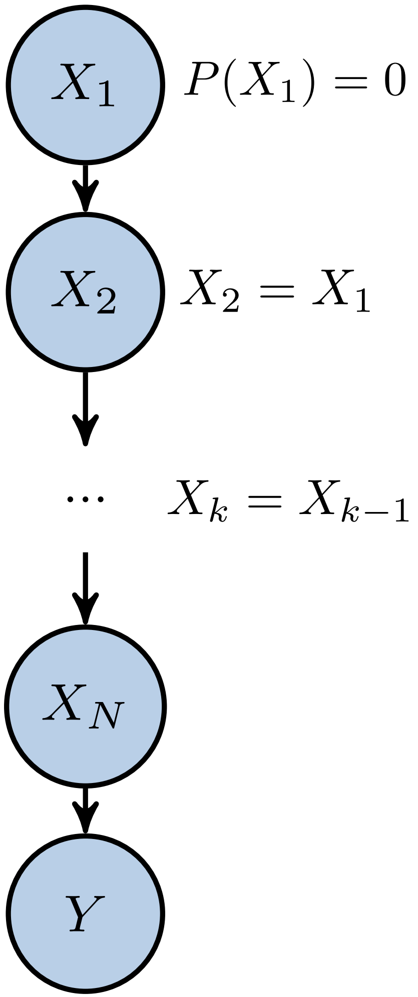
Standard bandit $$\Omega\left(\sqrt{\frac{N}{T}}\right) \geq R_T \leq \mathcal{O}\left(\sqrt{\frac{N}{T}\log T}\right)$$

$m(\boldsymbol{q}) < N$ can be thought of as the *effective* number of arms

# General graphs - challenges

In general, $P(Y|X_i = j) \neq P(Y|do(X_i = j))$

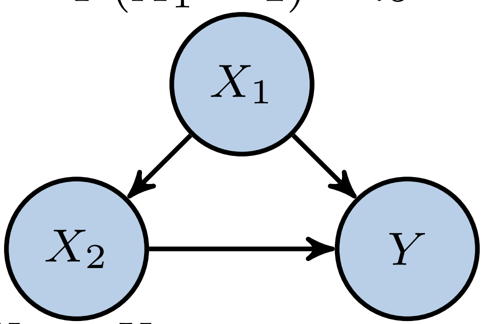
- We could map from observation to intervention via the do-calculus but,
- Its no longer optimal ignore the information intervening on one variable can provide about another.



# General graphs - challenges

- The variance of the observational estimate on $P(Y|do(X_i = j))$ no longer depends only on $P(X_i = j)$

$P(X_1 = 1) = .5$




$X_2 = X_1$

$P(X_2 = 1) = 0.5$

$$P(Y|do(X_2 = 1)) = P(X_1 = 1)P(Y|X_1 = 1, X_2 = 1) + P(X_1 = 0)P(Y|X_1 = 0, X_2 = 1)$$

# General graphs - solution

We assume the distribution over the parents of $Y$ given each action, $\mathrm{P}\left(\mathcal{P}\mathrm{a}_Y \,|a\right)$, is known for all $a \in \mathcal{A}$

Let $\eta$ be a distribution of interventions $a \in \mathcal{A}$

for $t \in \{1, ..., T\}$ :

Sample action $a_t$ from $\eta$

for $a \in \mathcal{A}$ :

$$\hat{\mu}_a = \frac{1}{T} \sum_{t=1}^{T} Y_t \frac{\mathrm{P}\left\{\mathcal{P}\mathrm{a}_Y(X_t)|a\right\}}{\sum_{b \in \mathcal{A}} \eta_b \mathrm{P}\left(\mathcal{P}\mathrm{a}_Y(X_t)|b\right)}$$

# General graphs - solution

$$m(\eta) = \max_{a \in \mathcal{A}} \mathbb{E}_a \left[ \frac{\mathrm{P}\left\{\mathcal{P}\mathrm{a}_Y(X) | a\right\}}{\sum_{b \in \mathcal{A}} \eta_b \mathrm{P}\left\{\mathcal{P}\mathrm{a}_Y(X) | b\right\}} \right]$$

$$\eta^* = \arg\min_{\eta} m(\eta) \qquad m(\eta^*) \leq |\mathcal{A}|$$

$$R_T \in \mathcal{O}\left( \sqrt{\frac{m(\eta)}{T} \log\left(2T|\mathcal{A}|\right)} \right)$$

General CB upper bound

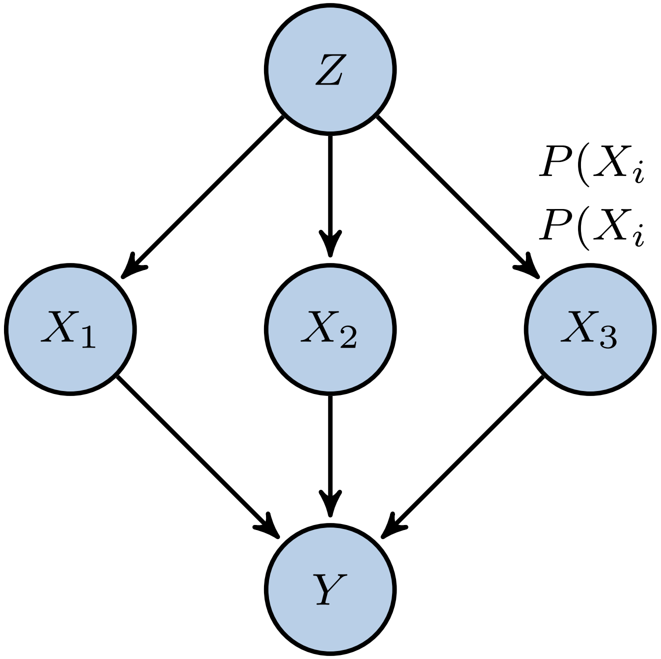
$$R_T \in \Omega\left( \sqrt{\frac{|\mathcal{A}|}{T}} \right)$$

Standard lower bound

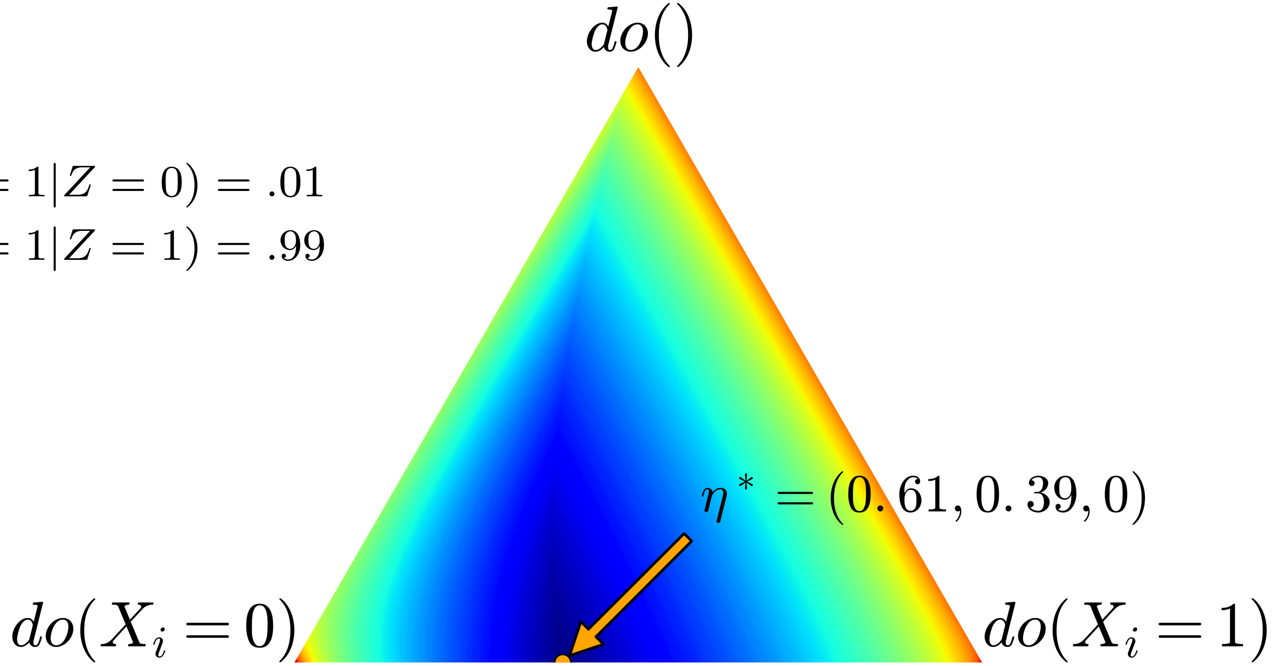
# General graphs - solution

$$\boxed{m(\eta)} = \max_{a \in \mathcal{A}} \mathbb{E}_a \left[ \frac{\mathrm{P}\left\{\mathcal{P}a_Y(X)|a\right\}}{\sum_{b \in \mathcal{A}} \eta_b \mathrm{P}\left\{\mathcal{P}a_Y(X)|b\right\}} \right]$$

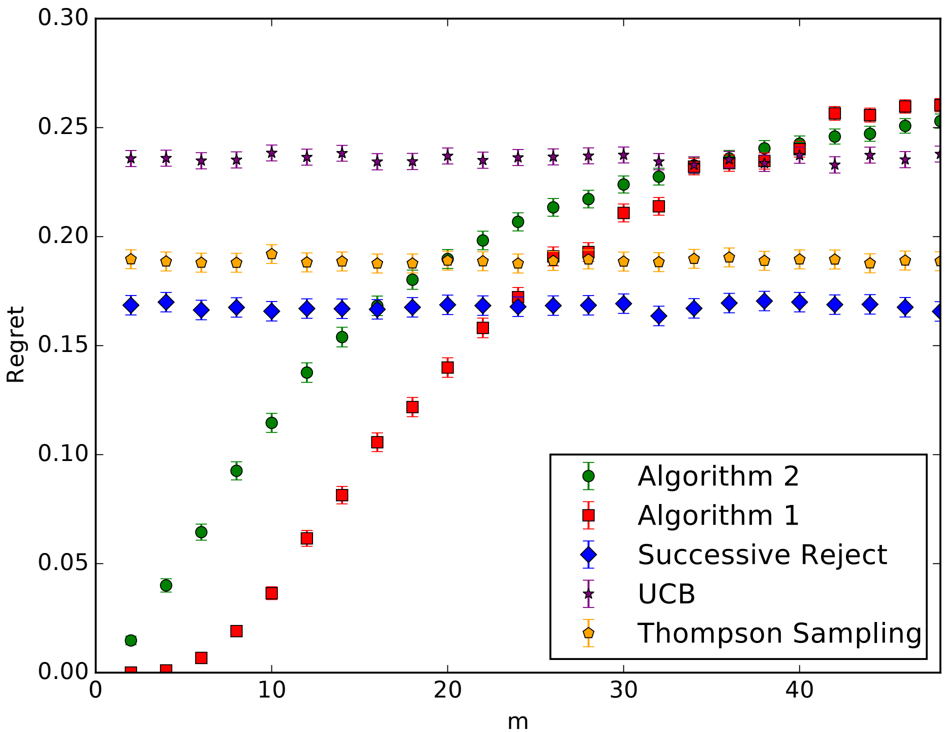
$P(Z=1) = 0.7$



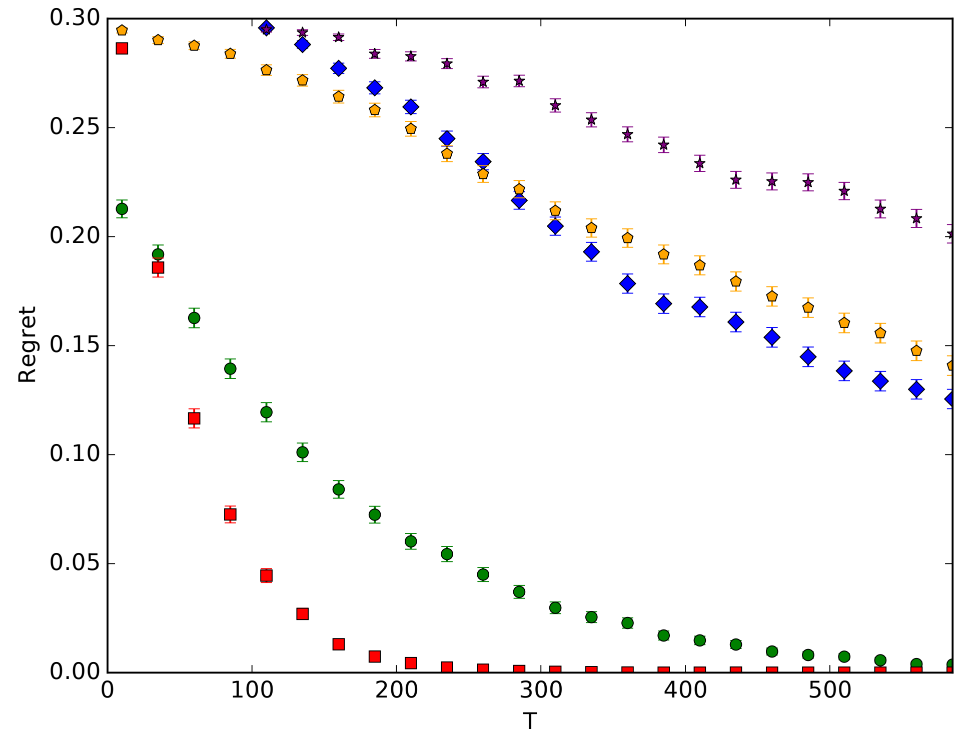

$P(X_i = 1|Z=0) = .01$
$P(X_i = 1|Z=1) = .99$



# Experiments



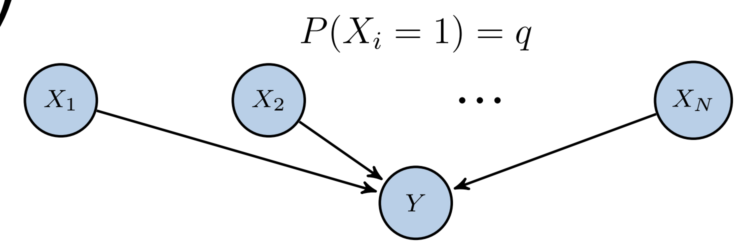

(a) Simple regret vs $m(\boldsymbol{q})$ for fixed horizon $T = 400$ and number of variables $N = 50$




(b) Simple regret vs horizon, $T$, with $N = 50$, $m = 2$ and fixed $\epsilon = .3$

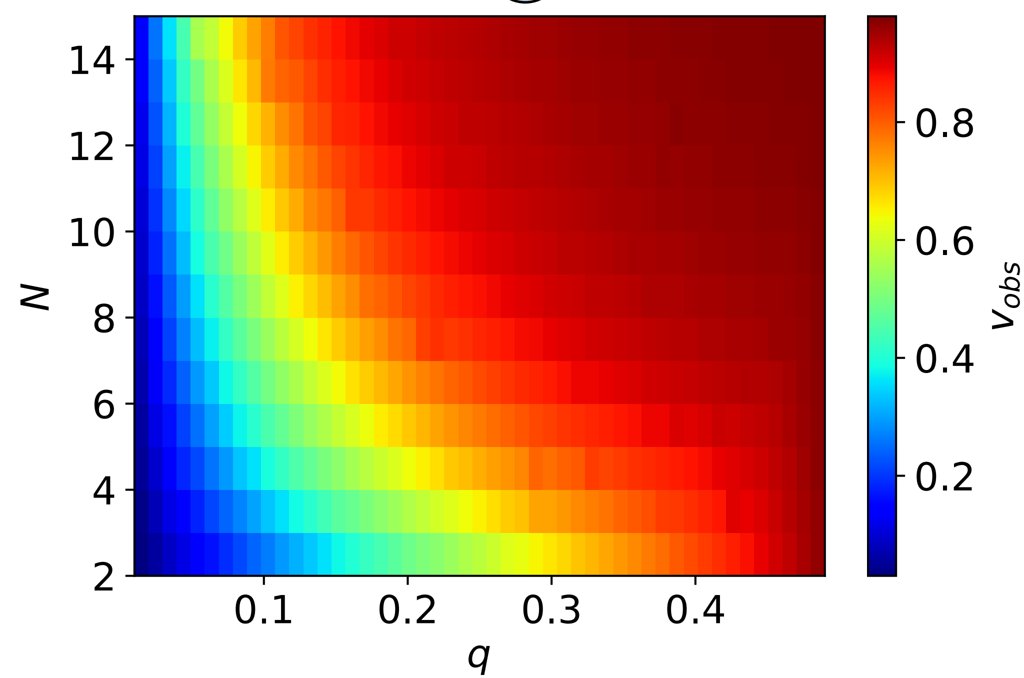
# Quantifying the value of intervention

$$R_T \in \mathcal{O}\left(\sqrt{\frac{m(\eta)}{T}\log\left(2T|\mathcal{A}|\right)}\right) \Leftarrow \text{holds for all } \eta$$





$$= \frac{m(\eta^*)}{m(\eta_{do})}$$



# Conclusion

- What is causality?
  - Predicting the outcome of interventions
- Why should we care in machine learning?
  - Because we want to act in response to our models
- Observational causal inference
  - Leveraging invariance assumptions to map from properties from a system pre-intervention to the system post-intervention
- Interventional causal inference
  - Learning by doing
- A unified approach - Causal Bandits
  - Bandits and causal observational inference fundamentally solve the same problem – learning to intervene
  - Adding structure to bandit problems with causal graphical models allows us to explore large action spaces much more quickly

# Open questions

- Relaxing the assumption that the interventional distributions are known
- Contextual causal bandit problems
- Cumulative regret
- Extensions to MDPs
- Connections between causal effect estimation and off-policy evaluation

# Questions

?

# The do calculus (simplified)

1. D-separation still applies after intervention.

   $(Cancer \perp\!\!\!\perp Asthma | Smoke)_{G_{\overline{X}}} \implies P(Cancer|do(Smoke), Asthma) = P(Cancer|do(Smoke))$

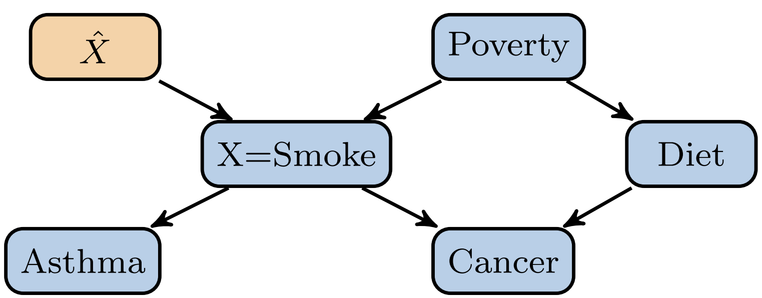
2. If there are no backdoor paths from $X$ to $Y$ then intervention≡observation.

   $(\hat{X} \perp\!\!\!\perp Cancer | X, Poverty)_{G^{\dagger}} \implies P(Cancer|do(Smoke), Poverty) = P(Cancer|Smoke, Poverty)$

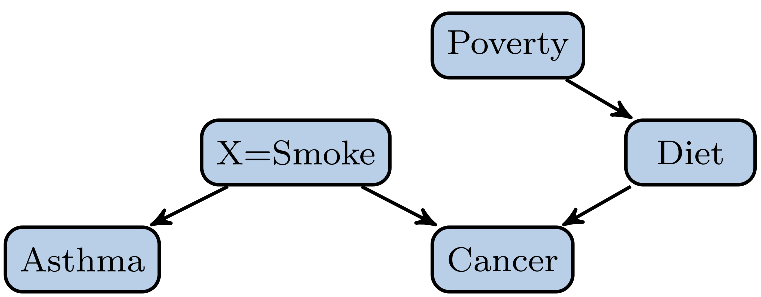
3. If there are only backdoor paths from $X$ to $Y$ then intervention doesn't change $P(Y)$.

   $(\hat{X} \perp\!\!\!\perp Diet)_{G^{\dagger}} \implies P(Diet|do(Smoke)) = P(Diet)$

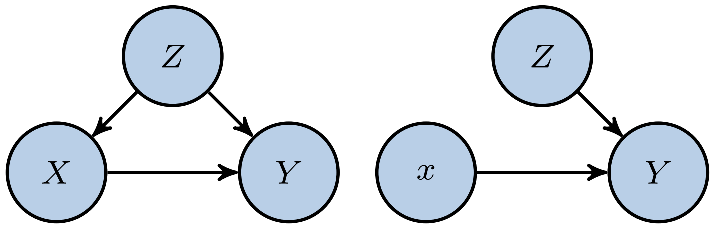
(a) $G^{\dagger}$




(b) $G_{\overline{X}}$



# Pesky statistical issues




$$\text{Training data} \quad (\boldsymbol{z}_i, \boldsymbol{x}_i, y_i) \sim \mathrm{P}\left(Z\right)\mathrm{P}\left(X|Z\right)\mathrm{P}\left(Y|X,Z\right)$$

$$\text{Test data} \quad (\boldsymbol{z}_i, \boldsymbol{x}_i, y_i) \sim \mathrm{P}\left(Z\right)\delta(X=x)\,\mathrm{P}\left(Y|X,Z\right)$$