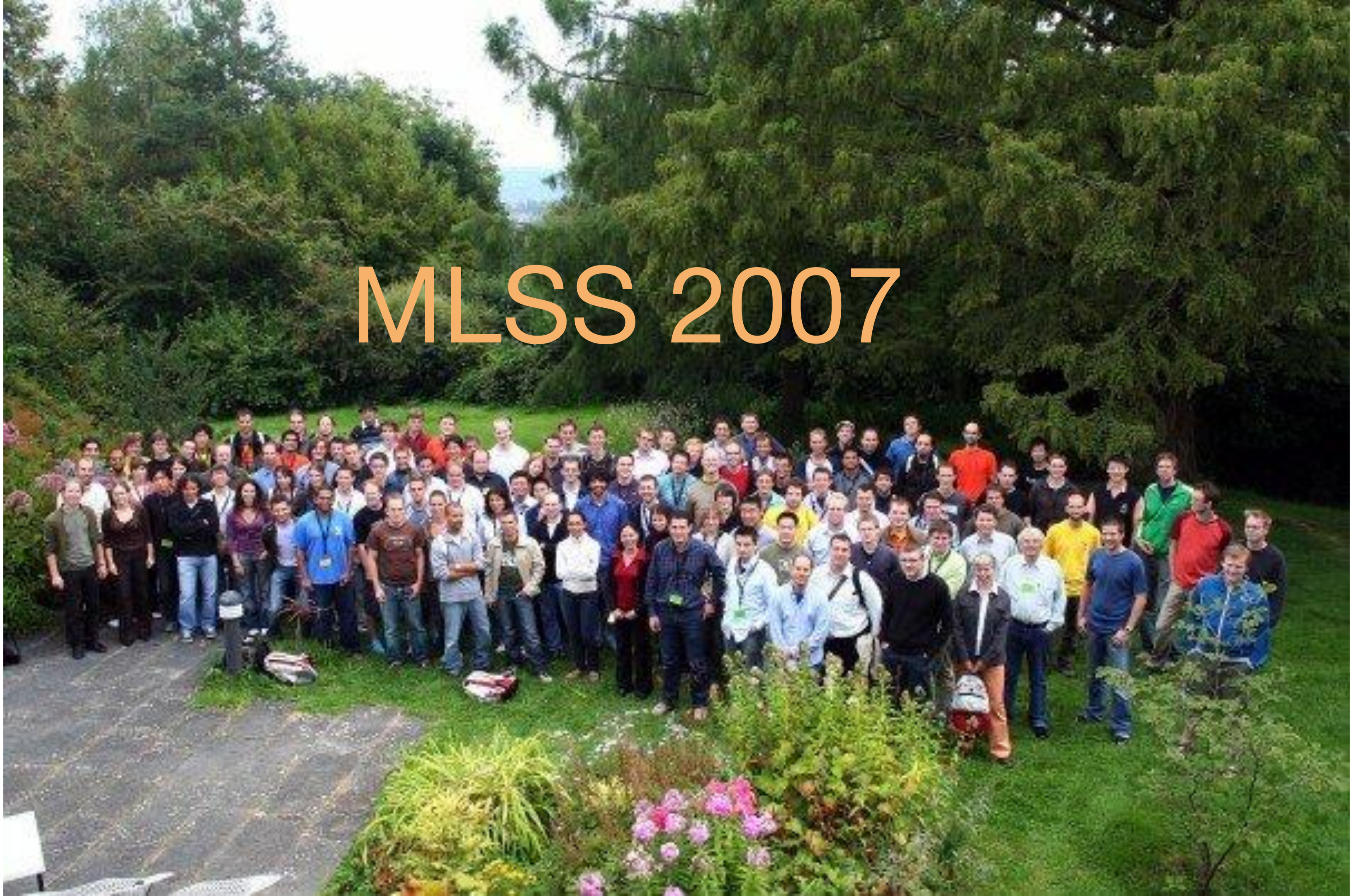
Causal Inference in Everyday Machine Learning

Ferenc Huszar
Twitter, London


MLSS 2007

Outline

High level motivation

Basic cases

Interventions, do-calculus

Counterfactuals

Offline A/B testing

Counterfactual Risk Minimization

Practical: toy examples in python

blindspot

JUDEA PEARL
WINNER OF THE TURING AWARD
AND DANA MACKENZIE

THE BOOK OF WHY

α  β

THE NEW SCIENCE OF CAUSE AND EFFECT

To Build Truly Intelligent Machines, Teach Them Cause and Effect

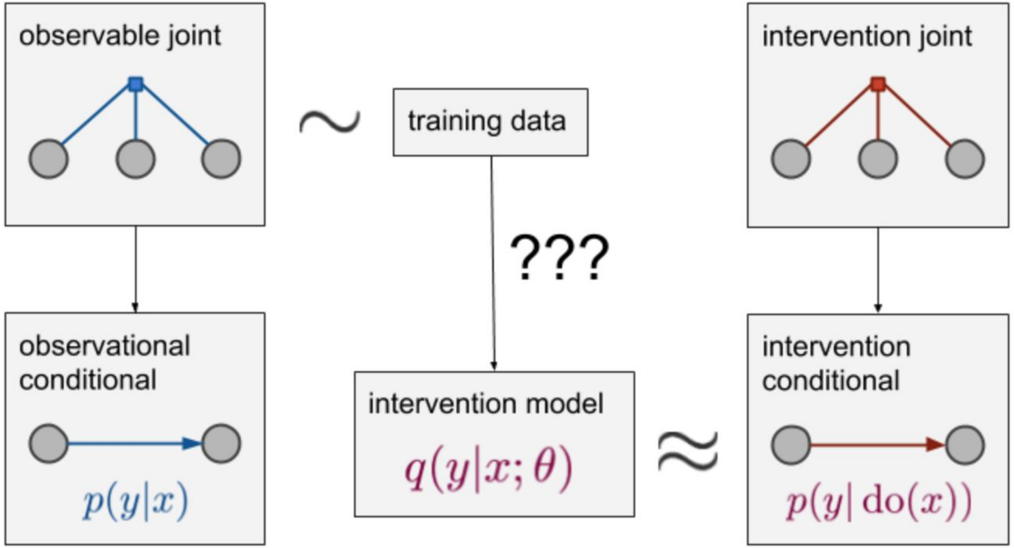
53

Judea Pearl, a pioneering figure in artificial intelligence, argues that AI has been stuck in a decades-long rut. His prescription for progress? Teach machines to understand the question why.

But as Pearl sees it, the field of AI got mired in probabilistic associations. These days, headlines tout the latest breakthroughs in machine learning and neural networks. We read about computers that can master ancient games and drive cars. Pearl is underwhelmed. As he sees it, the state of the art in artificial intelligence today is merely a souped-up version of what machines could already do a generation ago: find hidden regularities in a large set of data. “All the impressive achievements of deep learning amount to just curve fitting,” he said recently.

May 24th, 2018

ML beyond Curve Fitting: An Intro to Causal Inference and do-Calculus



inference.vc/untitled

blindspot

despite its importance, causal inference was not widely understood

basic hygiene

basic hygiene

think of it like understanding overfitting

not rocket science

not rocket science

main ideas quite intuitive, methods are often simple

not just healthcare

not just healthcare

it is relevant to many applications of machine learning

not either or

not either or

Causal or counterfactual reasoning can and should be combined with deep learning

Take homes

blindspot

basic hygiene

not rocket science

not just healthcare

not either or

Three levels of questions

Observational questions

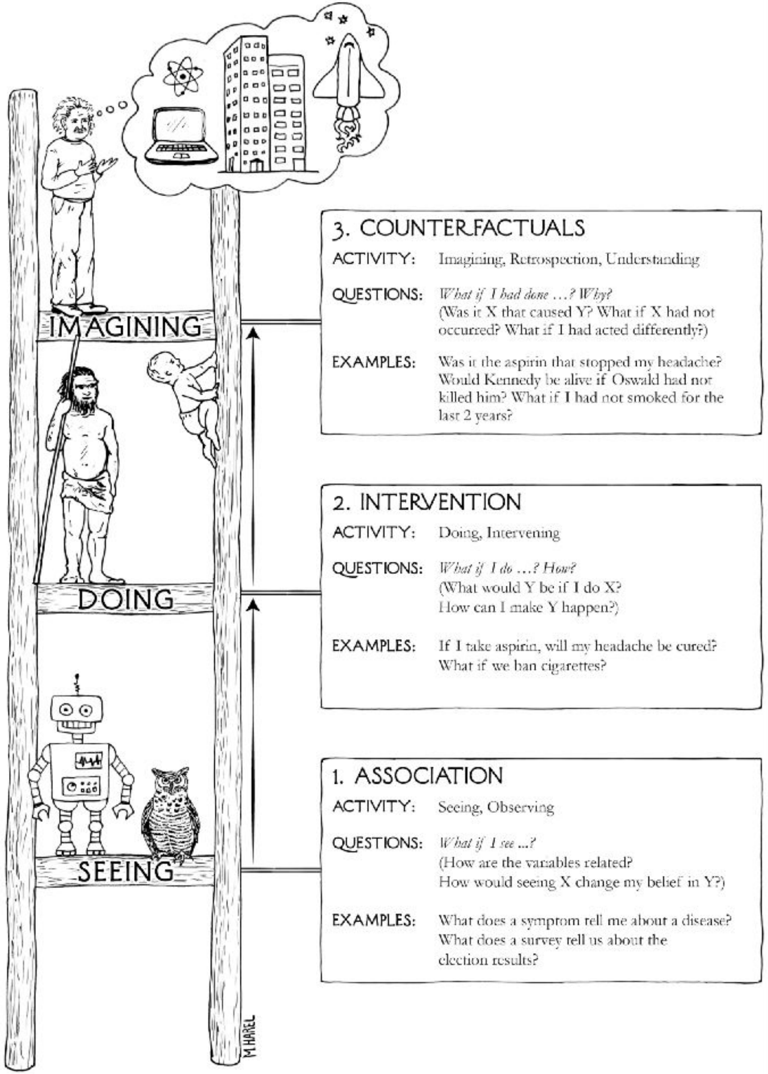
Do people who are given the drug tend to recover?

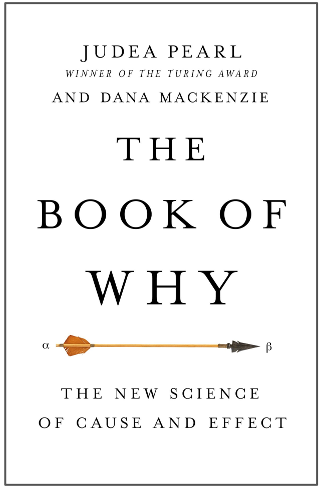
Action/Intervention Questions

If I give people this drug, how likely it is that they recover?

Counterfactuals

Had I given the patient the drug two weeks ago, would he have recovered?

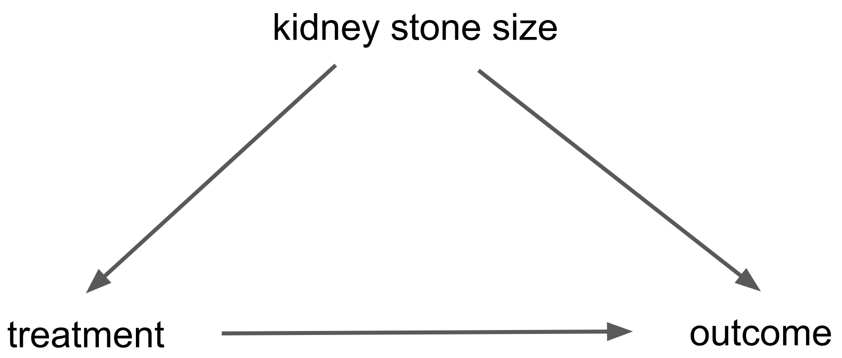

IMAGINING
DOING
SEEING
3. COUNTERFACTUALS
ACTIVITY: Imagining, Retrospection, Understanding
QUESTIONS: *What if I had done …? Why?*
(Was it X that caused Y? What if X had not occurred? What if I had acted differently?)
EXAMPLES: Was it the aspirin that stopped my headache? Would Kennedy be alive if Oswald had not killed him? What if I had not smoked for the last 2 years?
2. INTERVENTION
ACTIVITY: Doing, Intervening
QUESTIONS: *What if I do …? How?*
(What would Y be if I do X? How can I make Y happen?)
EXAMPLES: If I take aspirin, will my headache be cured? What if we ban cigarettes?
1. ASSOCIATION
ACTIVITY: Seeing, Observing
QUESTIONS: *What if I see …?*
(How are the variables related? How would seeing X change my belief in Y?)
EXAMPLES: What does a symptom tell me about a disease? What does a survey tell us about the election results?
M.HAREL


JUDEA PEARL
WINNER OF THE TURING AWARD
AND DANA MACKENZIE
THE
BOOK OF
WHY
α
β
THE NEW SCIENCE
OF CAUSE AND EFFECT

Simpson’s paradox

	outcome
Treatment A (minimally invasive)	78% (273/350)
Treatment B (invasive)	**83%** (289/350)

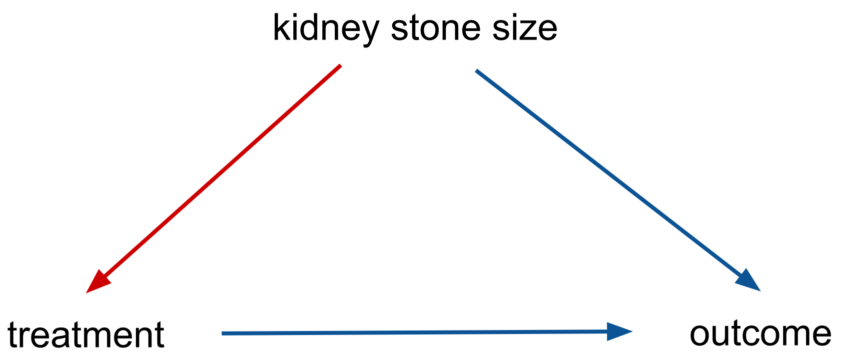
Causal diagram - Confounder



Observed association = causal association “+” spurious association

$p(y|x)$ $\qquad$ $p(y|do(x))$

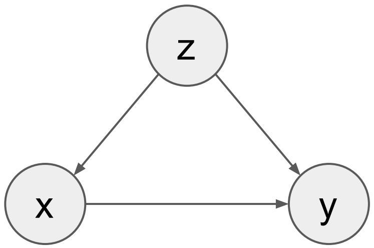
Stability/robustness under policy change



Observed association = causal association "+" spurious correlation

$p(y|x)$ $\quad$ $p(y|do(x))$

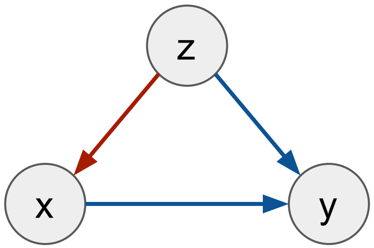
Reminder: confounder adjustment formula



$$p(y|do(x)) = \sum_z p(z)p(y|x,z)$$

$$p(y|x) = \sum_z p(z|x)p(y|x,z)$$

Reminder: confounder adjustment formula



$$p(y|do(x)) = \sum_z p(z)p(y|x,z)$$

$$p(y|x) = \sum_z p(z|x)p(y|x,z)$$

Quiz: Do you need p(y | do(x)) or p(y| x)?

Scenario 1: Doctor
You work for the hospital. You do the analysis in order to improve your decision making around which treatment to give patients.

Scenario 2: Insurance company
You work for an insurance company developing an insurance product for people who are about to undergo surgery (treatment 2). You have to estimate the probability of a negative outcome so you can set the price of the product

Scenario 3: Scientist
You are a researcher, and you want to understand kidney stone disease.

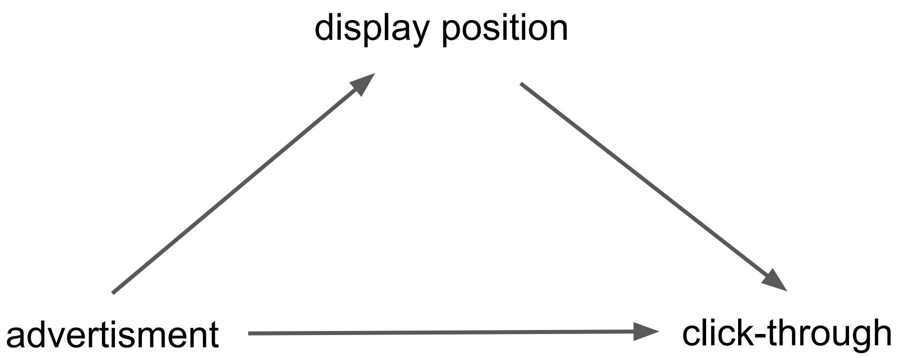
Online advertising example

	click-through
Advertisement A	6.2% (124/2000)
Advertisement B	**7.5%** (149/2000)

Is position a confounding variable?

	click-through	position=1	position=2
Advertisement A	6.2% (124/2000)	**18.1%** (32/176)	**5.1%** (92/1823)
Advertisement B	**7.5%** (149/2000)	15.6% (78/500)	4.8% (71/1500)

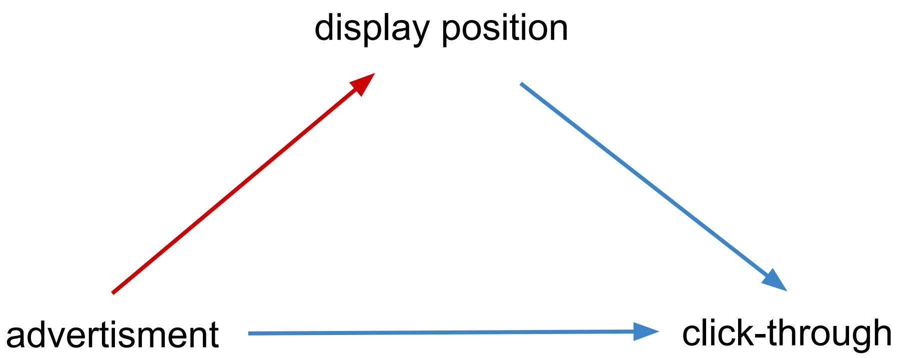
Causal diagram - Mediator variable



Observed association = direct causal effect “+” indirect causal effect

$$p(y|x) = p(y|do(x))$$

Causal diagram - Mediator variable



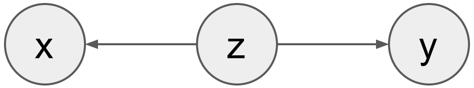
Observed association = direct causal effect “+” indirect causal effect

Patterns in causal diagrams

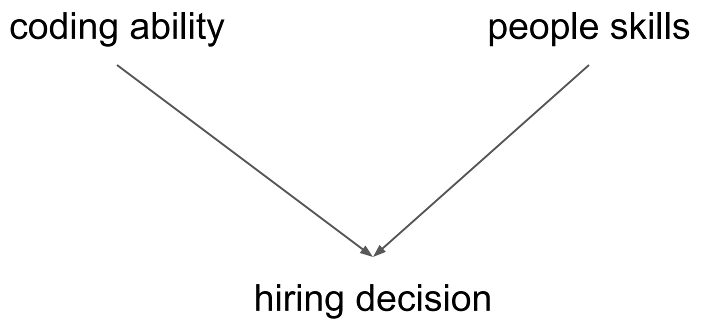
Confounder

(hidden) common cause

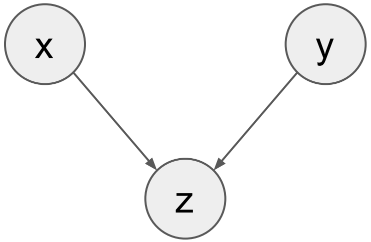
Introduces spurious associations



Explaining away - example



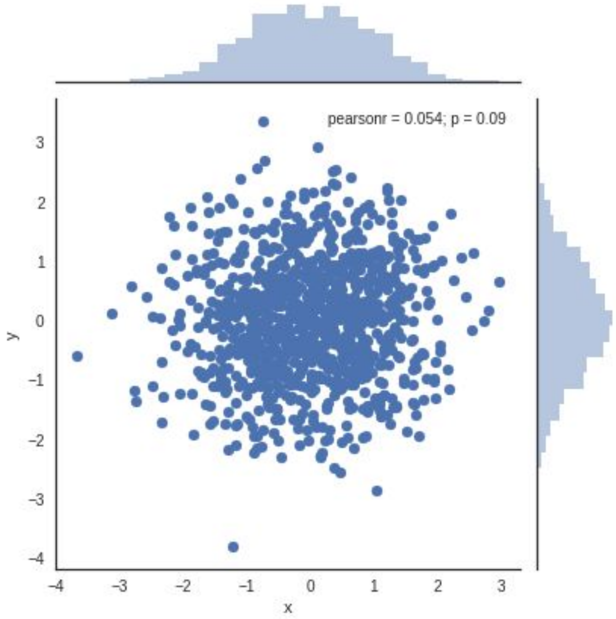
Collider bias toy example



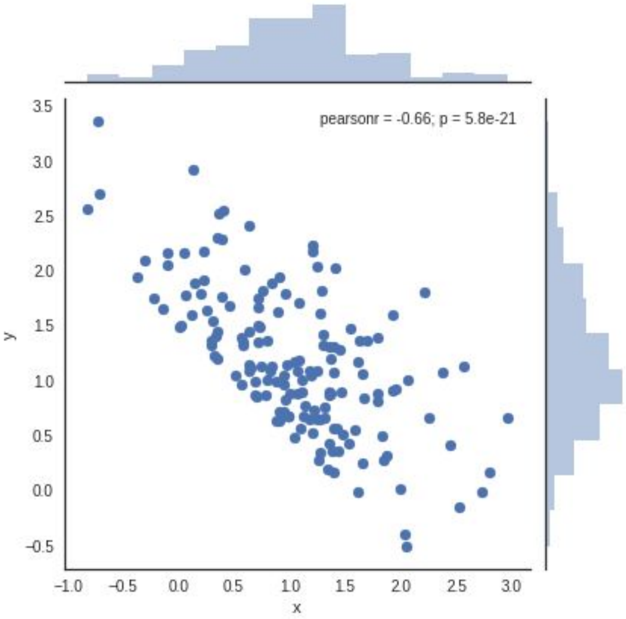
$$p(x) = \mathcal{N}(0, 1)$$
$$p(y) = \mathcal{N}(0, 1)$$
$$p(z = 1|x, y) = \Phi(3 * x + 3 * y - 1.5)$$

Collider bias illustration

$$p(x, y)$$



$$p(x, y | z = 1)$$



But why would I condition on a collider then?

We always implicitly condition on the datapoint being in the dataset

- In ML, we assume data is sampled i.i.d. from $P_{X,Y}$
- In reality, it is always sampled from $P_{X,Y|O=1}$
- O: binary variable, whether or not you observe this datapoint
- If O is independent of X and Y, you can ignore it
- If O is a collider, you have a problem!
- Sampling biases:
  - Availability bias
  - Survivorship bias
  - etc...

Basic patterns in causal diagrams

Confounder

(hidden) common cause

introduces spurious associations

$x \leftarrow z \rightarrow y$

Mediator

creates causal association between two variables

direct causal effect vs indirect causal effect

$x \rightarrow z \rightarrow y$

$x \leftarrow z \leftarrow y$

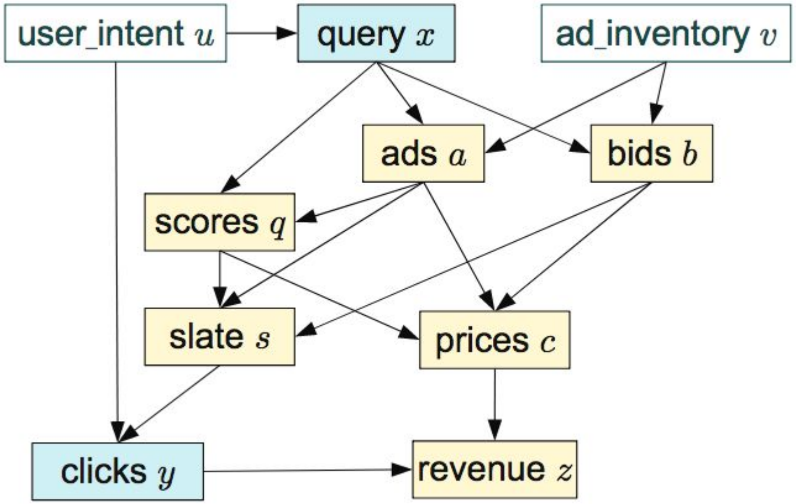
Collider

introduces spurious association when conditioned on

$x \rightarrow z \leftarrow y$

Causal diagrams in real life

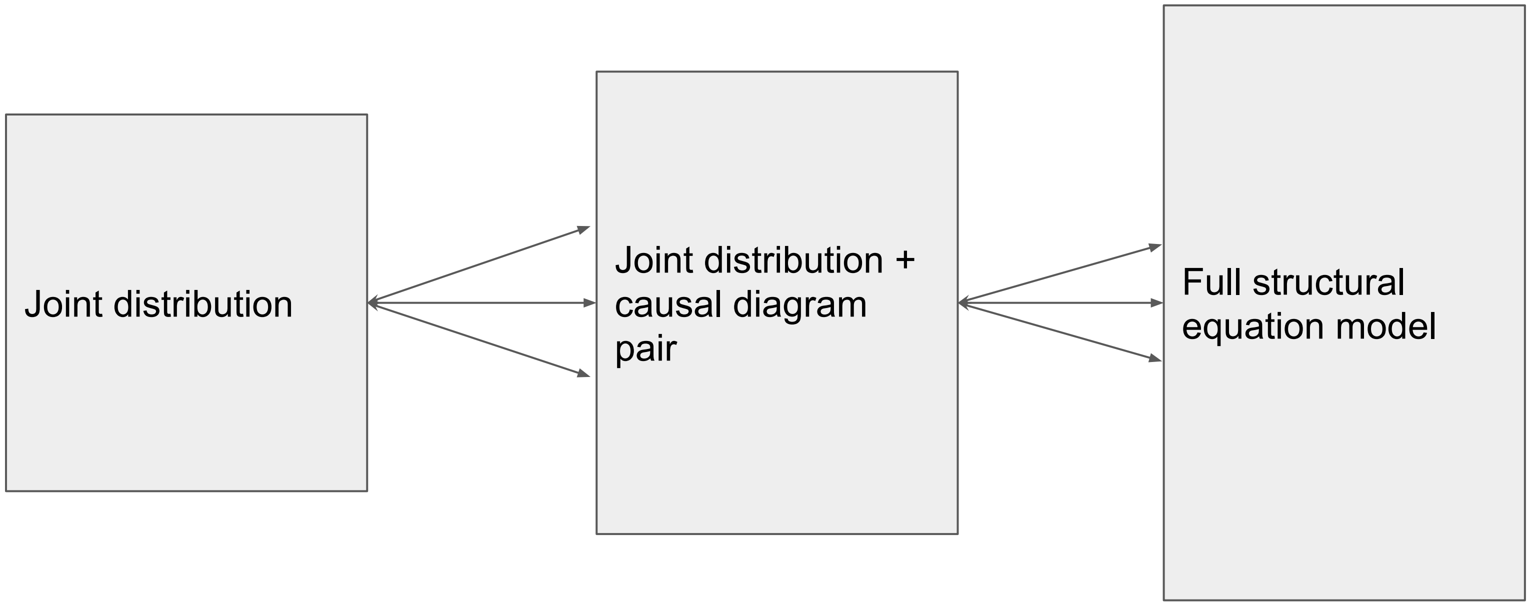
BOTTOU, PETERS, ET AL.



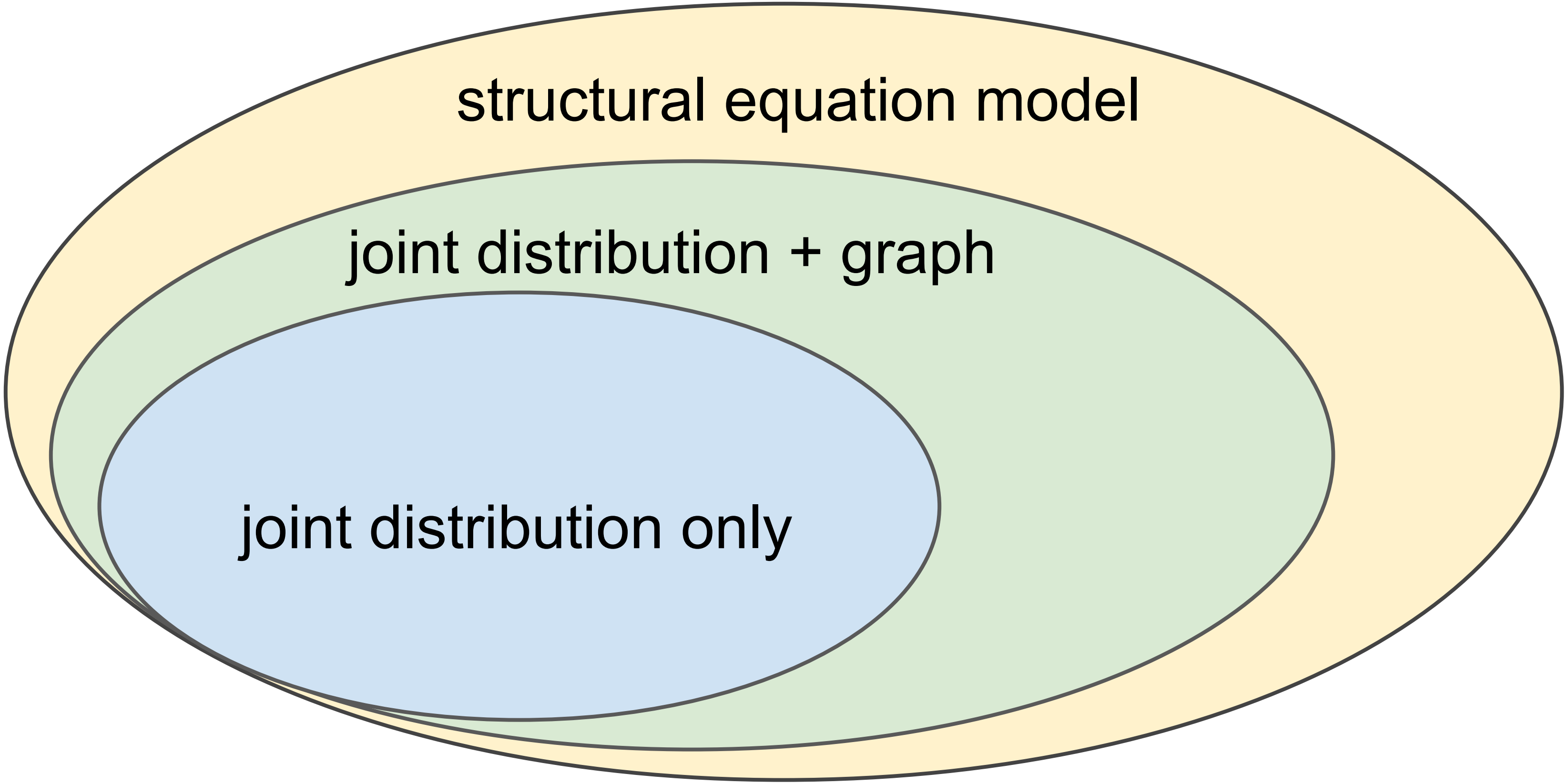
Interventions

do-calculus notation

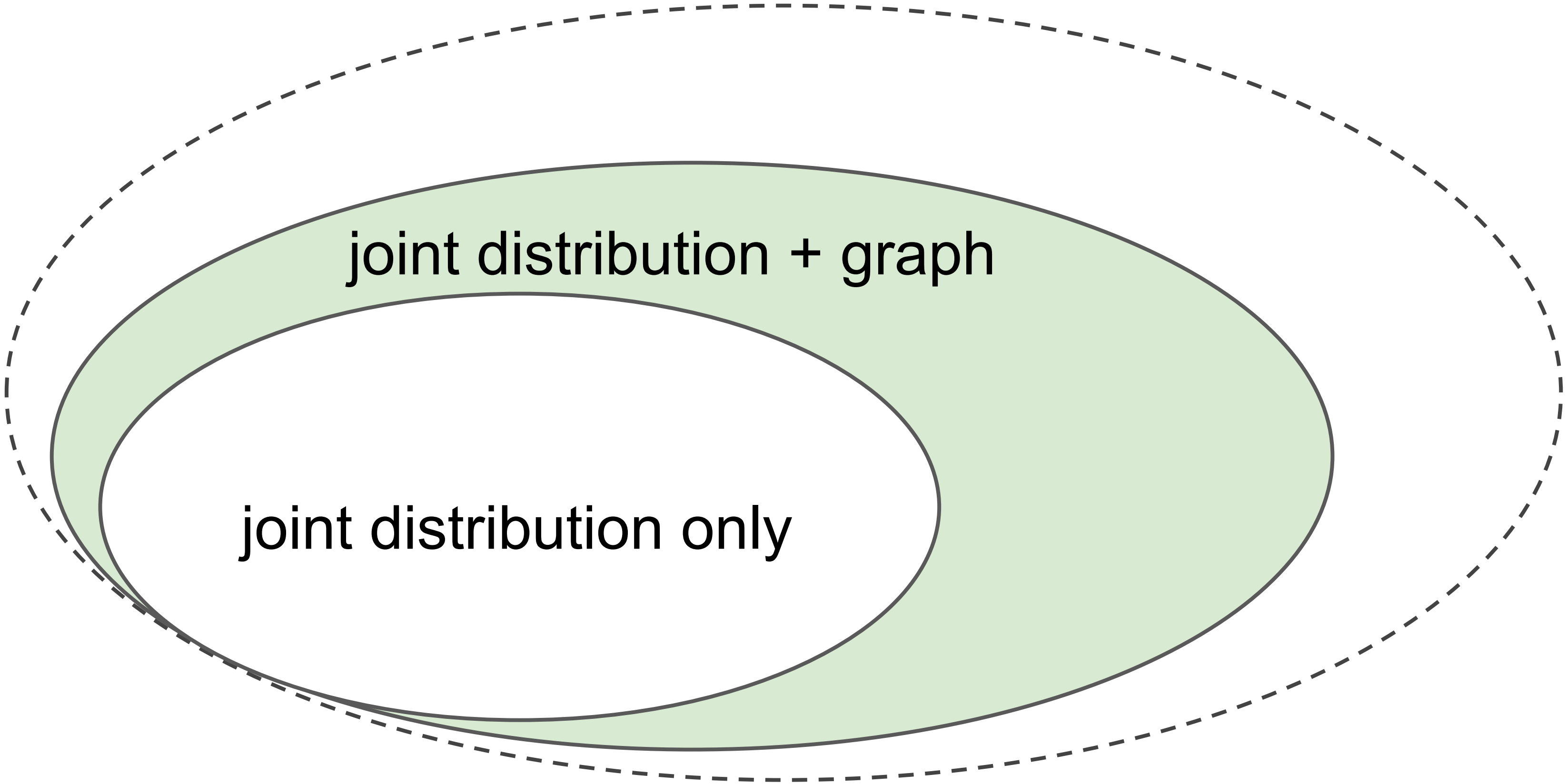
Different Granularity of Description



The questions you can ask/answer



Problems you can solve if you have a causal graph

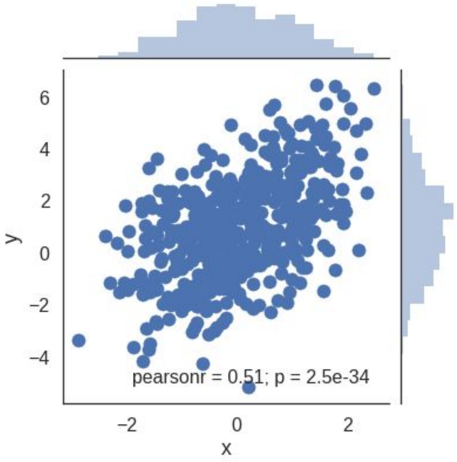


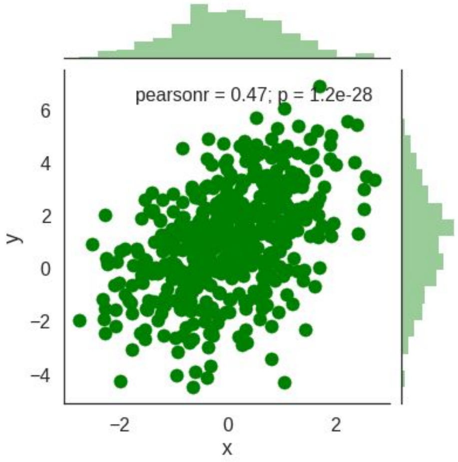
Toy example: three scripts

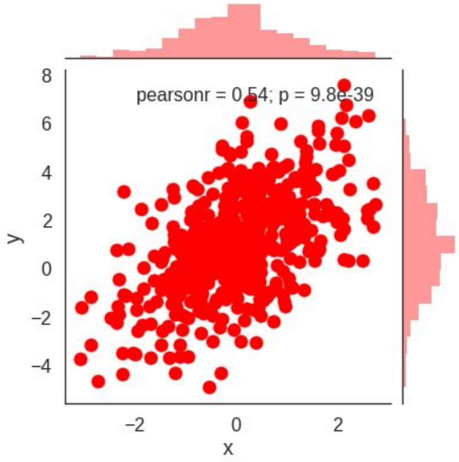
```
x = randn()
y = x + 1 + sqrt(3)*randn()
```

```
y = 1 + 2*randn()
x = (y-1)/4 + sqrt(3)*randn()/2
```

```
z = randn()
y = z + 1 + sqrt(3)*randn()
x = z
```





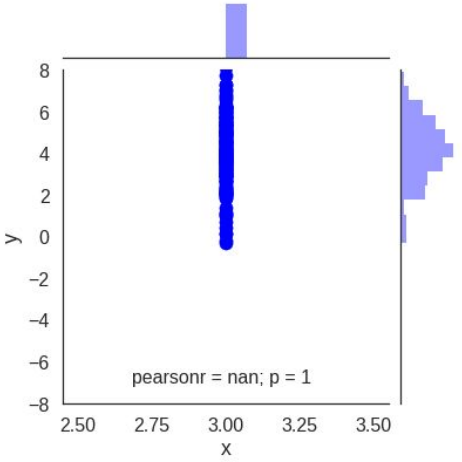


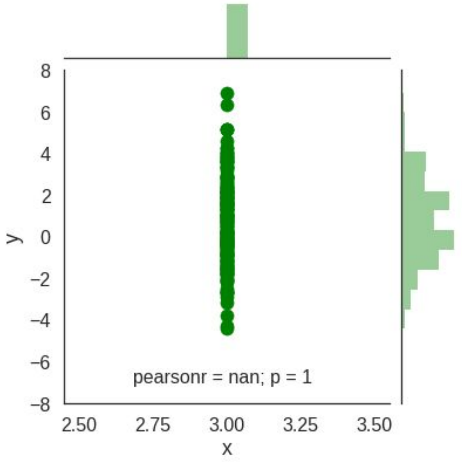
Intervention: what if I interfere and set the value x=3

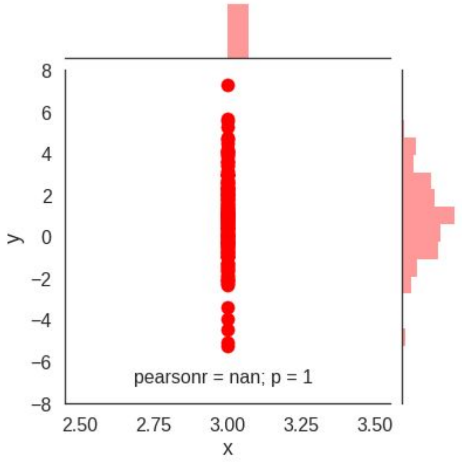
```
x = randn()
x = 3
y = x + 1 + sqrt(3)*randn()
x = 3
```

```
y = 1 + 2*randn()
x = 3
x = (y-1)/4 + sqrt(3)*randn()/2
x = 3
```

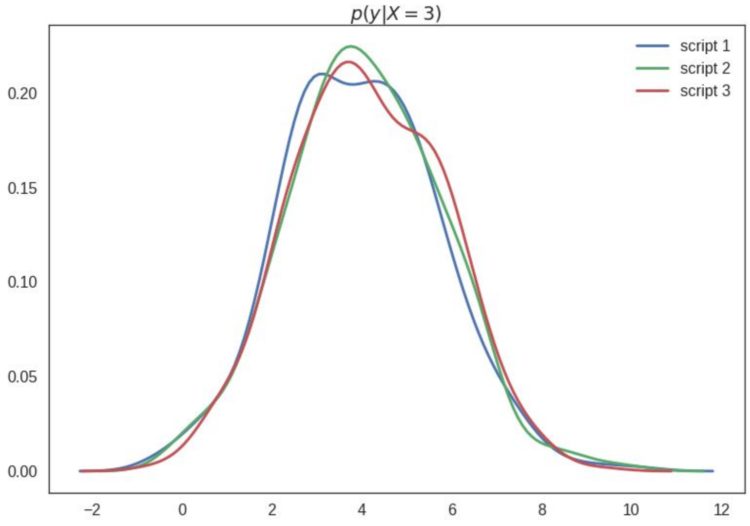
```
z = randn()
x = 3
x = z
x = 3
y = z + 1 + sqrt(3)*randn()
x = 3
```

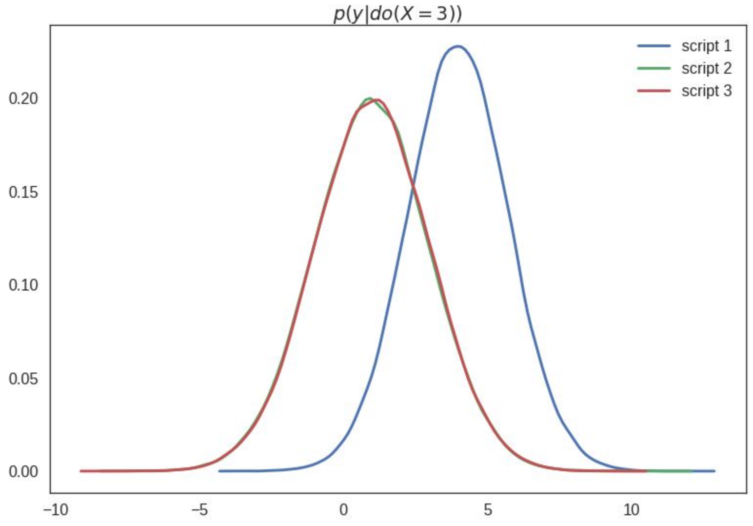






Three scripts: “intervention conditionals”



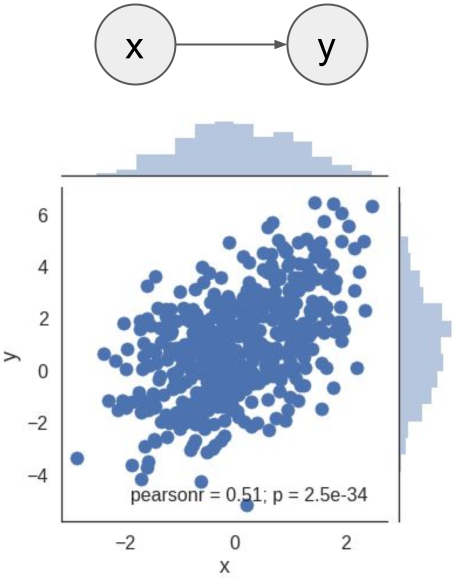


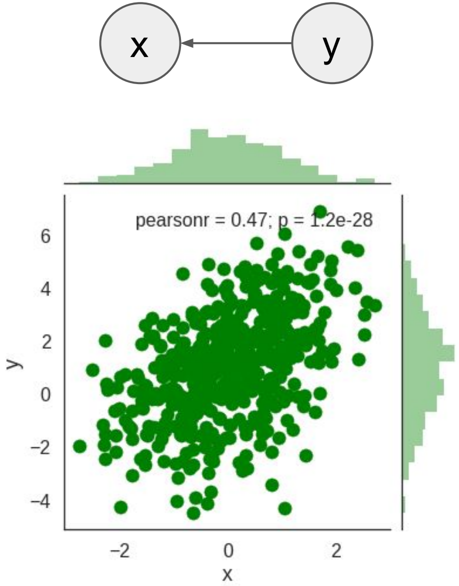
Toy example: three scripts

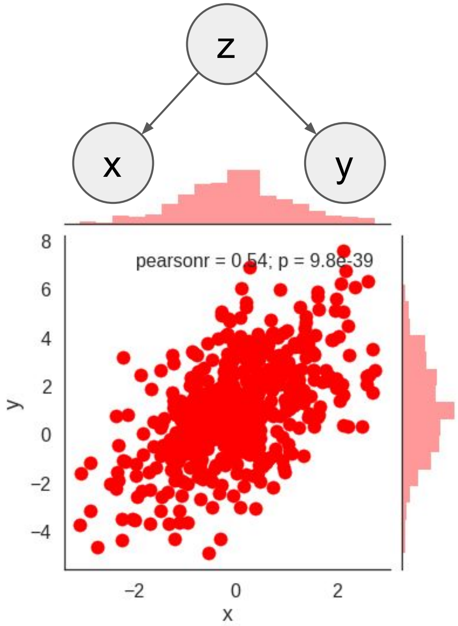
```
x = randn()
y = x + 1 + sqrt(3)*randn()
```

```
y = 1 + 2*randn()
x = (y-1)/4 + sqrt(3)*randn()/2
```

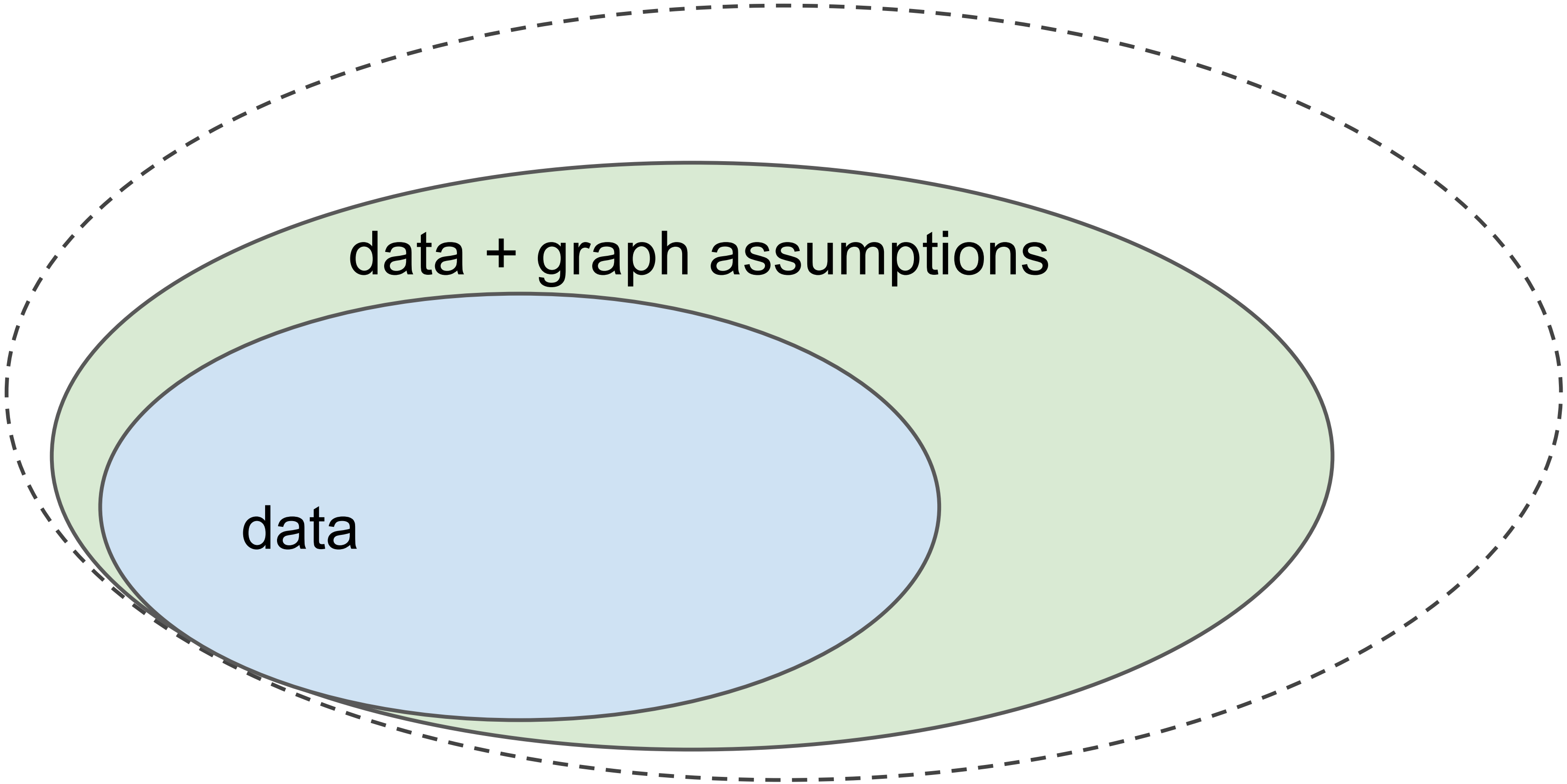
```
z = randn()
y = z + 1 + sqrt(3)*randn()
x = z
```







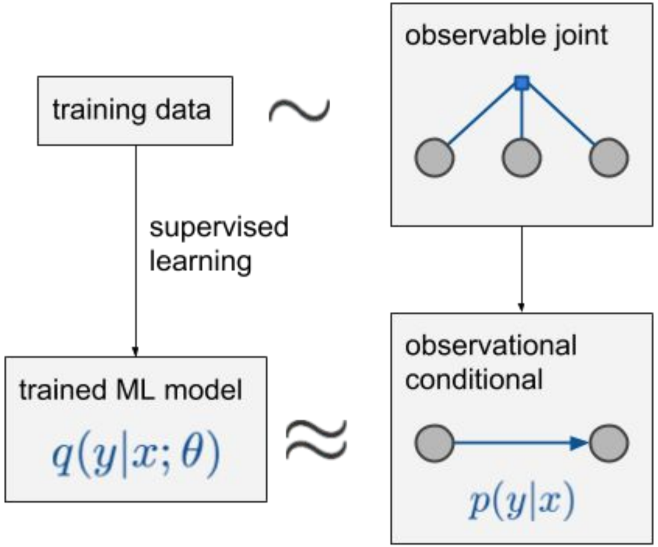
Causal graph -> predict behaviour under intervention



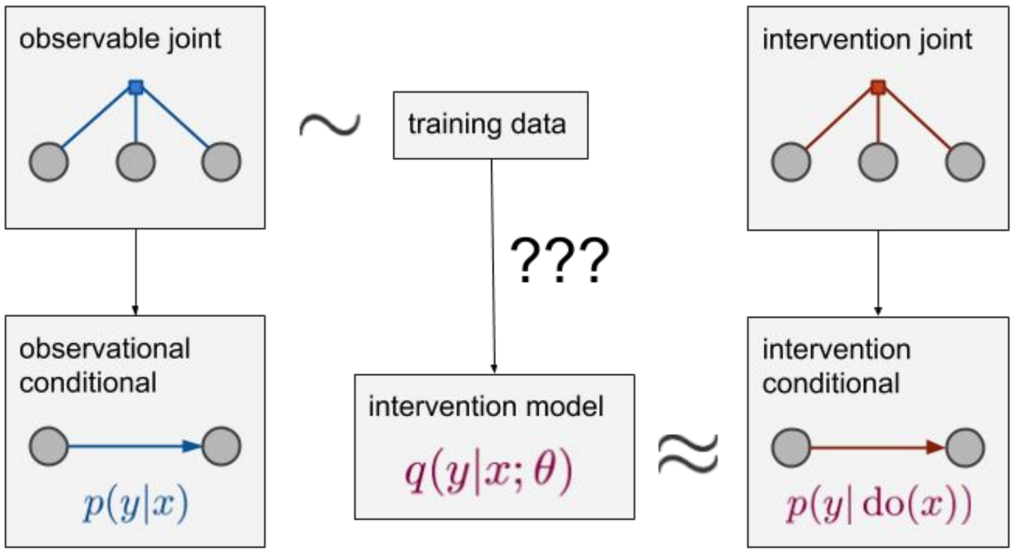
Lesson so far

- usually in ML deals with a joint distribution
- e.g. supervised learning approximates p(y|x) from samples
- the joint distribution obscures details of data generating process
- joint distribution does not allow you to predict effects of interventions
- p(y| do(x)): distribution of y under intervention on x
- p(y | x) and p(y | do(x)) are not always the same
- multiple p(y|do(x)) are consistent with the same p(x,y)
- finding p(y|do(x)) from p(x,y) is under-specified or ill-posed problem
- to solve such problems one needs additional constraints or priors
- **Can the graph structure be sufficient constraint?**

What we do in ML: estimate p(y|x) from data

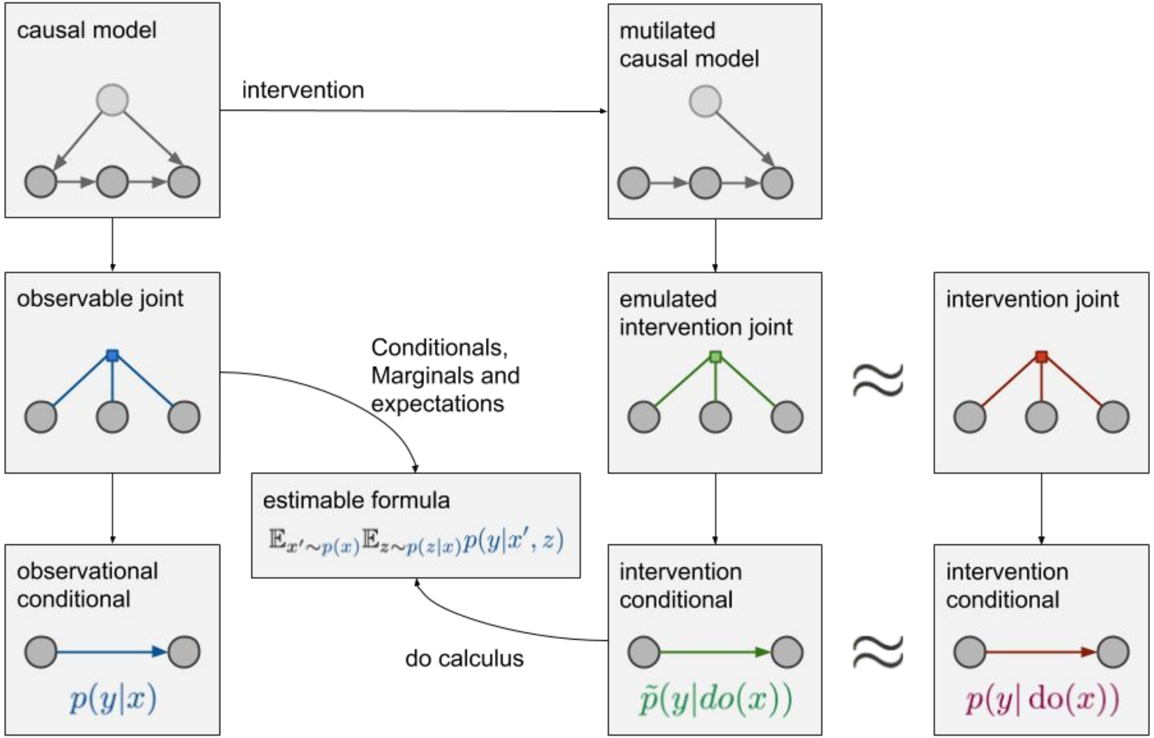


What we want to do in Causal ML



Thinking with causal models

do-calculus

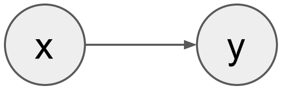

causal model
intervention
mutilated
causal model
observable joint
Conditionals,
Marginals and
expectations
emulated
intervention joint
≈
intervention joint
estimable formula
$\mathbb{E}_{x' \sim p(x)} \mathbb{E}_{z \sim p(z|x)} p(y|x', z)$
observational
conditional
$p(y|x)$
do calculus
intervention
conditional
$\tilde{p}(y|do(x))$
≈
intervention
conditional
$p(y|\,\mathrm{do}(x))$

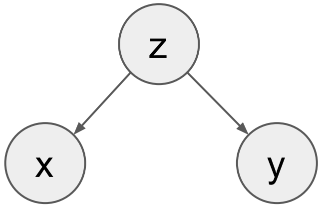
Toy example: three scripts

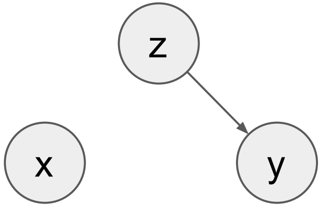
```
x = randn()
y = x + 1 + sqrt(3)*randn()
```

```
y = 1 + 2*randn()
x = (y-1)/4 + sqrt(3)*randn()/2
```

```
z = randn()
y = z + 1 + sqrt(3)*randn()
x = z
```





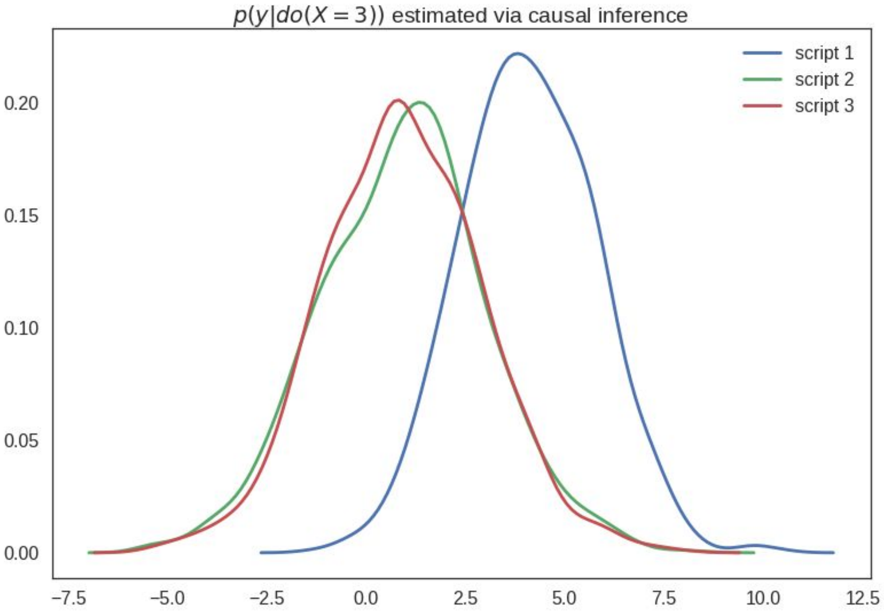


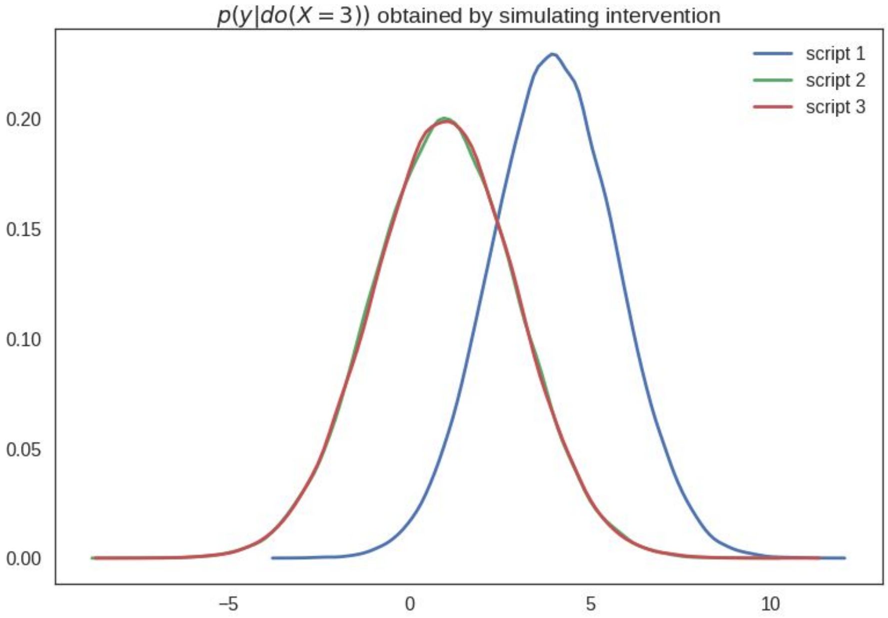
$$P(y|do(X)) = p(y|x)$$

$$P(y|do(X)) = p(y)$$

$$P(y|do(X)) = p(y)$$

Three scripts: “intervention conditionals”





estimated from joint + causal diagram

actually running the experiment

do-calculus

Rule 1 (Insertion/deletion of observations):

$$P(y|do(x), z, w) = P(y|do(x), w) \text{ if } Y \perp\!\!\!\perp Z|X, W||X$$

Rule 2 (Action/observation exchange):

$$P(y|do(x), do(z), w) = P(y|do(x), z, w) \text{ if } Y \perp\!\!\!\perp I_Z|X, Z, W||X$$

Rule 3 (Insertion/deletion of actions):

$$P(y|do(x), do(z), w) = P(y|do(x), w) \text{ if } Y \perp\!\!\!\perp I_Z|X, W||X$$

Rule 4 (Marginalization/sum-rule):

$$P(y|do(x), w) = \textstyle\sum_z P(y, z|do(x), w)$$

Rule 5 (Conditioning):

$$P(y|do(x), z, w) = \frac{P(y, z|do(x), w)}{\sum_y P(y, z|do(x), w)}$$

Rule 6 (product/chain-rule):

$$P(y, z|do(x), w) = P(y|do(x), w, z)P(z|do(x), w)$$

Reference: Hyttinen et al "Do Calculus when the True Graph is Known"

Causal inference by do-calculus

Step 0: Recognise you have a causal inference problem

Step 1: Find out what “causal query” you want to estimate, e.g. $p(y| do(x))$

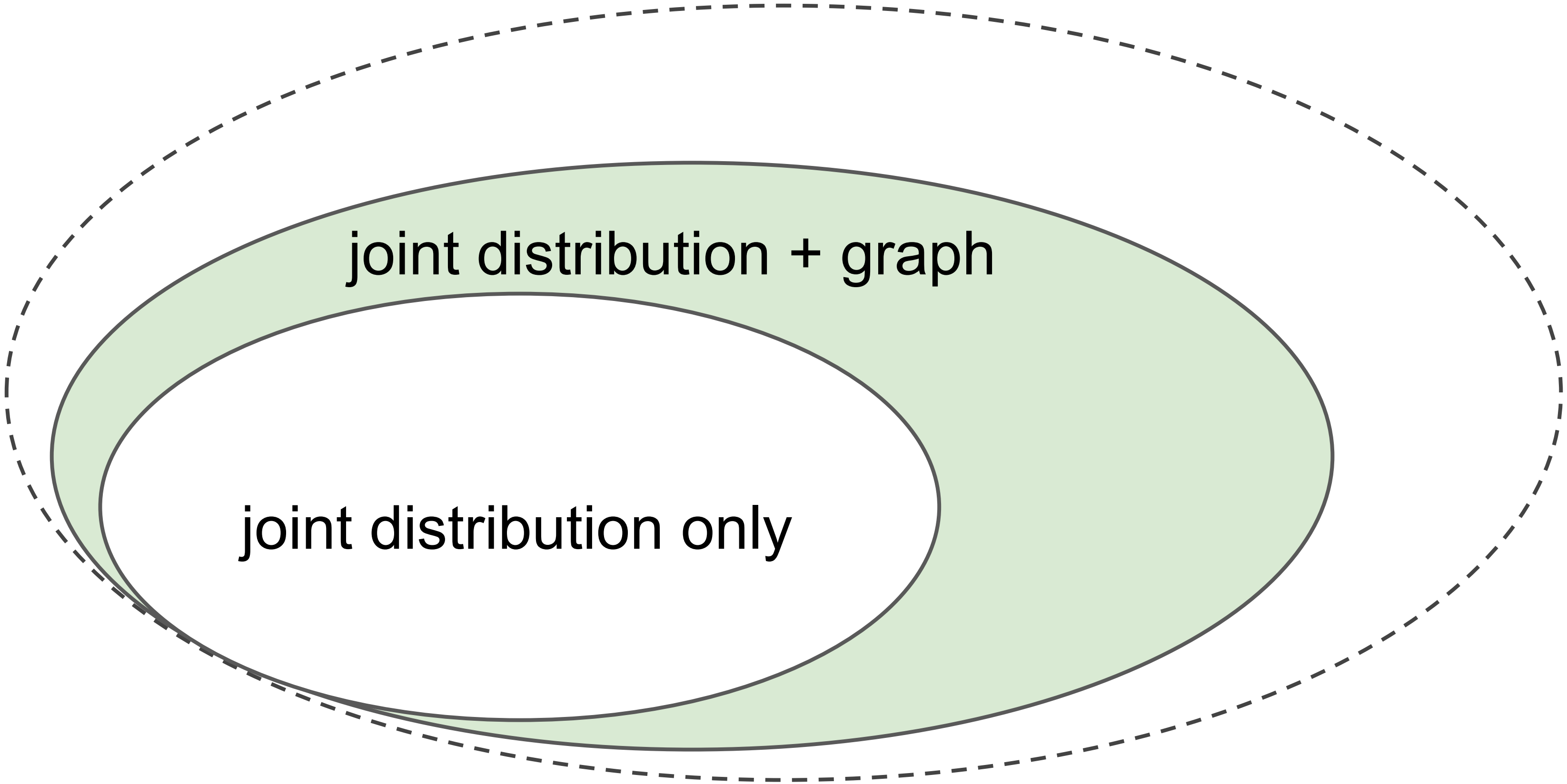
Step 2: Draw causal graph (assumptions, and/or causal discovery)

Step 3: Apply rules of do-calculus and probabilities until you get rid of all do’s

(if you can’t your problem is *non-identifiable*)

Step 4: Find a smart way to estimate approximate from data **(machine learning)**

Problems you can solve if you have a causal graph



Second Lecture

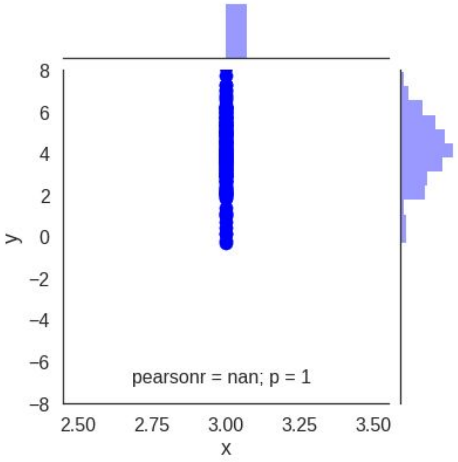
Bugfixing

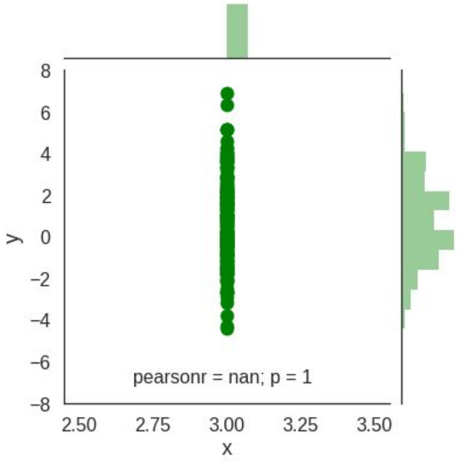
Intervention: what if I interfere and set the value x=3

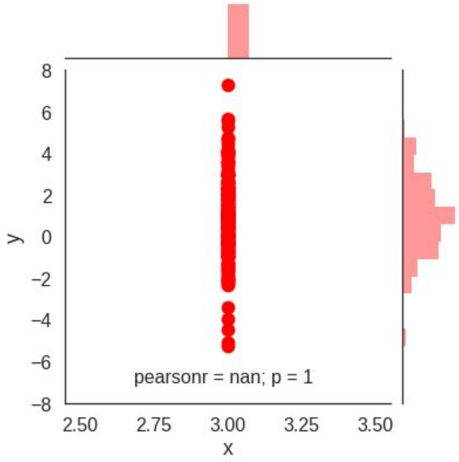
x = randn()
x = 3
y = x + 1 + sqrt(3)*randn()
x = 3

y = 1 + 2*randn()
x = 3
x = (y-1)/4 + sqrt(3)*randn()/2
x = 3

z = randn()
x = 3
x = z
x = 3
y = x + 1 + sqrt(3)*randn()
x = 3







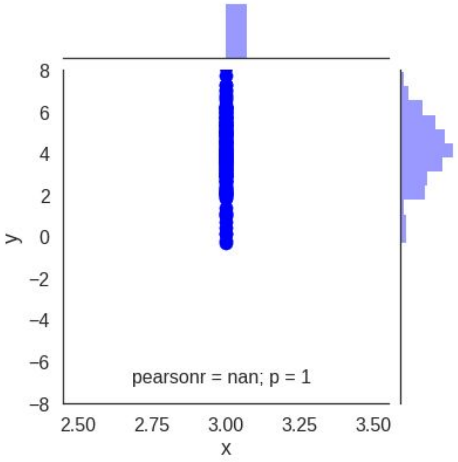
Intervention: what if I interfere and set the value x=3

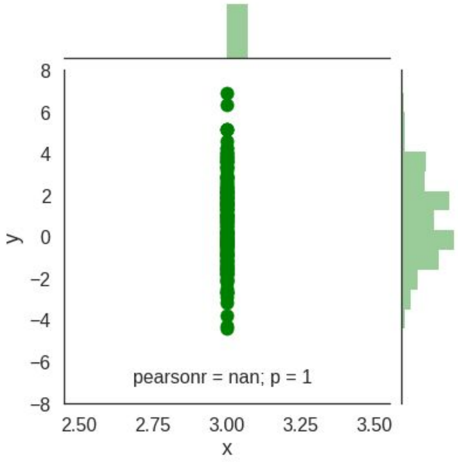
x = randn()
x = 3
y = x + 1 + sqrt(3)*randn()
x = 3

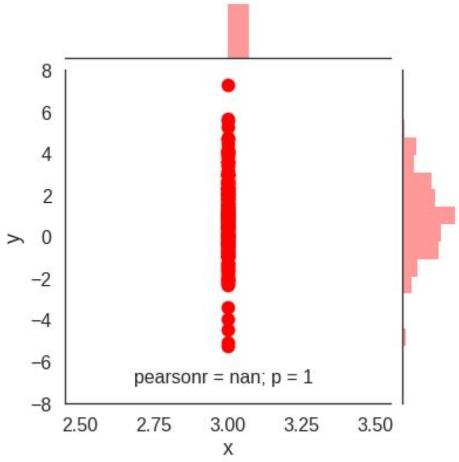
y = 1 + 2*randn()
x = 3
x = (y-1)/4 + sqrt(3)*randn()/2
x = 3

z = randn()
x = 3
x = z
x = 3
y = z + 1 + sqrt(3)*randn()
x = 3

correction



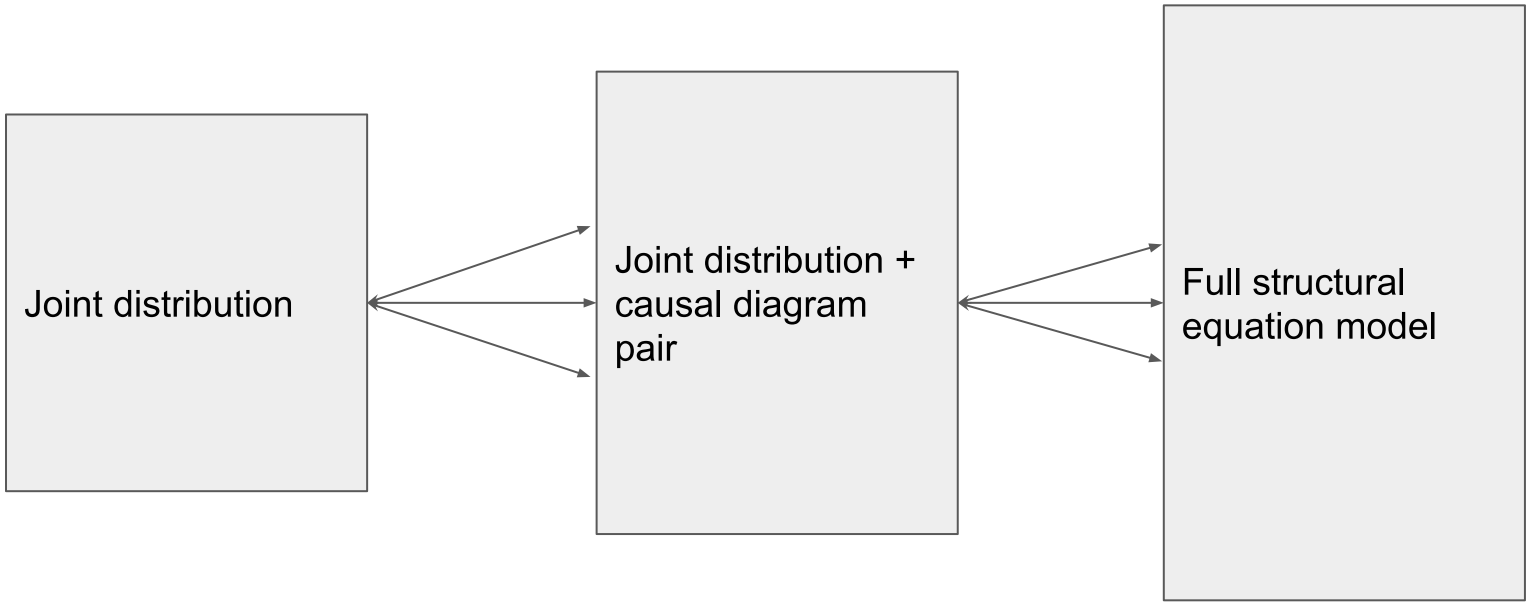




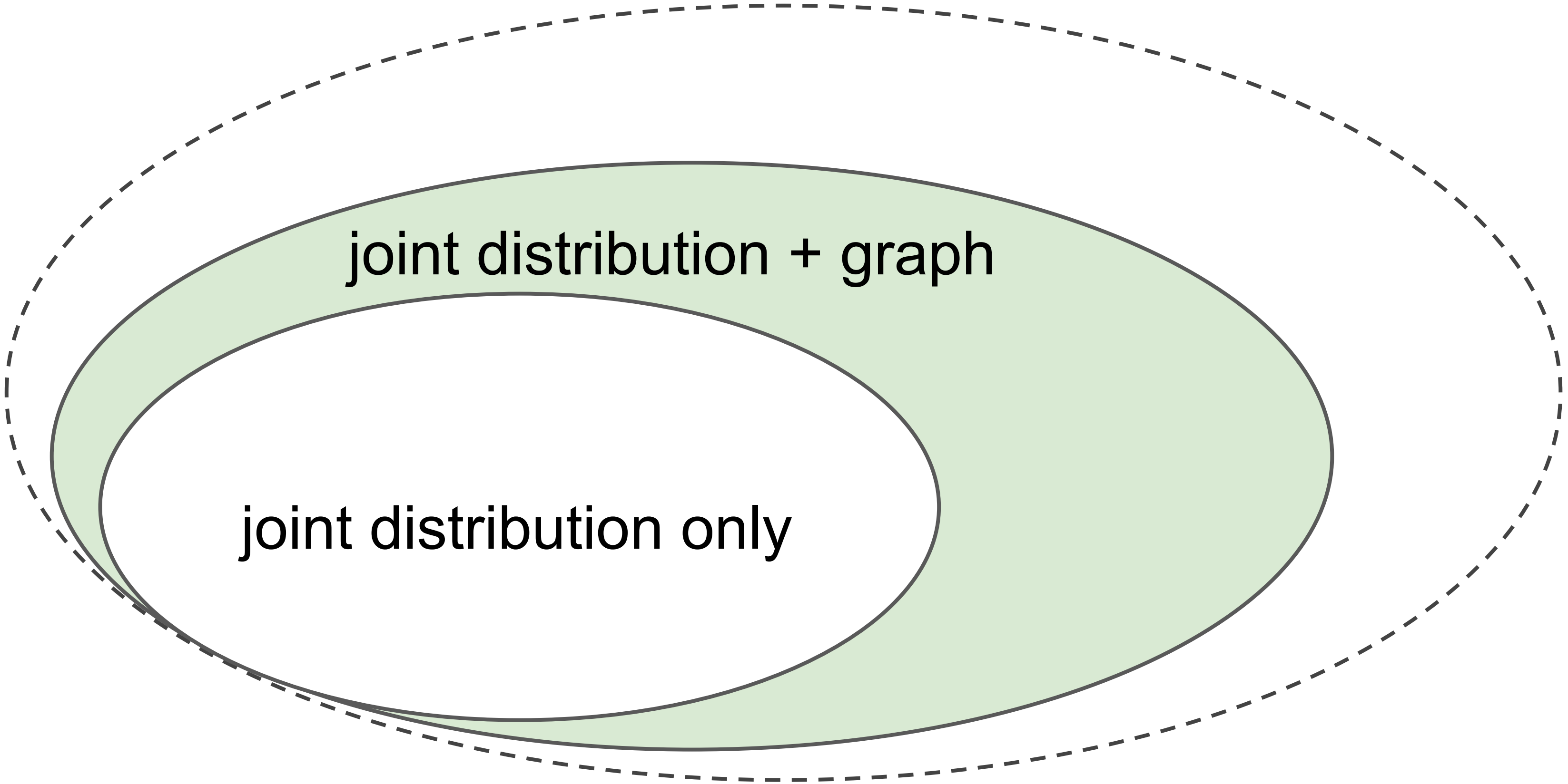
Apology: availability bias example

Recap

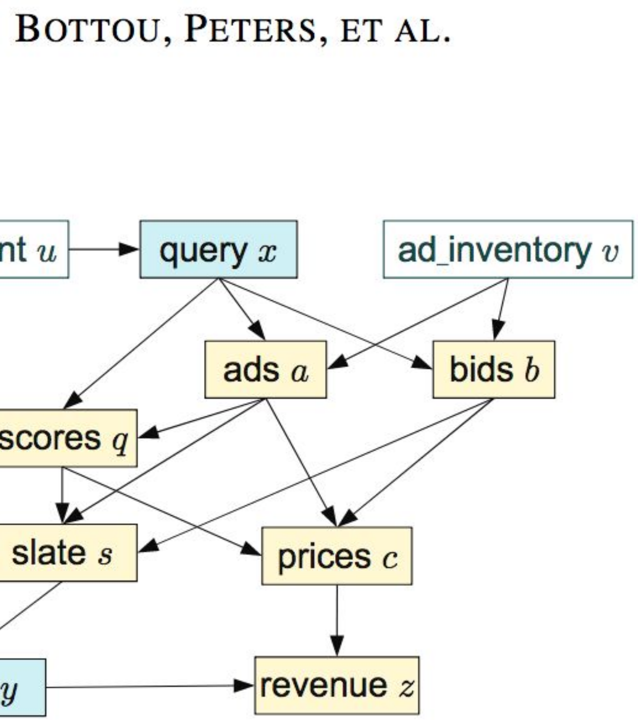
Different Granularity of Description



Problems you can solve if you have a causal graph



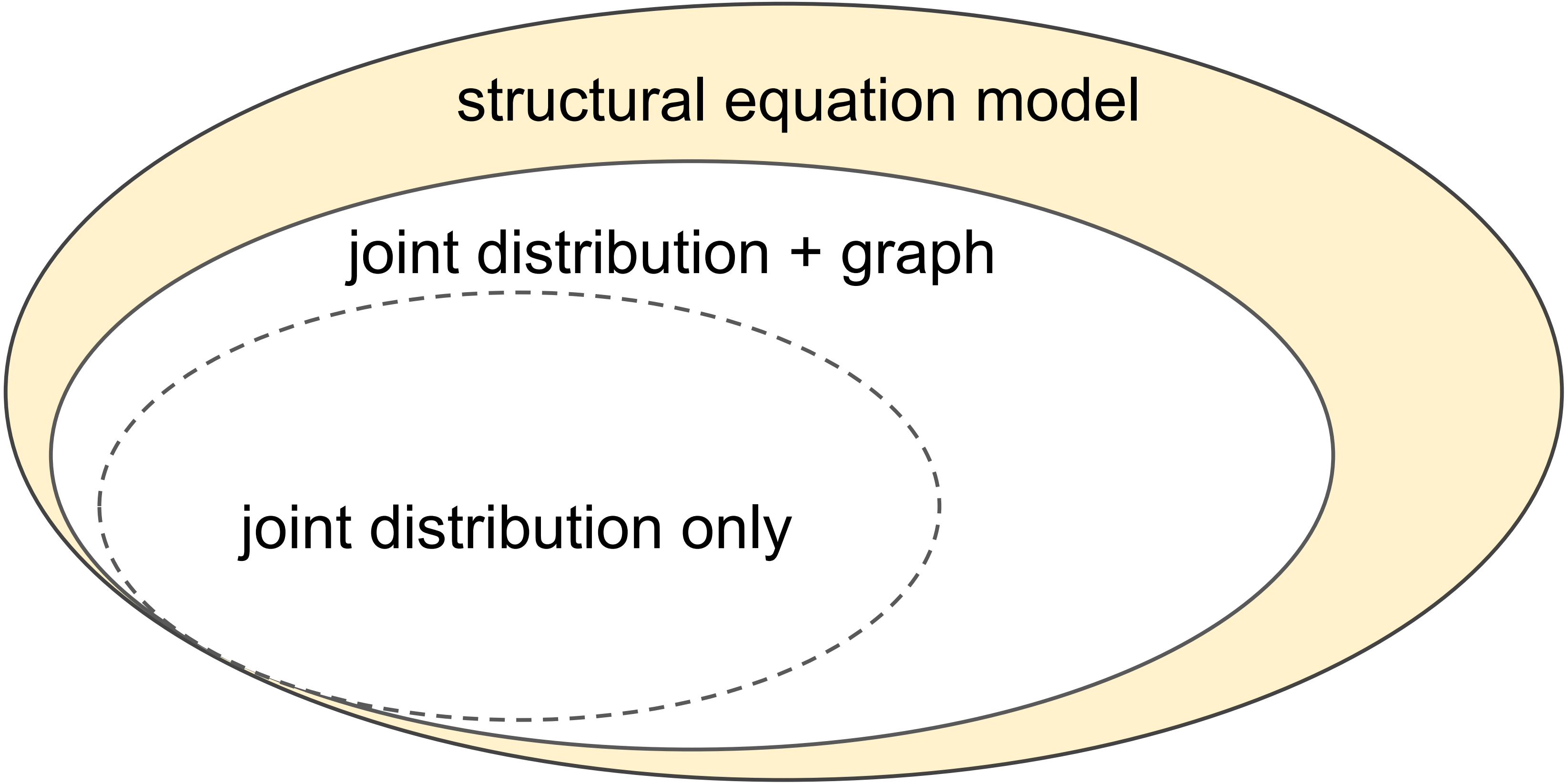
Causal diagram for online advertising



Corresponding structural equation model

$$
\begin{array}{rclll}
x & = & f_1(u, \varepsilon_1) & & \text{Query context } x \text{ from user intent } u. \\
a & = & f_2(x, v, \varepsilon_2) & & \text{Eligible ads } (a_i) \text{ from query } x \text{ and inventory } v. \\
b & = & f_3(x, v, \varepsilon_3) & & \text{Corresponding bids } (b_i). \\
q & = & f_4(x, a, \varepsilon_4) & & \text{Scores } (q_{i,p}, R_p) \text{ from query } x \text{ and ads } a. \\
s & = & f_5(a, q, b, \varepsilon_5) & & \text{Ad slate } s \text{ from eligible ads } a \text{, scores } q \text{ and bids } b. \\
c & = & f_6(a, q, b, \varepsilon_6) & & \text{Corresponding click prices } c. \\
y & = & f_7(s, u, \varepsilon_7) & & \text{User clicks } y \text{ from ad slate } s \text{ and user intent } u. \\
z & = & f_8(y, c, \varepsilon_8) & & \text{Revenue } z \text{ from clicks } y \text{ and prices } c.
\end{array}
$$

What problems you can solve with S.E.M.



Counterfactuals

Counterfactuals

David Blei's example:

Given that Hillary Clinton did not win the election

and that he did not visit Michigan 3 days before the election

and everything else we observed about the situation

what can we say about the probability of her winning the election

had she visited that state 2 days before the election

Counterfactuals

Counterfactual fairness

given that you are a woman

and that your promotion was rejected

and everything else known about you

would your promotion have been successful

if you were a man

Counterfactuals

My example

Given that I have a beard

and that I have a PhD

and everything else known about me

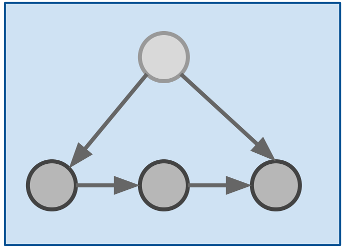
what can we say about the probability of my not getting a PhD

had I decided to shave my beard 5 years ago

🧔	💍	💪	🎓
0	1	1	0
0	0	1	1
1	0	1	0
1	1	1	1
1	1	0	0

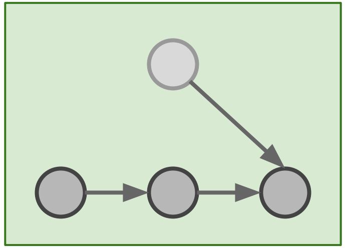
p(🎓 | 🧔 = 0)

observed, factual



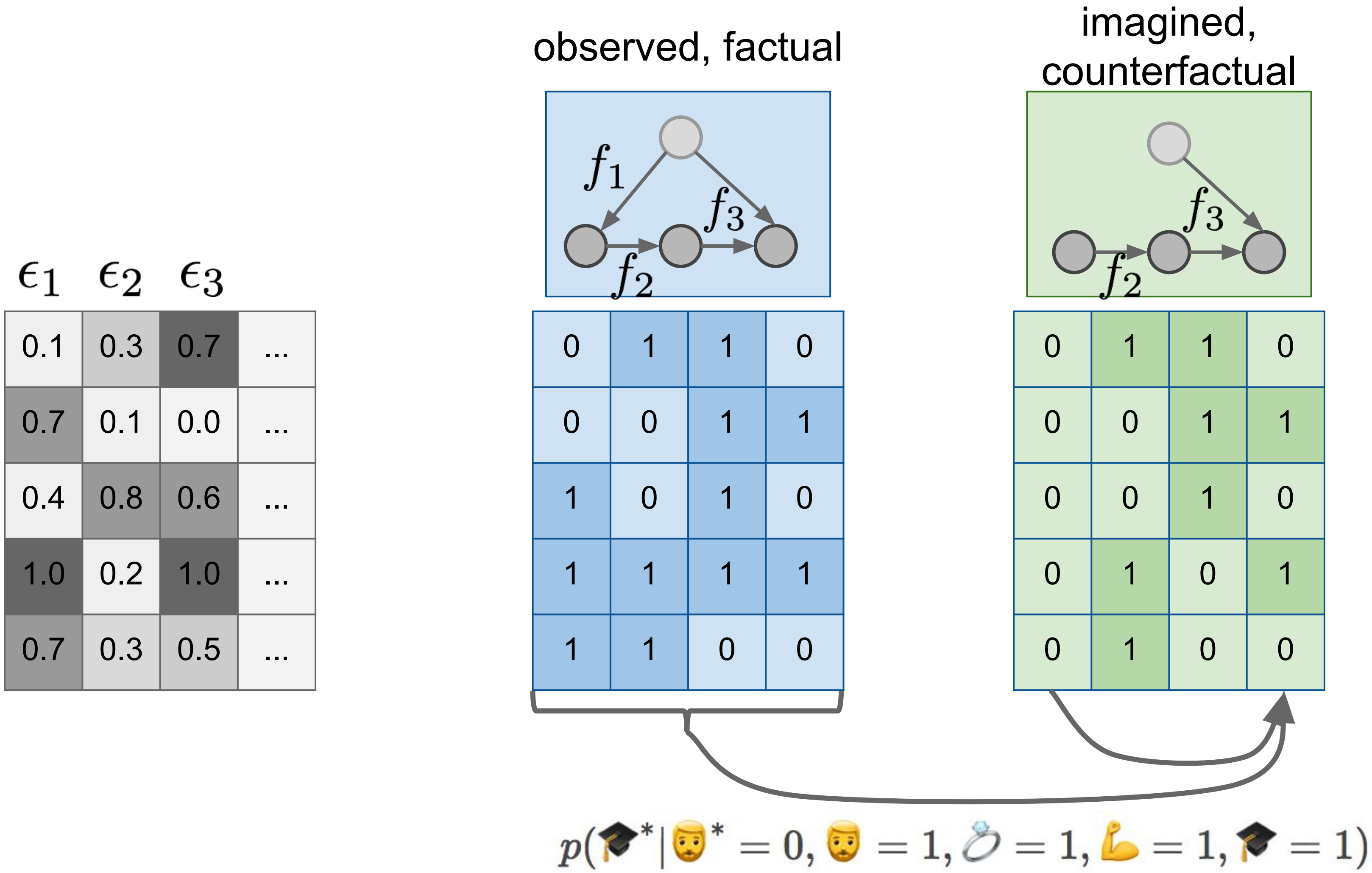
0	1	1	0
0	0	1	1
1	0	1	0
1	1	1	1
1	1	0	0

imagined, counterfactual



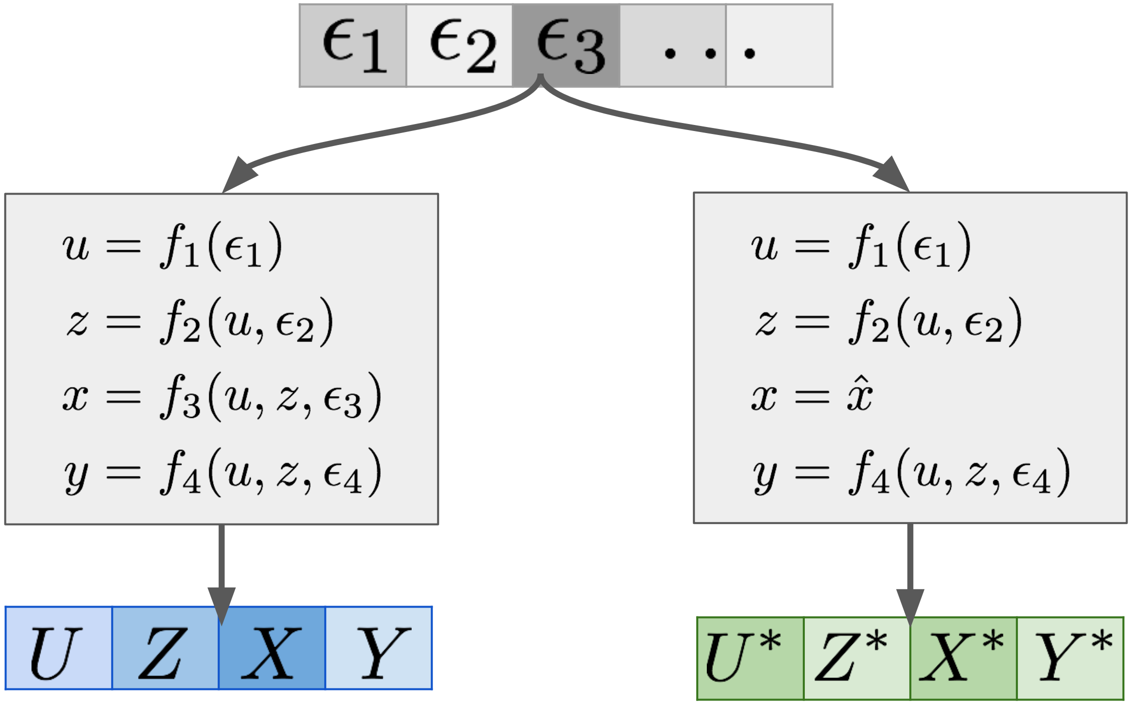
0	1	1	0
0	0	1	1
0	0	1	0
0	1	0	1
0	1	0	0

$p($🎓$|do($👨$=0))$



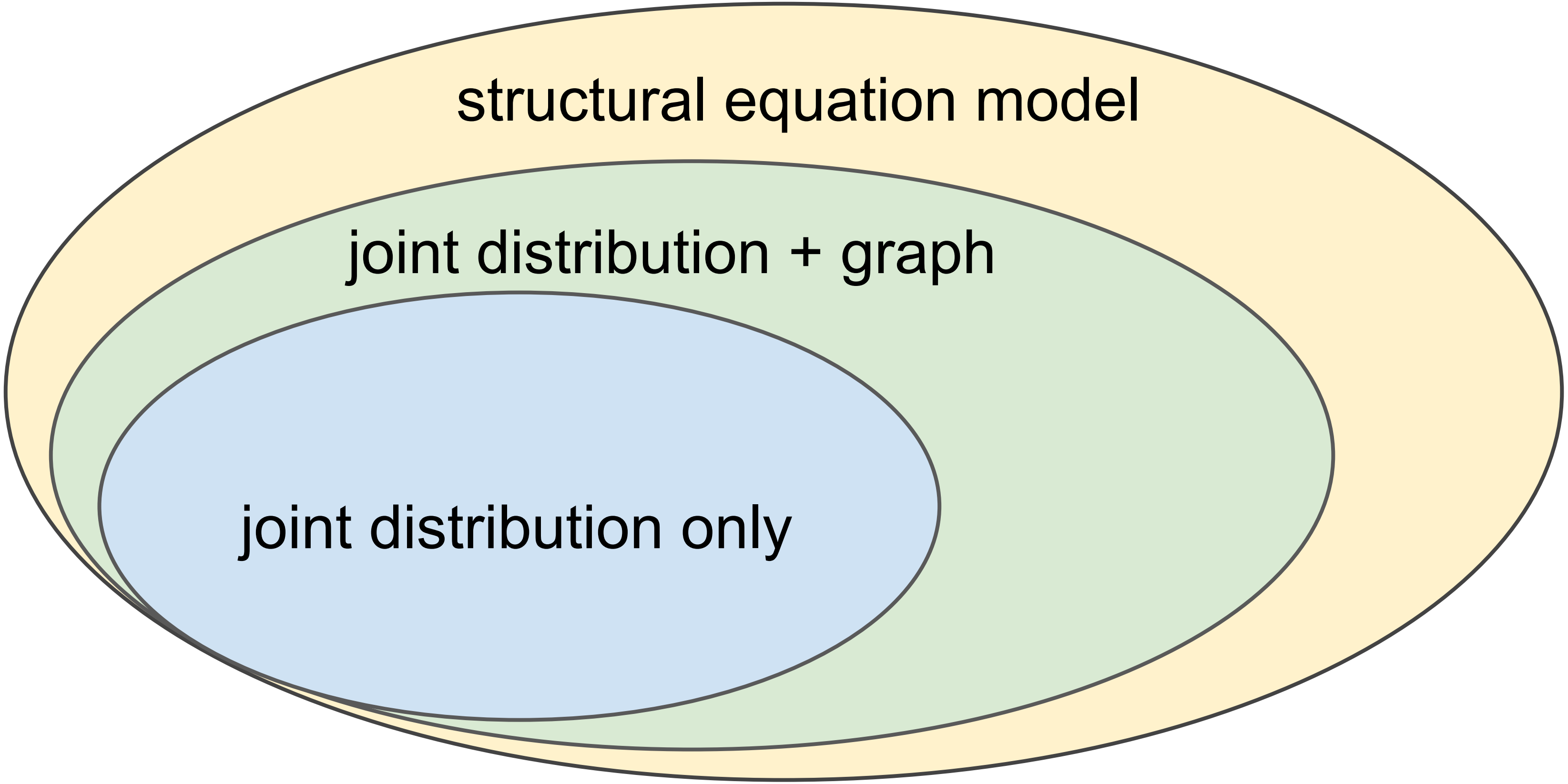
ϵ_1 ϵ_2 ϵ_3
0.1 0.3 0.7 ...
0.7 0.1 0.0 ...
0.4 0.8 0.6 ...
1.0 0.2 1.0 ...
0.7 0.3 0.5 ...
observed, factual
f_1 f_2 f_3
0 1 1 0
0 0 1 1
1 0 1 0
1 1 1 1
1 1 0 0
imagined, counterfactual
f_2 f_3
0 1 1 0
0 0 1 1
0 0 1 0
0 1 0 1
0 1 0 0
$p($🎓$^*|$👨$^* = 0,$ 👨 $= 1,$ 💍 $= 1,$ 💪 $= 1,$ 🎓 $= 1)$

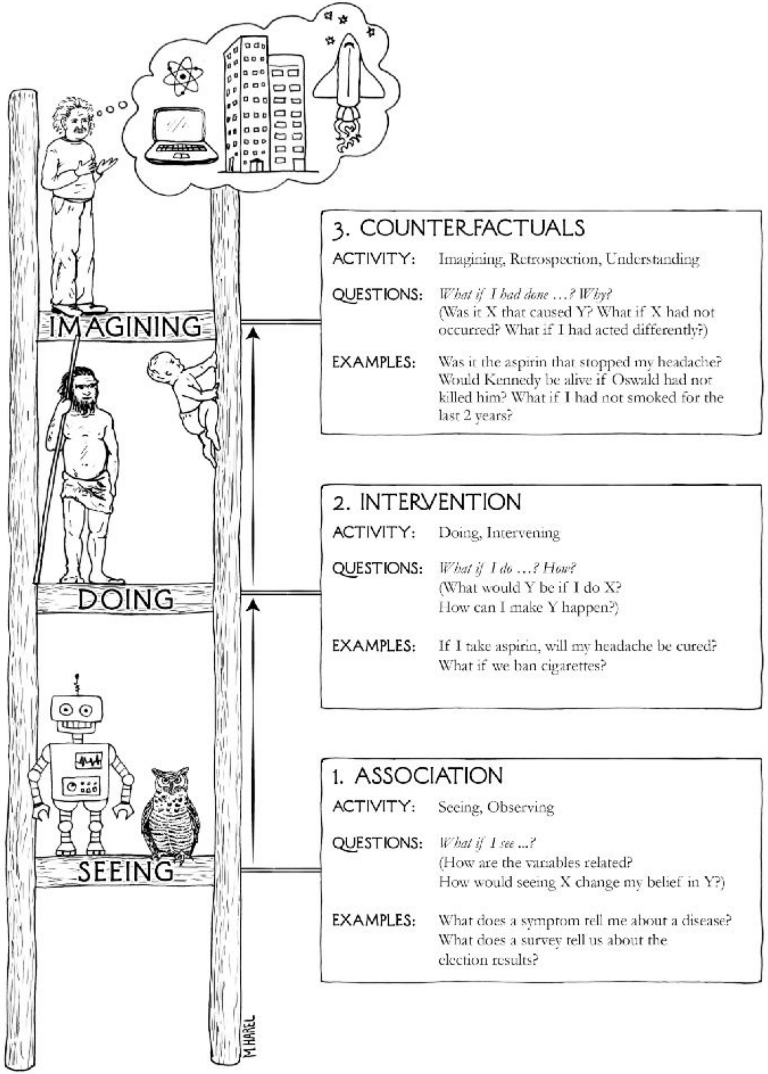
Another way to draw this

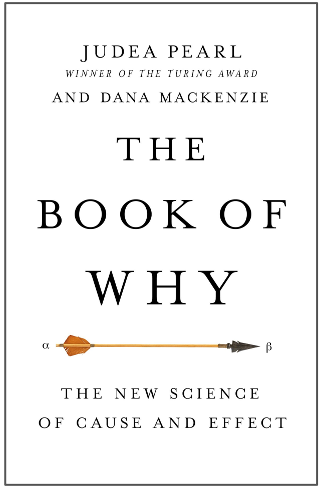


$$p(y^* | X^* = \hat{x}, X = x, Y = y, U = u, V = v)$$

The questions you can ask/answer




3. COUNTERFACTUALS
ACTIVITY: Imagining, Retrospection, Understanding
QUESTIONS: What if I had done ...? Why?
(Was it X that caused Y? What if X had not occurred? What if I had acted differently?)
EXAMPLES: Was it the aspirin that stopped my headache? Would Kennedy be alive if Oswald had not killed him? What if I had not smoked for the last 2 years?
IMAGINING
2. INTERVENTION
ACTIVITY: Doing, Intervening
QUESTIONS: What if I do ...? How?
(What would Y be if I do X? How can I make Y happen?)
EXAMPLES: If I take aspirin, will my headache be cured? What if we ban cigarettes?
DOING
1. ASSOCIATION
ACTIVITY: Seeing, Observing
QUESTIONS: What if I see ...?
(How are the variables related? How would seeing X change my belief in Y?)
EXAMPLES: What does a symptom tell me about a disease? What does a survey tell us about the election results?
SEEING
M.HAREL


JUDEA PEARL
WINNER OF THE TURING AWARD
AND DANA MACKENZIE
THE BOOK OF WHY
α β
THE NEW SCIENCE OF CAUSE AND EFFECT

Counterfactual Reasoning in Online Systems

Journal of Machine Learning Research 14 (2013) 3207-3260 Submitted 9/12; Revised 3/13; Published 11/13

Counterfactual Reasoning and Learning Systems: The Example of Computational Advertising

Léon Bottou LEON@BOTTOU.ORG
Microsoft
1 Microsoft Way
Redmond, WA 98052, USA

Jonas Peters* PETERS@STAT.MATH.ETHZ.CH
Max Planck Institute
Spemannstraße 38
72076 Tübingen, Germany

Joaquin Quiñonero-Candela† JQUINONERO@GMAIL.COM
Denis X. Charles CDX@MICROSOFT.COM
D. Max Chickering DMAX@MICROSOFT.COM
Elon Portugaly ELONP@MICROSOFT.COM
Dipankar Ray DIPANRAY@MICROSOFT.COM
Patrice Simard PATRICE@MICROSOFT.COM
Ed Snelson EDSNELSO@MICROSOFT.COM
Microsoft
1 Microsoft Way
Redmond, WA 98052, USA

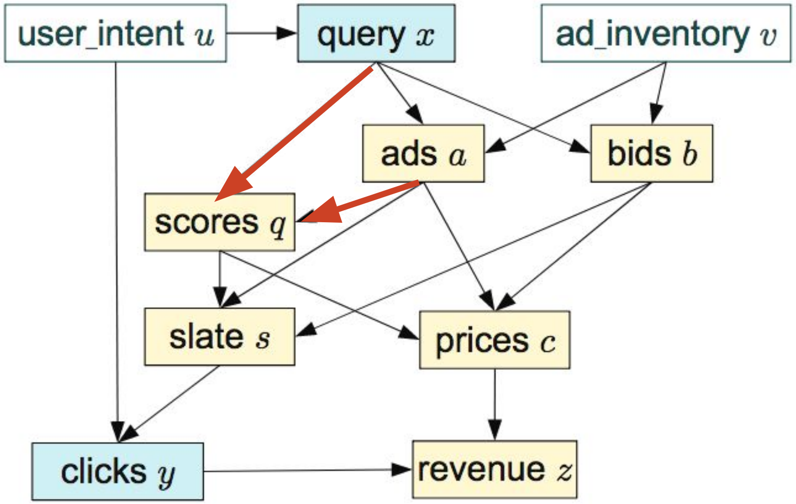
Abstract

This work shows how to leverage causal inference to understand the behavior of complex learning systems interacting with their environment and predict the consequences of changes to the system. Such predictions allow both humans and algorithms to select the changes that would have improved the system performance. This work is illustrated by experiments on the ad placement system associated with the Bing search engine.

Keywords: causation, counterfactual reasoning, computational advertising

causal diagram for online advertising

BOTTOU, PETERS, ET AL.



structural equation model for online advertising

$$
\begin{array}{rcll}
x & = & f_1(u,\varepsilon_1) & \text{Query context } x \text{ from user intent } u. \\
a & = & f_2(x,v,\varepsilon_2) & \text{Eligible ads } (a_i) \text{ from query } x \text{ and inventory } v. \\
b & = & f_3(x,v,\varepsilon_3) & \text{Corresponding bids } (b_i). \\
\boxed{q = f_4(x,a,\varepsilon_4)} & & & \text{Scores } (q_{i,p}, R_p) \text{ from query } x \text{ and ads } a. \\
s & = & f_5(a,q,b,\varepsilon_5) & \text{Ad slate } s \text{ from eligible ads } a\text{, scores } q \text{ and bids } b. \\
c & = & f_6(a,q,b,\varepsilon_6) & \text{Corresponding click prices } c. \\
y & = & f_7(s,u,\varepsilon_7) & \text{User clicks } y \text{ from ad slate } s \text{ and user intent } u. \\
z & = & f_8(y,c,\varepsilon_8) & \text{Revenue } z \text{ from clicks } y \text{ and prices } c.
\end{array}
$$

Markov factorization for online advertising

$$
\mathrm{P}\begin{pmatrix} u,v,x,a,b \\ q,s,c,y,z \end{pmatrix} = \left\{ \begin{array}{ll}
\mathrm{P}(u,v) & \text{Exogenous vars.} \\
\times\, \mathrm{P}(x \mid u) & \text{Query.} \\
\times\, \mathrm{P}(a \mid x,v) & \text{Eligible ads.} \\
\times\, \mathrm{P}(b \mid x,v) & \text{Bids.} \\
\boxed{\times\, \mathrm{P}(q \mid x,a)} & \text{Scores.} \\
\times\, \mathrm{P}(s \mid a,q,b) & \text{Ad slate.} \\
\times\, \mathrm{P}(c \mid a,q,b) & \text{Prices.} \\
\times\, \mathrm{P}(y \mid s,u) & \text{Clicks.} \\
\times\, \mathrm{P}(z \mid y,c) & \text{Revenue.}
\end{array} \right.
$$

What is the question we ask

“How will the system perform if we replace model M by model M ?”*

Usual solution in industry: A/B testing, a.k.a. randomized controlled trial

- Actually try model M* on a percentage of traffic
- **Difficult:** Requires you to develop model M* to production quality
- **Time consuming:** How many days until statistical significance
- **Inefficient:** You can only test so many models at once
- **Ethical?:** Providing different user experience to people
- **Doesn’t always work:** e.g. advertiser budget, training data distribution

Offline A/B testing

Can we predict the results without actually running an A/B test?

Precisely what **Causal Inference** and **Counterfactual Reasoning** are for

Phrase problem as “what if” question:

“How would the system have performed if, when the data was collected, we had replaced model M by model M ?”*

How did model M perform

$$
\begin{aligned}
\mathrm{P}(\omega) \quad = \quad & \mathrm{P}(u,v)\,\mathrm{P}(x\,|\,u)\,\mathrm{P}(a\,|\,x,v)\,\mathrm{P}(b\,|\,x,v)\,\mathrm{P}(q\,|\,x,a) \\
& \times \mathrm{P}(s\,|\,a,q,b)\,\mathrm{P}(c\,|\,a,q,b)\,\mathrm{P}(y\,|\,s,u)\,\mathrm{P}(z\,|\,y,c)\ .
\end{aligned}
$$

Click-through rate

$$
Y = \int_{\omega} y\,\mathrm{P}(\omega)\ .
$$

How did model M perform

$$\begin{aligned} \mathrm{P}(\omega) &= \mathrm{P}(u,v)\mathrm{P}(x\,|\,u)\mathrm{P}(a\,|\,x,v)\mathrm{P}(b\,|\,x,v)\mathrm{P}(q\,|\,x,a) \\ &\quad \times \mathrm{P}(s\,|\,a,q,b)\mathrm{P}(c\,|\,a,q,b)\mathrm{P}(y\,|\,s,u)\mathrm{P}(z\,|\,y,c)\,. \end{aligned}$$

Click-through rate

$$Y = \int_{\omega} y P(w) \approx \frac{1}{N} \sum_{n=1}^{N} y_n$$

How would model M* have performed

$$\boxed{\mathrm{P}^*}(\omega) = \mathrm{P}(u,v)\,\mathrm{P}(x\,|\,u)\,\mathrm{P}(a\,|\,x,v)\,\mathrm{P}(b\,|\,x,v)\,\boxed{\mathrm{P}^*(q\,|\,x,a)}$$
$$\times\,\mathrm{P}(s\,|\,a,q,b)\,\mathrm{P}(c\,|\,a,q,b)\,\mathrm{P}(y\,|\,s,u)\,\mathrm{P}(z\,|\,x,c)\,.$$

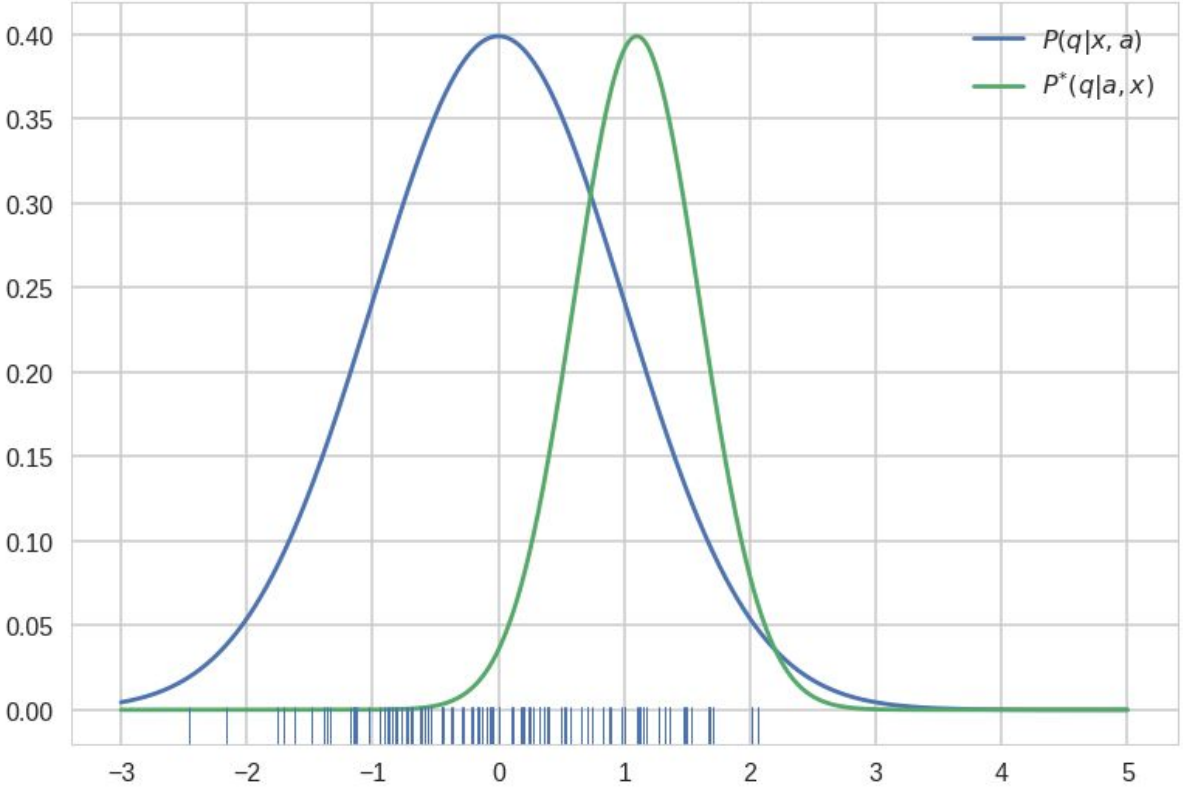
Click-through rate of M*

$$\boxed{Y^*} = \int_\omega y\,\boxed{\mathrm{P}^*}(\omega)$$

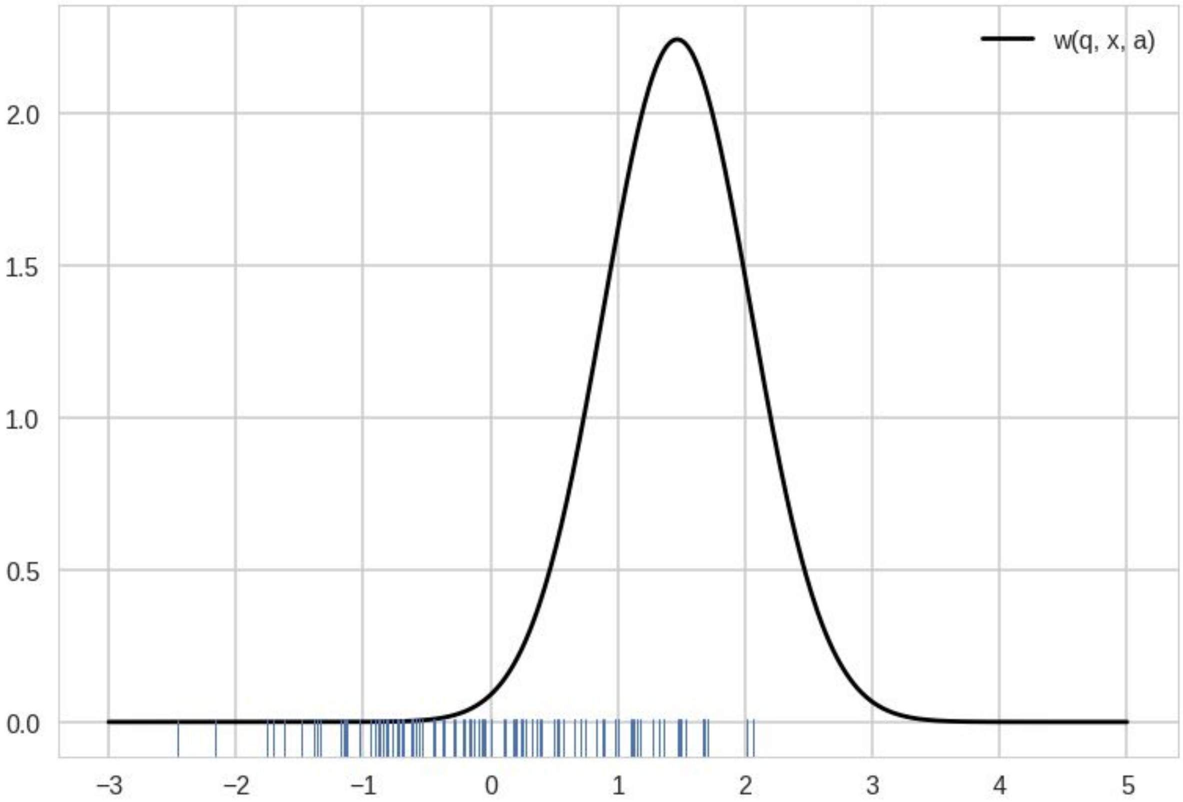
How to estimate Y* from data

$$Y^* = \int y P^*(\omega) d\omega$$

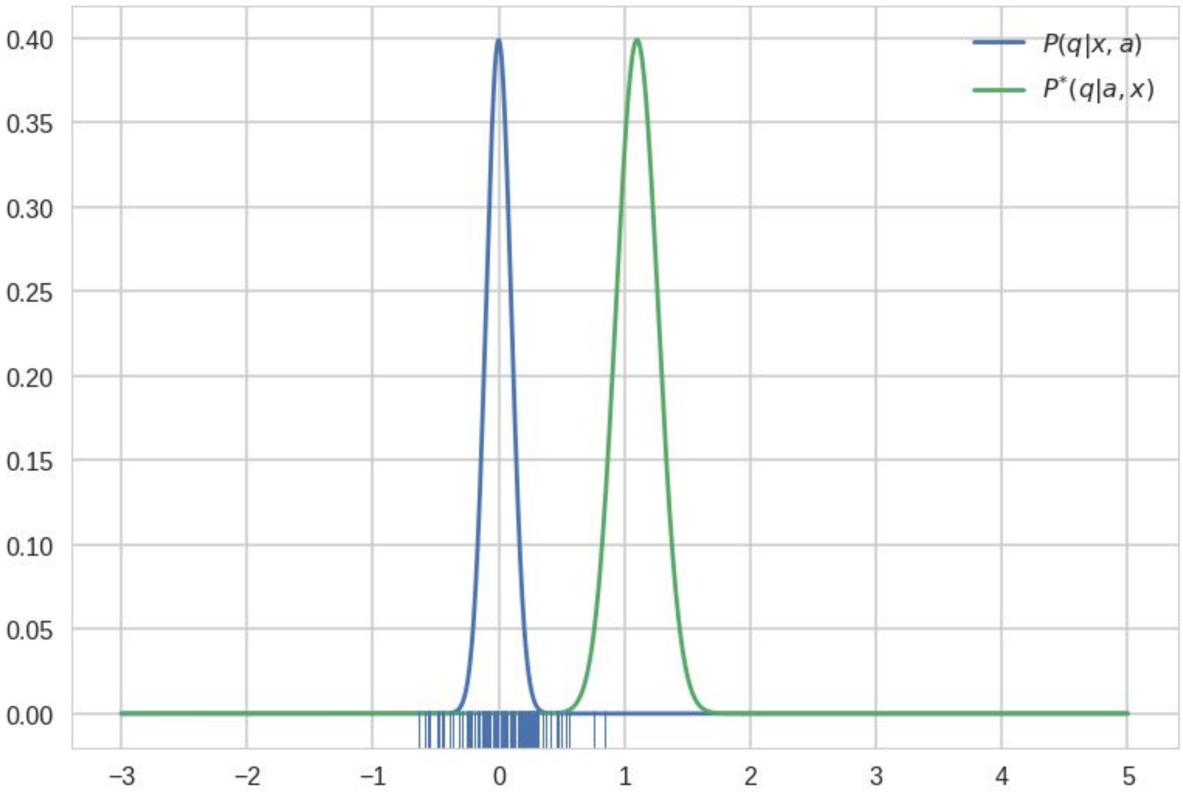
Importance sampling in pictures



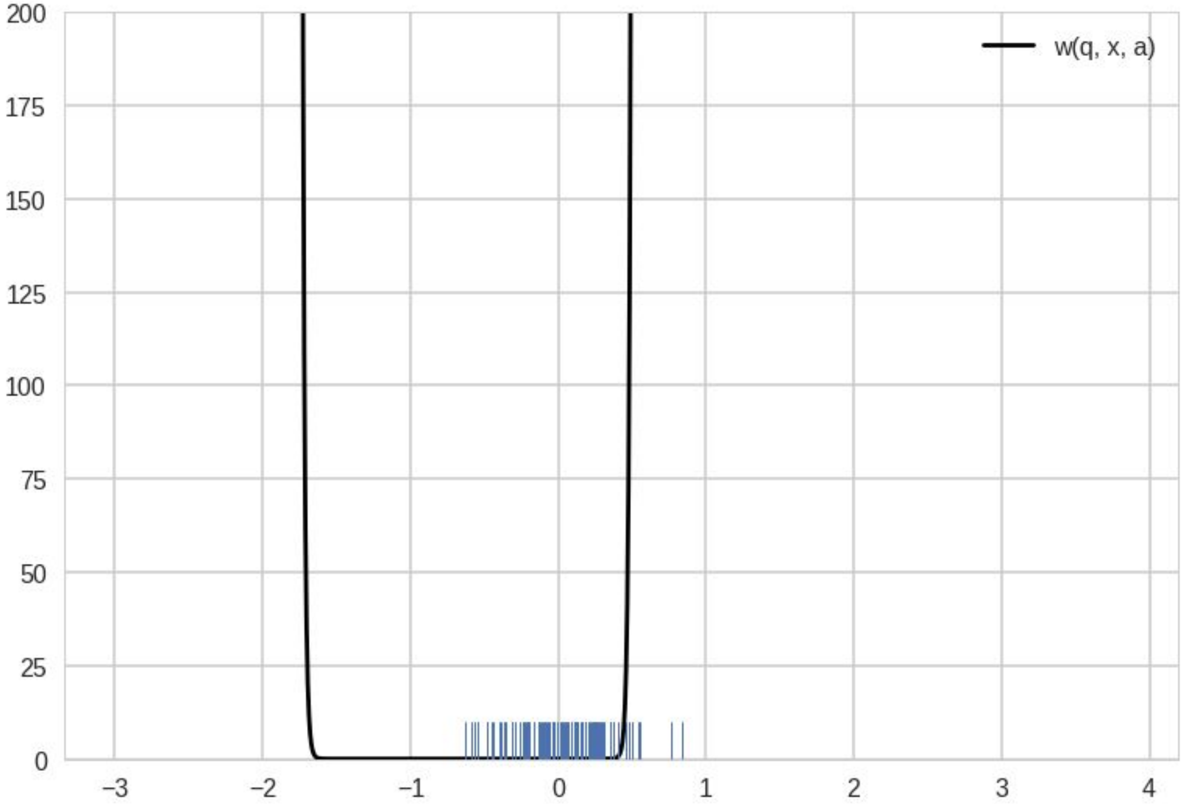
Importance sampling in pictures



Importance sampling in pictures



Importance sampling in pictures



Importance sampling

Requires models M and M* to produce non-deterministic scores

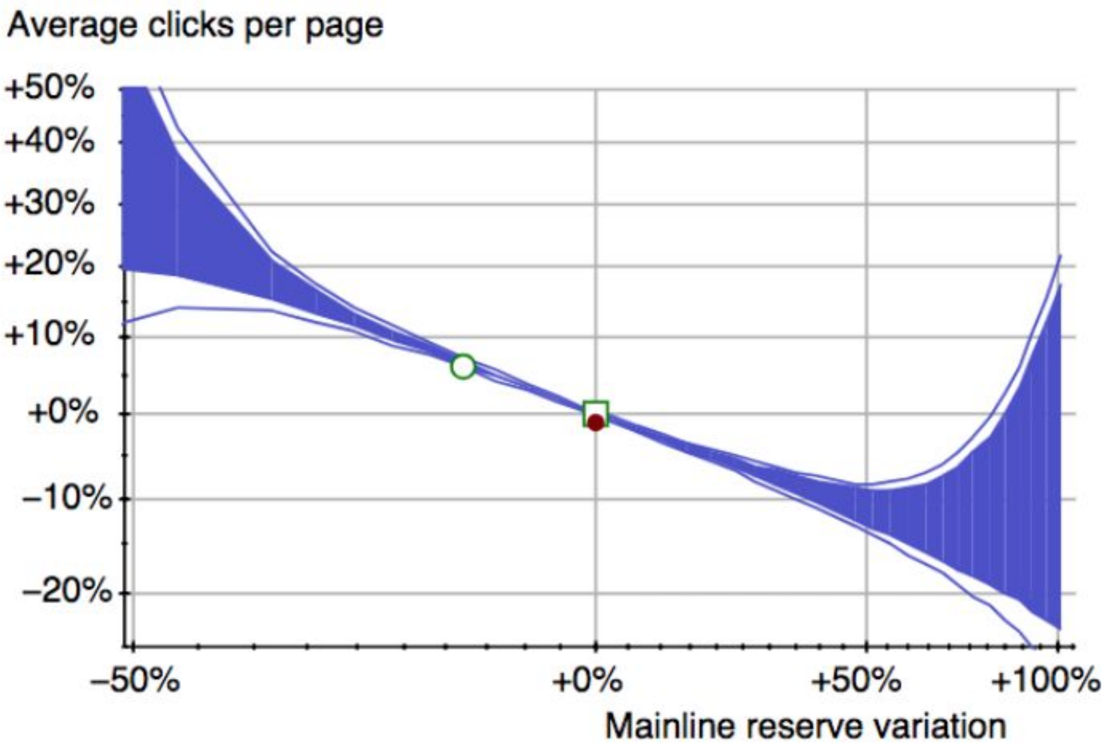
Requires “overlap” between M and M*

The more dissimilar M and M* are, the higher the variance of importance weights

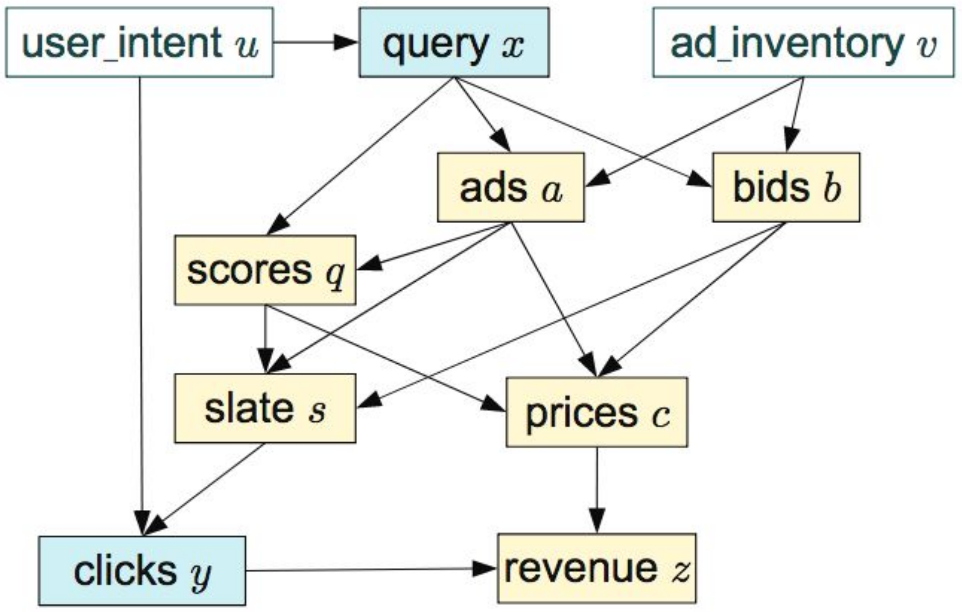
It has excessively high variance in practice

Capped Importance Weighting: reduces variance, but introduces bias

Clipped Importance Weighting results



Leveraging Predictors for Variance Reduction



Leveraging Predictors for Variance Reduction

$$Y^* = \int_\omega yP^*(\omega)$$

Offline A/B testing for Recommender Systems

Alexandre Gilotte, Clément Calauzènes, Thomas Nedelec, Alexandre Abraham, Simon Dollé
Criteo Research
f.name@criteo.com

ABSTRACT

Online A/B testing evaluates the impact of a new technology by running it in a real production environment and testing its performance on a subset of the users of the platform. It is a well-known practice to run a preliminary offline evaluation on historical data to iterate faster on new ideas, and to detect poor policies in order to avoid losing money or breaking the system. For such offline evaluations, we are interested in methods that can compute offline an estimate of the potential uplift of performance generated by a new technology. Offline performance can be measured using estimators known as *counterfactual* or *off-policy* estimators. Traditional counterfactual estimators, such as *capped importance sampling* or *normalised importance sampling*, exhibit unsatisfying bias-variance compromises when experimenting on personalized product recommendation systems. To overcome this issue, we model the bias incurred by these estimators rather than bound it in the worst case, which leads us to propose a new counterfactual estimator. We provide a benchmark of the different estimators showing their correlation with business metrics observed by running online A/B tests on a large-scale commercial recommender system.

Online A/B tests became ubiquitous in tech companies in order to make informed decisions on the rollout of a new technology such as a recommender system. Each new software implementation is tested by comparing its performance with the previous production version through randomised experiments. In practice, to compare two technologies, a pool of units (e.g. users, displays or servers) of the platform is split in two populations and each of them is exposed to one of the tested technologies. The careful choice of the unit reflects independence assumptions under which the test is run. At the end of the experiments, business metrics such as the generated revenue, the number of clicks or the time spent on the platform are compared to make a decision on the future of the new technology.

However, online A/B tests take time and cost money. Indeed, to gather a sufficient amount of data to reach statistical sufficiency and be able to study periodic behaviours (the signal can be different from one day to the other), an A/B test is usually implemented over several weeks. On top of this, prototypes need to be brought to production standard to be tested. These reasons prevent companies from iterating quickly on new ideas.

To solve these pitfalls, people historically relied on offline experiments based on some rank-based metrics, such as NDCG [13],

Different importance sampling-like estimators

Table 2: Summary table of the different estimators. First column sum up the formulae of the estimators, second one the approximation $\tilde{\mathcal{B}}$ of the bias term $\mathcal{B}$ in the general case and the case of zero capping

	$\hat{\mathcal{R}}(\pi_t, c)$	approx $\tilde{\mathcal{B}}^{\text{CIS}}(\pi_t, c, x)$
CIS	$\frac{1}{n}\sum_{S_n} r\overline{w}(a,x)$	0
NCIS	$\frac{\sum_{S_n} r\overline{w}(a,x)}{\sum_{S_n} \overline{w}(a,x)}$	$\mathbb{E}_{\pi_t}\left[\frac{R\overline{W}}{W}\right]\frac{1-\mathbb{E}_{\pi_t}\left[\frac{\overline{W}}{W}\right]}{\mathbb{E}_{\pi_t}\left[\frac{\overline{W}}{W}\right]}$
PieceNCIS	$\sum_{g\in\mathcal{G}} \alpha_g \hat{\mathcal{R}}\|_g^{\text{NCIS}}(\pi_t, c)$	$\mathbb{E}_{\pi_t}\left[\frac{R\overline{W}}{W}\middle\|X\in g\right]\frac{1-\mathbb{E}_{\pi_t}\left[\frac{\overline{W}}{W}\middle\|X\in g\right]}{\mathbb{E}_{\pi_t}\left[\frac{\overline{W}}{W}\middle\|X\in g\right]}$
PointNCIS	$\frac{1}{n}\sum_{S_n} \hat{IP}_c(x)\overline{w}(a,x)r$	$\mathbb{E}_{\pi_t}\left[\frac{R\overline{W}}{W}\middle\|X=x\right]\frac{1-\mathbb{E}_{\pi_t}\left[\frac{\overline{W}}{W}\middle\|X=x\right]}{\mathbb{E}_{\pi_t}\left[\frac{\overline{W}}{W}\middle\|X=x\right]}$

Different importance sampling-like estimators

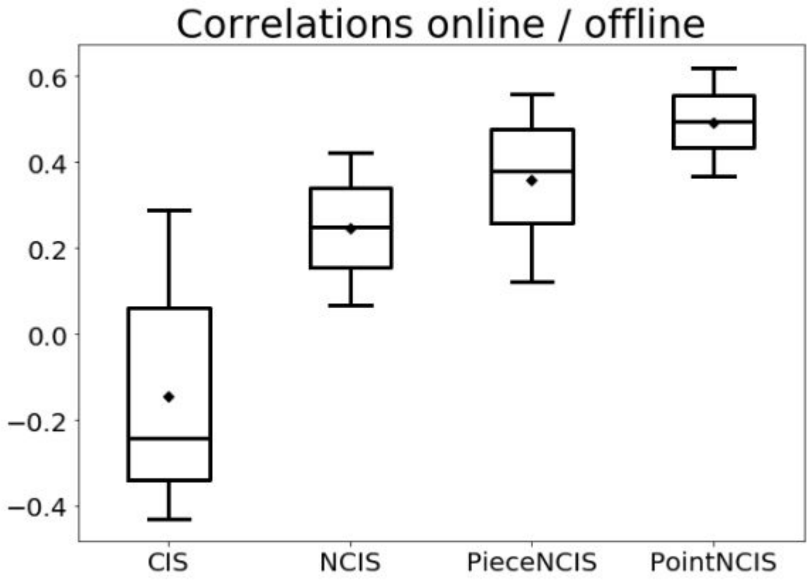


Figure 3: Correlation between online and offline uplifts. Confidence bounds are obtained using bootstraps. Whiskers are 10% and 90% quantiles and the boxes represent quartiles.

Counterfactual Risk Minimization: Learning from Logged Bandit Feedback

Adith Swaminathan ADITH@CS.CORNELL.EDU
Cornell University, Ithaca, NY 14853 USA

Thorsten Joachims TJ@CS.CORNELL.EDU
Cornell University, Ithaca, NY 14853 USA

Abstract

We develop a learning principle and an efficient algorithm for batch learning from logged bandit feedback. This learning setting is ubiquitous in online systems (e.g., ad placement, web search, recommendation), where an algorithm makes a prediction (e.g., ad ranking) for a given input (e.g., query) and observes bandit feedback (e.g., user clicks on presented ads). We first address the counterfactual nature of the learning problem through propensity scoring. Next, we prove generalization error bounds that account for the variance of the propensity-weighted empirical risk estimator. These constructive bounds give rise

number of ranked articles the user read) (Li et al., 2010). The feedback, however, provides only partial information – "bandit feedback"– limited to the particular prediction shown by the system. The feedback for all the other predictions the system could have made is typically not known. This makes learning from log data fundamentally different from supervised learning, where "correct" predictions (e.g., the best ranking of news articles for that user) together with a loss function provide full-information feedback.

We study the problem of batch learning from logged bandit feedback. Unlike online learning with bandit feedback, batch learning does not require interactive experimental control over the system. Furthermore, it enables the reuse of existing data and offline cross-validation techniques for

Bandit setting

context x

recommendation y

policy $h_0(y \mid x)$ selects the recommendation

$\delta(x,y)$: feedback from the user

counterfactual risk:

$$\hat{R}(h) = \frac{1}{n} \sum_{i=1}^{n} \delta_i \frac{h(y_i \mid x_i)}{p_i}.$$

Take home messages

blindspot

basic hygiene

not rocket science

not just healthcare

not either or

Practical