THE UNIVERSITY
of EDINBURGH

Methods for Causal Inference
Lecture 10: Pearl's adjustment formula

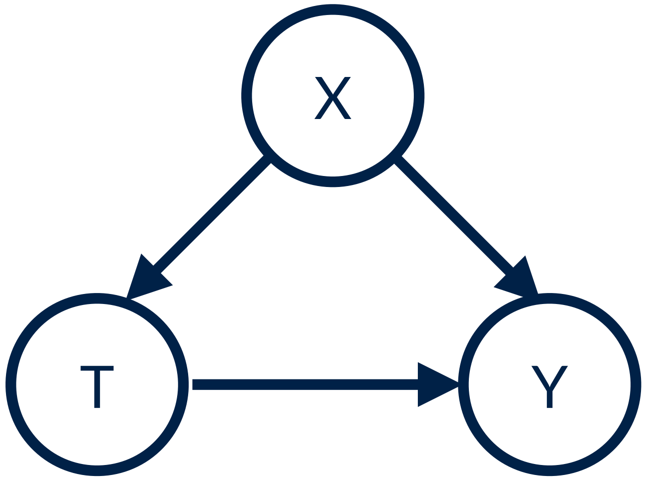
Ava Khamseh

School of Informatics

2025-2026

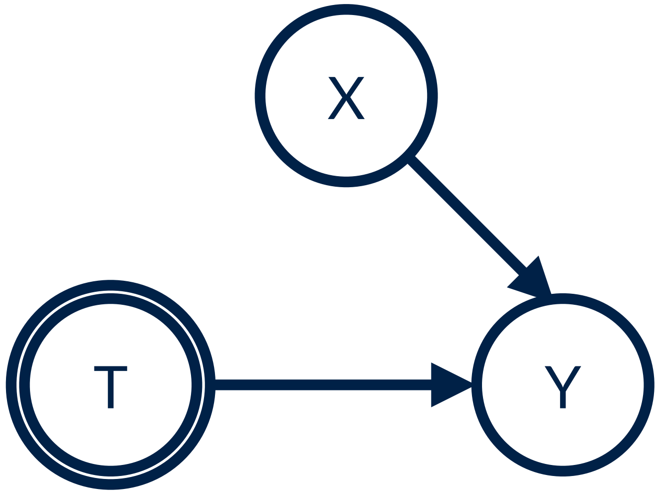
Observation (conditioning) vs intervention

Distinguish between: a variable T takes a value t naturally and cases where we **fix** T=t by denoting the latter do(T=t)



$$p(Y = y|T = t)$$

Probability that Y=y **conditional** on finding T=t
i.e., population distribution of Y among individuals whose T value is t (subset)



$$p(Y = y|do(T = t))$$

Probability that Y=y when we **intervene** to make T=t
i.e., population distribution of Y if **everyone in the population** had their T value fixed at t.

Graph surgery

Pearl, Causal Inference in Statistics (2016)

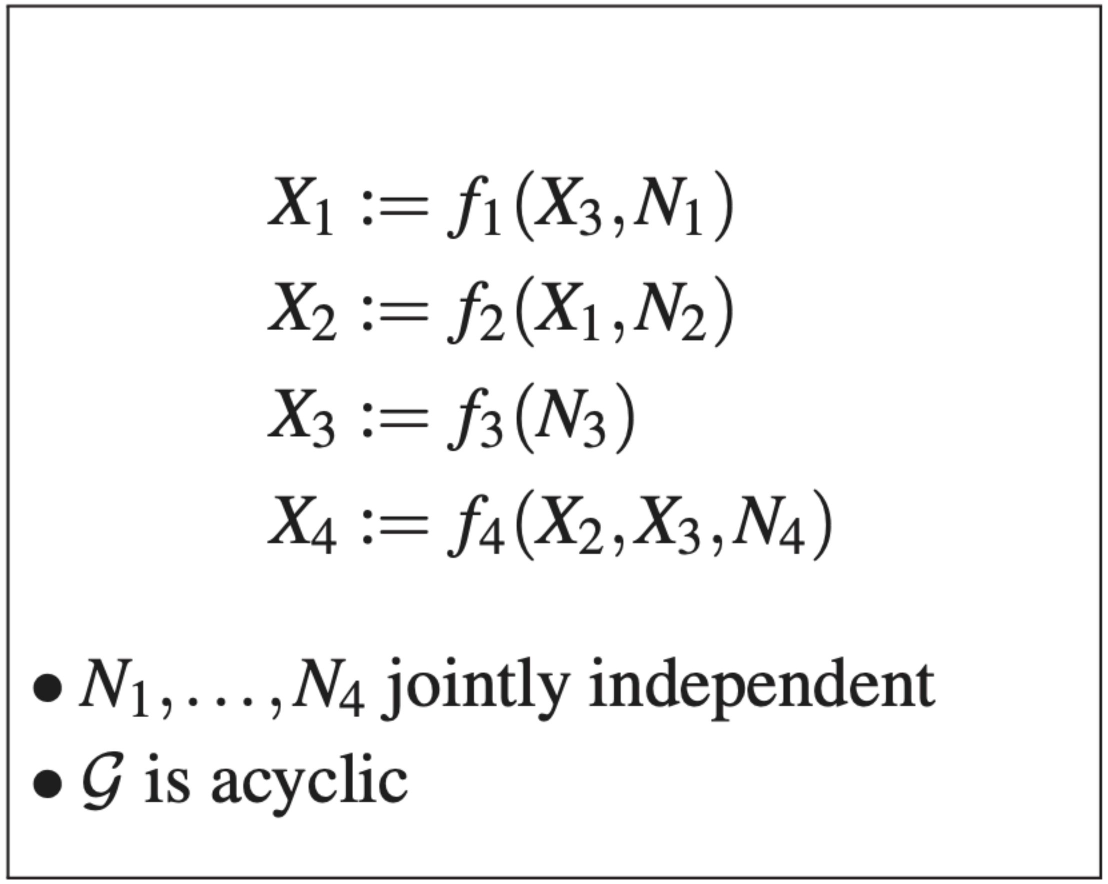
Structural Causal Models (SCM)

An SCM consists of d structural assignments

$$X_j := f_j(PA_j, N_j) \quad , \quad j = 1, \cdots, d$$

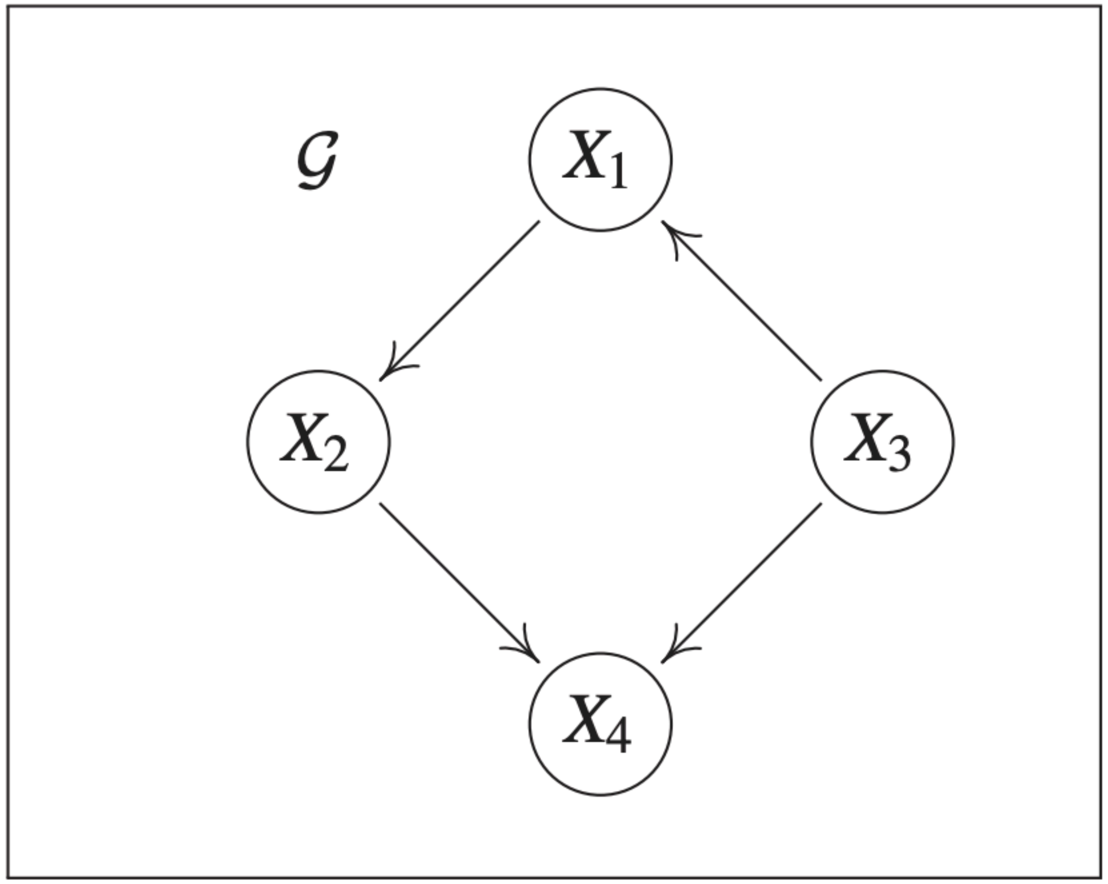
Parents of X_j , i.e., direct causes of X_j

Jointly independent noise variables



$$X_1 := f_1(X_3, N_1)$$
$$X_2 := f_2(X_1, N_2)$$
$$X_3 := f_3(N_3)$$
$$X_4 := f_4(X_2, X_3, N_4)$$

- $N_1, \ldots, N_4$ jointly independent
- $\mathcal{G}$ is acyclic



Jonas Peters et al, Elements of Causal Inference (2017)

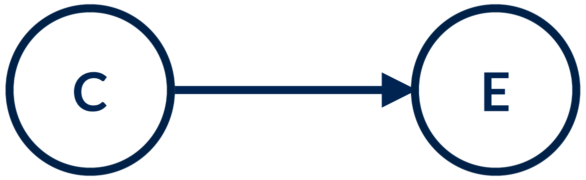
Intervention vs observation: Example

- Consider the following causal model with structure equations:

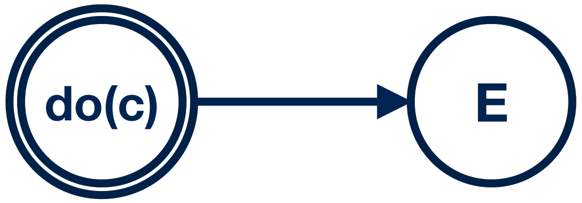
Random Variables

$$C := N_C$$

$$E := 4 \cdot C + N_E$$



where, $N_C, N_E \sim \mathcal{N}(0,1)$, are independent and iid. **We expect**:



- Apply do(C):
  - The new distribution $p(E|do(C)) \neq p(E)$
  - Since there are no other confounders: $p(E|do(C)) = p(E|C)$

Jonas Peters et al, Elements of Causal Inference (2017)

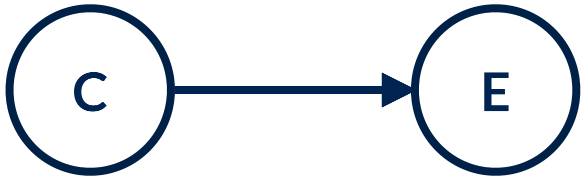
Intervention vs observation: Example

- Consider the following causal model with structure equations:

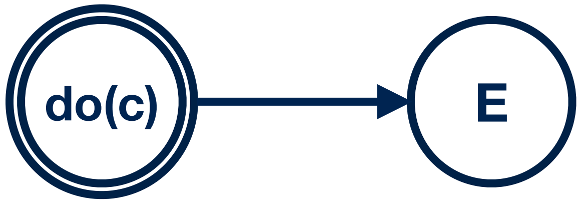
Random Variables

$$C := N_C$$

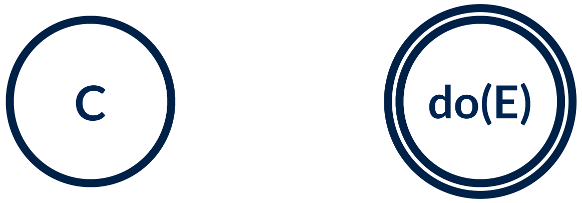
$$E := 4 \cdot C + N_E$$



where, $N_C, N_E \sim \mathcal{N}(0,1)$, are independent and iid. **We expect**:



- Apply do(C):
  - The new distribution $p(E|do(C)) \neq p(E)$
  - Since there are no other confounders: $p(E|do(C)) = p(E|C)$

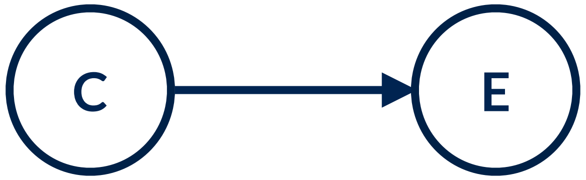


- Apply do(E):
  - The new distribution $p(C|do(E)) = p(C)$
  - Since there are no other confounders: $p(C|do(E)) \neq p(C|E)$

Jonas Peters et al, Elements of Causal Inference (2017)

Intervention vs observation: Analytical computation

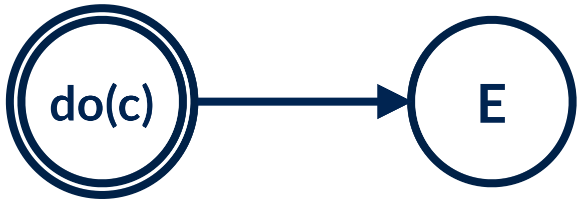
$$C := N_C$$
$$E := 4 \cdot C + N_E$$
$$N_C, N_E \sim \mathcal{N}(0,1), N_C \perp\!\!\!\perp N_E$$



Using, $\mathrm{Var}[aX] = a^2\mathrm{Var}[X]$, $4C \sim \mathcal{N}(0,16)$.

Using, $4C \perp\!\!\!\perp N_E$, and the sum of two normally distributed random variables is another normally distributed random variable (by **convolution**):

$$E \sim \mathcal{N}\left(\mu_{4C} + \mu_{N_E}, \sigma^2_{4C} + \sigma^2_{N_E}\right)$$
$$\Rightarrow E \sim \mathcal{N}\left(0, 17\right)$$



A fixed number

$$p(E) = \mathcal{N}\left(0,17\right) \neq \mathcal{N}\left(8,1\right) = p(E|do(C=2)) = p(E|C=2)$$
$$\neq \mathcal{N}\left(12,1\right) = p(E|do(C=3)) = p(E|C=3)$$

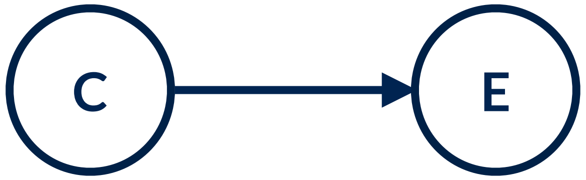
Jonas Peters et al, Elements of Causal Inference (2017)

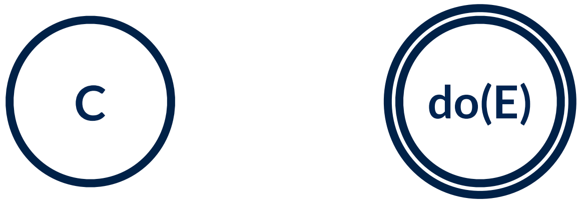
Intervention vs observation: Analytical computation

$$C := N_C$$

$$E := 4 \cdot C + N_E$$

$$N_C, N_E \sim \mathcal{N}(0,1), N_C \perp\!\!\!\perp N_E$$





$$p(C|do(E=2)) = \mathcal{N}\,(0,1) = p(C|do(E = \text{Any } r > 0)) = p(C)$$

$\neq p(C|E=2)$ in the original distribution above

Proof: Use product rule: $p(C|E) = \dfrac{p(C,E)}{p(E)}$

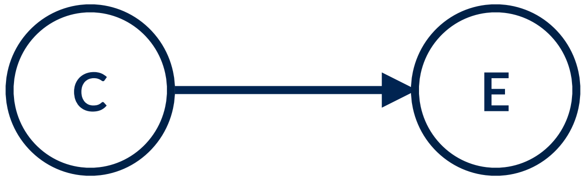
For a bivariate normal distribution (2 joint normal distributions), the marginal:

$$p(C|E) = \mathcal{N}(\tilde{\mu}, \tilde{\sigma}^2) \;\; \text{s.t.} \;\; \tilde{\mu} = \mu_C + \rho\frac{\sigma_C}{\sigma_E}(E - \mu_E), \;\tilde{\sigma}^2 = \sigma_C^2\,(1-\rho^2)$$

Jonas Peters et al, Elements of Causal Inference (2017)

Intervention vs observation: Analytical computation

$$C := N_C$$
$$E := 4 \cdot C + N_E$$
$$N_C, N_E \sim \mathcal{N}(0,1), N_C \perp\!\!\!\perp N_E$$



Proof (Cont.): Use $\mathrm{Cov}(aX, bY + cZ) = ab\ \mathrm{Cov}(X,Y) + ac\ \mathrm{Cov}(X,Z)$

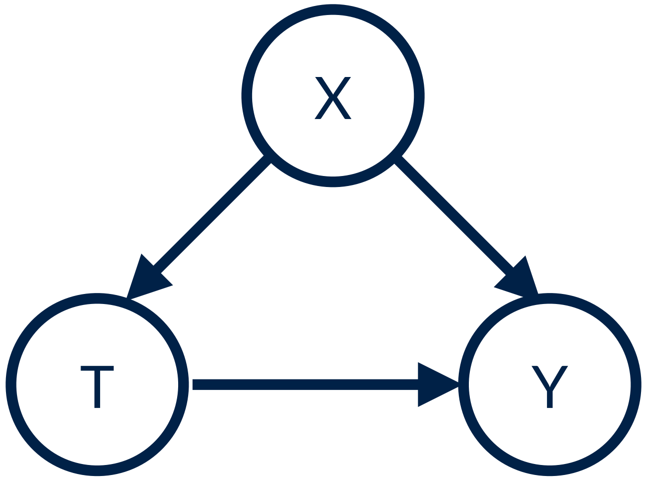
$$\Rightarrow \rho = \frac{\mathrm{Cov}(C,E)}{\sigma_C \sigma_E} = \frac{4\mathrm{Cov}(N_C, N_C) + \mathrm{Cov}(N_C, N_E)}{\sigma_C \sigma_E} = \frac{4}{\sqrt{17}}$$

$$\Rightarrow p(C|E=2) = \mathcal{N}\left(\frac{8}{17}, \sigma^2 = \frac{1}{17}\right) \Rightarrow \quad p(C|do(E)) \neq p(C|E)$$

Jonas Peters et al, Elements of Causal Inference (2017)

The adjustment formula

T: Drug usage
X: Sex
Y: Recovery



To know how effective the drugs is in the population, compare the **hypothetical interventions** by which

(i) the drug is administered uniformly to the entire population do(T=1) **vs**

(ii) complement, i.e., everyone is prevented from taking the drug do(T=0)

Aim: Estimate the difference (**Average Causal Effect ACE, aka ATE**)

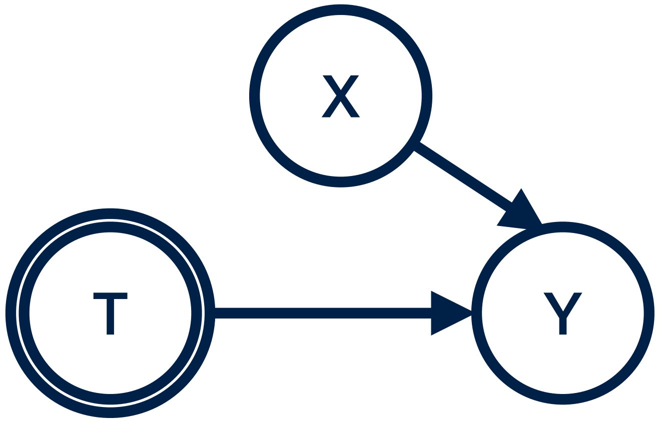
$$\boxed{p(Y = 1|do(T = 1)) - p(Y = 1|do(T = 0))}$$

Pearl, Causal Inference in Statistics (2016)

The adjustment formula

Using a **causal theory**, we aim to write $p(Y = y|do(T = t))$ in terms of quantities we can compute from the data, i.e., conditional probabilities.

The causal effect $p(Y = y|do(T = t))$ is equal to conditional probability in the manipulated graph $p_m(Y = y|T = t)$



Key observation: p_m shares 2 properties with p:

(i) $p_m(X = x) = p(X = x)$ is **invariant** under the intervention, X is not affected by removing the arrow from X to T, i.e. the proportion of males and females remain the same before and after the intervention

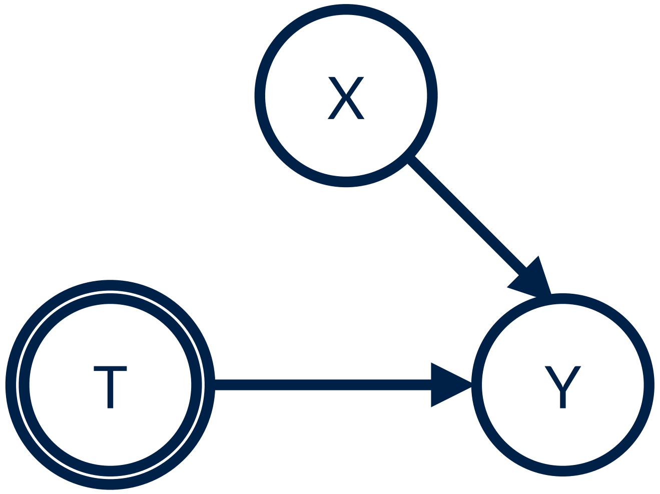
(ii) $p_m(Y = y|X = x, T = t) = p(Y = y|X = x, T = t)$ is **invariant**

Pearl, Causal Inference in Statistics (2016)

The adjustment formula

Moreover, T and X are d-separated in the modified model:

$$p_m(X = x|T = t) = p_m(X = x) = p(X = x) \quad *$$

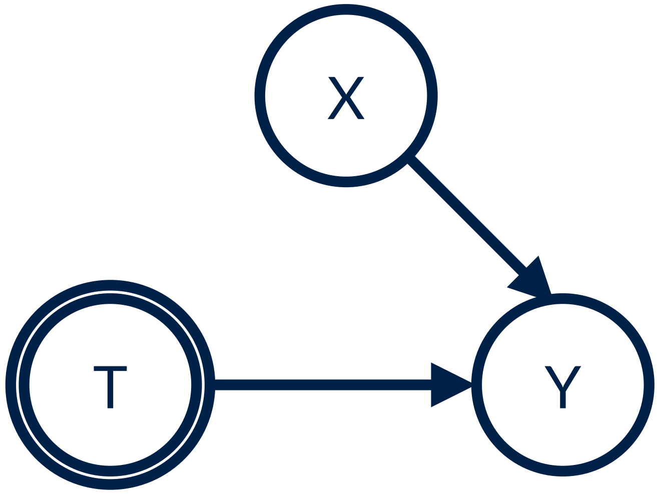


Pearl, Causal Inference in Statistics (2016)

The adjustment formula

Moreover, T and X are d-separated in the modified model:

$$p_m(X = x|T = t) = p_m(X = x) = p(X = x) \quad *$$



Putting these together:

$$p(Y = y|do(T = t)) = p_m(Y = y|T = t) \quad \text{by definition}$$

$$\sum_x p_m(Y = y|T = t, X = x)p_m(X = x|T = t) \quad \text{law of total prob}$$

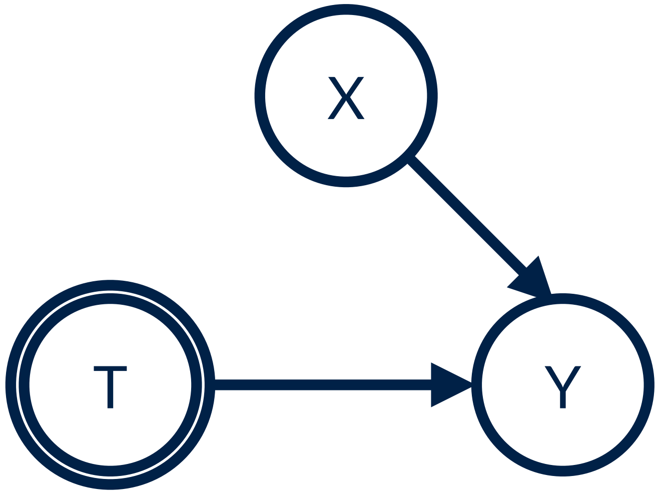
$$\sum_x p_m(Y = y|T = t, X = x)p_m(X = x) \quad *$$

Pearl, Causal Inference in Statistics (2016)

The adjustment formula

Moreover, T and X are d-separated in the modified model:

$$p_m(X = x|T = t) = p_m(X = x) = p(X = x) \quad *$$



Putting these together:

$$p(Y = y|do(T = t)) = p_m(Y = y|T = t) \quad \text{by definition}$$

$$\sum_x p_m(Y = y|T = t, X = x)p_m(X = x|T = t) \quad \text{law of total prob}$$

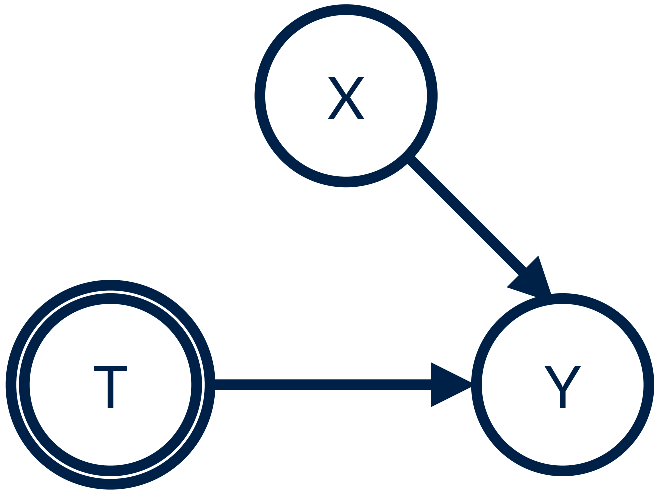
$$\sum_x p_m(Y = y|T = t, X = x)p_m(X = x) \quad *$$

Using the two invariance relations, we have the **adjustment formula**:

$$\boxed{p(Y = y|do(T = t)) = \sum_x p(Y = y|T = t, X = x)p(X = x)}$$

Pearl, Causal Inference in Statistics (2016)

The adjustment formula



Moreover, T and X are d-separated in the modified model:

$$p_m(X = x|T = t) = p_m(X = x) = p(X = x) \quad *$$

Putting these together:

$$p(Y = y|do(T = t)) = p_m(Y = y|T = t)$$ by definition

$$\sum_x p_m(Y = y|T = t, X = x)p_m(X = x|T = t)$$ law of total prob

$$\sum_x p_m(Y = y|T = t, X = x)p_m(X = x) \quad *$$

Use P_m as an intermediate tool

Using the two invariance relations, we have the **adjustment formula**:

$$\boxed{p(Y = y|do(T = t)) = \sum_x p(Y = y|T = t, X = x)p(X = x)}$$

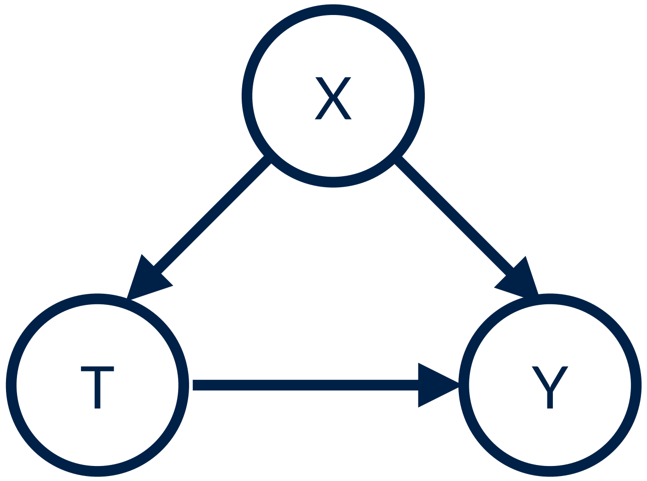
Pearl, Causal Inference in Statistics (2016)

The adjustment formula

$$\boxed{p(Y = y|do(T = t)) = \sum_x p(Y = y|T = t, X = x)p(X = x)}$$

Adjusting for X (controlling for X) … **seen before**?

Example: T=1 taking the drug, X=1 male, Y=1 recovery



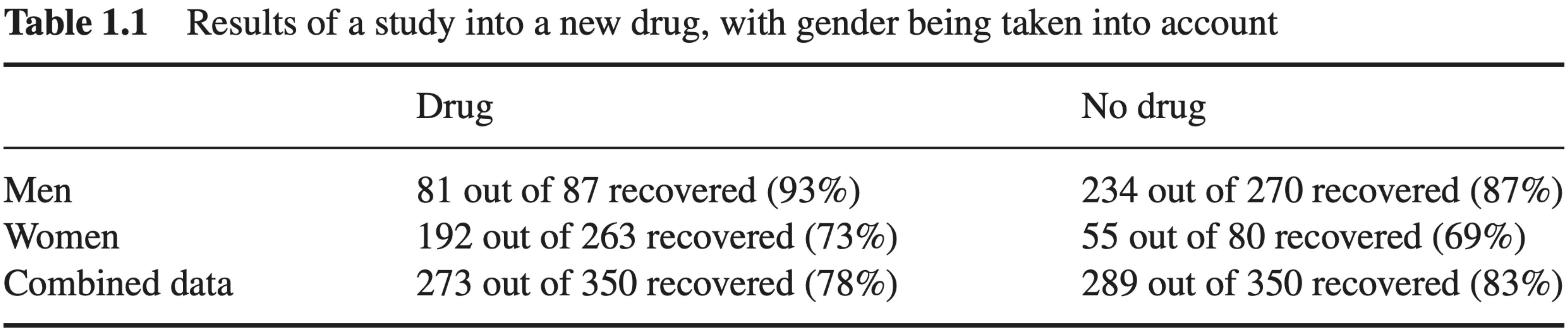


Table 1.1 Results of a study into a new drug, with gender being taken into account

	Drug	No drug
Men	81 out of 87 recovered (93%)	234 out of 270 recovered (87%)
Women	192 out of 263 recovered (73%)	55 out of 80 recovered (69%)
Combined data	273 out of 350 recovered (78%)	289 out of 350 recovered (83%)

Pearl, Causal Inference in Statistics (2016)

The adjustment formula

$$\boxed{p(Y=y|do(T=t)) = \sum_x p(Y=y|T=t, X=x)p(X=x)}$$

T=1 taking drug
X=1 male
Y=1 recovery

$$p(Y=y|do(T=1)) = p(Y=1|T=1, X=1)p(X=1) + p(Y=1|T=1, X=0)p(X=0)$$

$$p(Y=1|do(T=1)) = \frac{0.93(87+270)}{700} + \frac{0.73(263+80)}{700} = 0.832$$

$$p(Y=1|do(T=0)) = \frac{0.87(87+270)}{700} + \frac{0.69(263+80)}{700} = 0.7818$$

$$ACE: p(Y=1|do(T=1)) - p(Y=1|do(T=0)) = 0.832 - 0.7818 = 0.0505$$ ✓

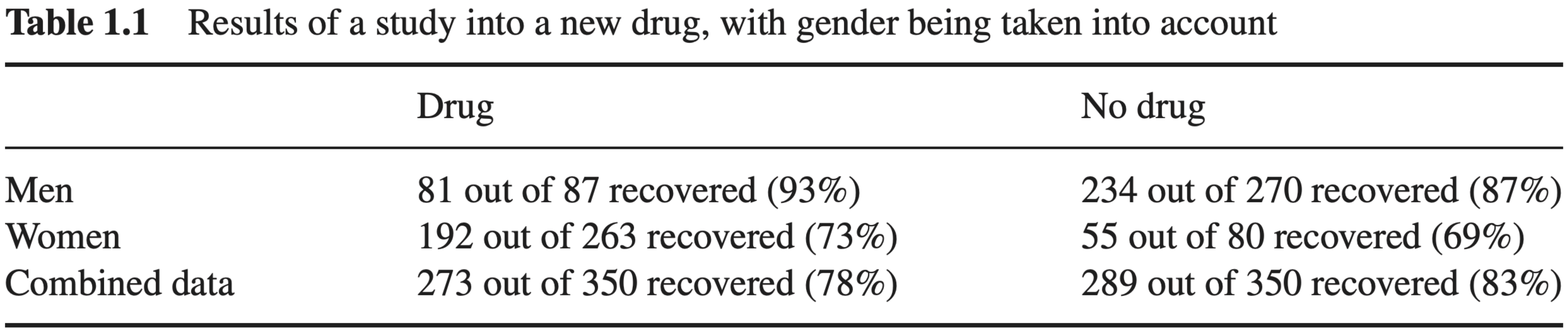


Table 1.1 Results of a study into a new drug, with gender being taken into account

	Drug	No drug
Men	81 out of 87 recovered (93%)	234 out of 270 recovered (87%)
Women	192 out of 263 recovered (73%)	55 out of 80 recovered (69%)
Combined data	273 out of 350 recovered (78%)	289 out of 350 recovered (83%)

Pearl, Causal Inference in Statistics (2016)

The adjustment formula

$$\boxed{p(Y=y|do(T=t)) = \sum_x p(Y=y|T=t,X=x)p(X=x)}$$

T=1 taking drug
X=1 male
Y=1 recovery

$$p(Y=y|do(T=1)) = p(Y=1|T=1,X=1)p(X=1) + p(Y=1|T=1,X=0)p(X=0)$$

$$p(Y=1|do(T=1)) = \frac{0.93(87+270)}{700} + \frac{0.73(263+80)}{700} = 0.832$$

Stratification!

$$p(Y=1|do(T=0)) = \frac{0.87(87+270)}{700} + \frac{0.69(263+80)}{700} = 0.7818$$

Note equivalence to Rubin's FW

$$ACE: p(Y=1|do(T=1)) - p(Y=1|do(T=0)) = 0.832 - 0.7818 = 0.0505$$ ✓

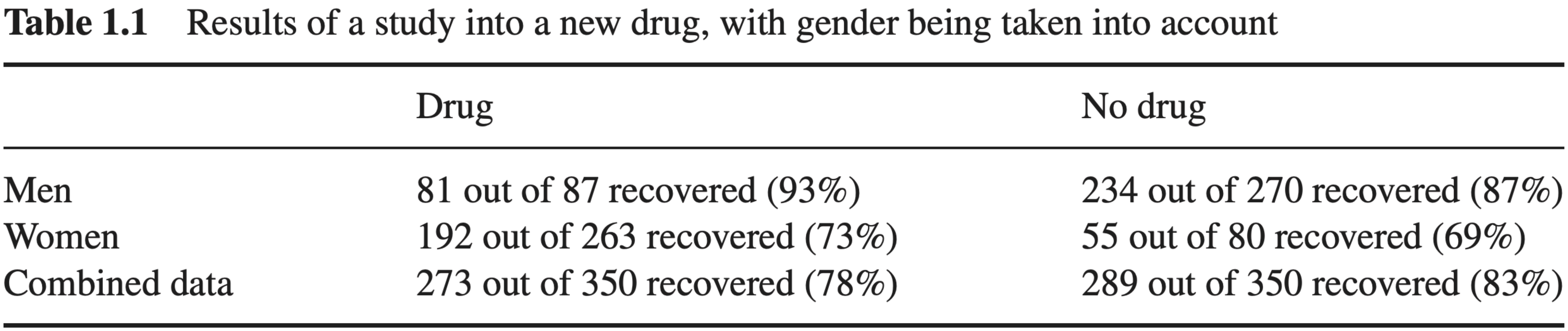


Table 1.1 Results of a study into a new drug, with gender being taken into account

	Drug	No drug
Men	81 out of 87 recovered (93%)	234 out of 270 recovered (87%)
Women	192 out of 263 recovered (73%)	55 out of 80 recovered (69%)
Combined data	273 out of 350 recovered (78%)	289 out of 350 recovered (83%)

Pearl, Causal Inference in Statistics (2016)

Pearl & Rubin

Pearl

$$\mathbb{E}(Y|do(T=1)) = \mathbb{E}(Y|T=1, X=1)p(X=1) + \mathbb{E}(Y|T=1, X=0)p(X=0)$$
$$\mathbb{E}(Y|do(T=0)) = \mathbb{E}(Y|T=0, X=1)p(X=1) + \mathbb{E}(Y|T=0, X=0)p(X=0)$$
$$\mathbb{E}(Y|do(T=1)) - \mathbb{E}(Y|do(T=0))$$

Rubin recall potential outcomes $y_0^{(i)}$ and $y_1^{(i)}$ and ATE:

$$\tau = \hat{\mathbb{E}}[\tau^{(i)}] = \hat{\mathbb{E}}[y_1^{(i)} - y_0^{(i)}] = \frac{1}{N}\sum_{i=0}^{N}\left(y_1^{(i)} - y_0^{(i)}\right)$$

Pearl, Causal Inference in Statistics (2016)

Pearl & Rubin

Pearl

$$\mathbb{E}(Y|do(T=1)) = \mathbb{E}(Y|T=1,X=1)p(X=1) + \mathbb{E}(Y|T=1,X=0)p(X=0)$$
$$\mathbb{E}(Y|do(T=0)) = \mathbb{E}(Y|T=0,X=1)p(X=1) + \mathbb{E}(Y|T=0,X=0)p(X=0)$$
$$\mathbb{E}(Y|do(T=1)) - \mathbb{E}(Y|do(T=0))$$

Rubin recall potential outcomes $y_0^{(i)}$ and $y_1^{(i)}$ and ATE:

$$\tau = \hat{\mathbb{E}}[\tau^{(i)}] = \hat{\mathbb{E}}[y_1^{(i)} - y_0^{(i)}] = \frac{1}{N}\sum_{i=0}^{N}\left(y_1^{(i)} - y_0^{(i)}\right)$$
$$= \frac{1}{N}\left(\sum_{i\in\text{males}}\left(y_1^{(i)} - y_0^{(i)}\right) + \sum_{i\in\text{females}}\left(y_1^{(i)} - y_0^{(i)}\right)\right)$$

Pearl, Causal Inference in Statistics (2016)

Pearl: To adjust or not to adjust

The previous example may give the impression that X-specific analysis, as compared to nonspecific, is the correct way forward. This is not the case. For example, let T=drug, Y=recovery, X= blood pressure **post-treatment,** i.e., important to take into account **how** the data is generated. Here, we know:

(i) the drug affects recovery by lowering the blood pressure

(ii) but it has a toxic effect for those who take it

NB: Data (numbers) in this table are identical to those in Table 1.1.

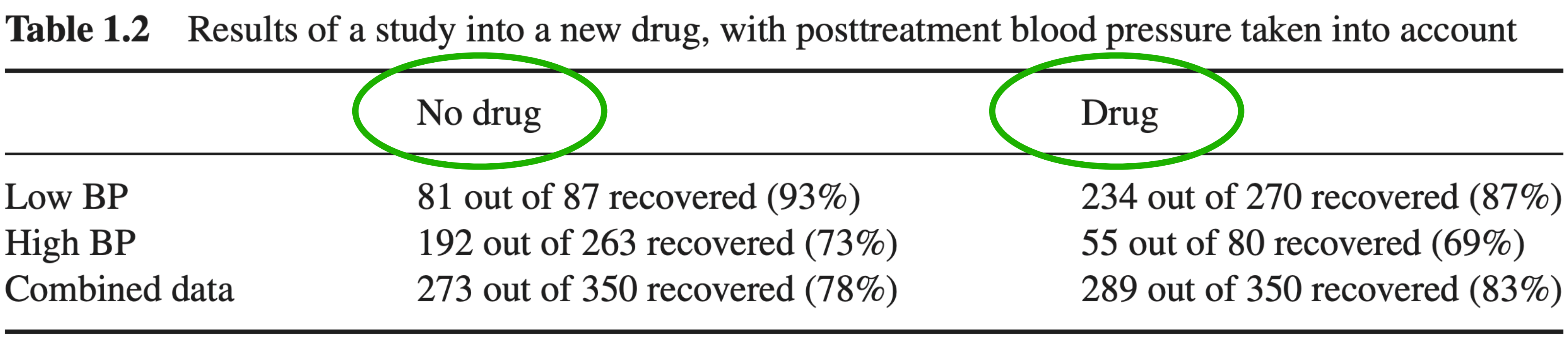
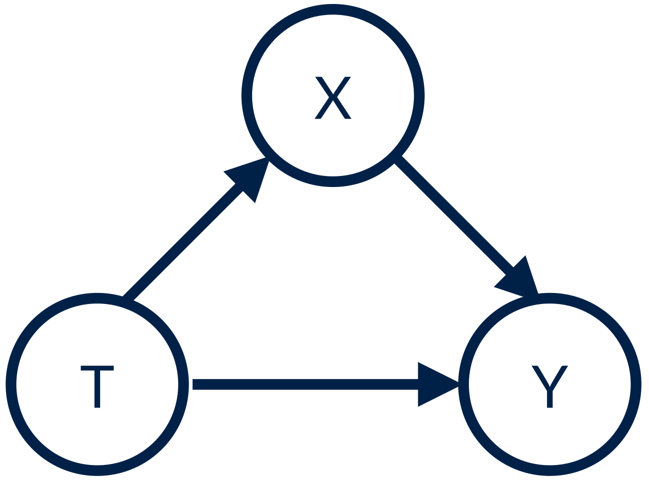


Table 1.2 Results of a study into a new drug, with posttreatment blood pressure taken into account

	No drug	Drug
Low BP	81 out of 87 recovered (93%)	234 out of 270 recovered (87%)
High BP	192 out of 263 recovered (73%)	55 out of 80 recovered (69%)
Combined data	273 out of 350 recovered (78%)	289 out of 350 recovered (83%)



Pearl, Causal Inference in Statistics (2016)

Pearl: To adjust or not to adjust

For general population, the drug might improve recovery rates because of its effect on blood pressure. But in low BP/high BP **post-treatment** subpopulations, we only observe the toxic effect of the drug.

Aim, as before, to gauge the overall causal effect of the drug on recovery.
Unlike before, it does **not** make sense to separate results by blood pressure as treatment affect recovery via reducing BP.
Contrast this with the a situation per BP is measure **before** treatment and direction of arrow from T to X is reversed.

Therefore, we **should** recommend treatment in this case because 78% < 83% .

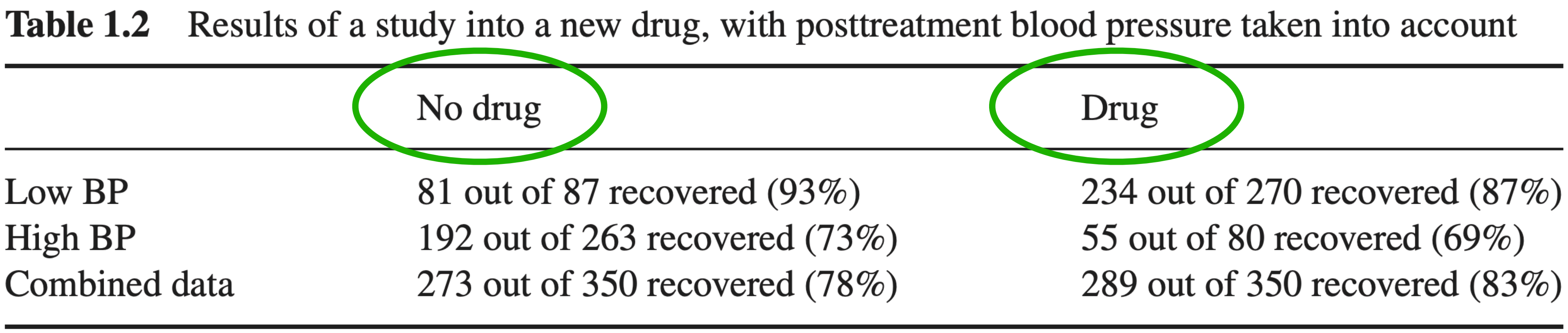
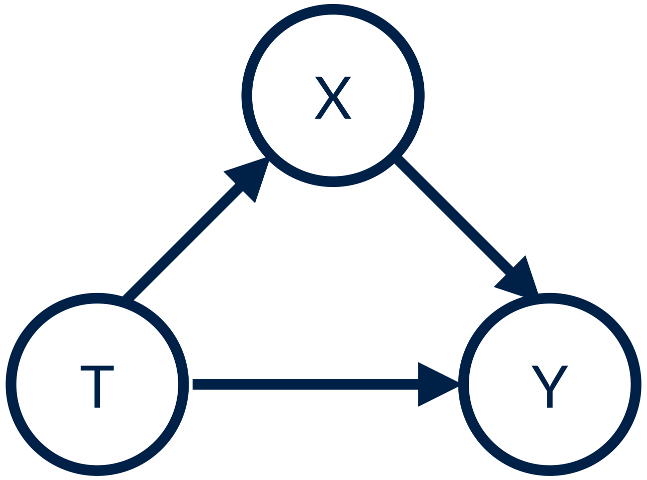


Table 1.2 Results of a study into a new drug, with posttreatment blood pressure taken into account

	No drug	Drug
Low BP	81 out of 87 recovered (93%)	234 out of 270 recovered (87%)
High BP	192 out of 263 recovered (73%)	55 out of 80 recovered (69%)
Combined data	273 out of 350 recovered (78%)	289 out of 350 recovered (83%)



Pearl, Causal Inference in Statistics (2016)

Pearl: To adjust or not to adjust

Pearls algorithmic approach tells us to adjust or not. Starting with: $p(Y = 1|do(T = 1))$, intervene on T. But since no arrow is entering T, there will be no change in the graph: $p(Y = 1|do(T = 1)) = p(Y = 1|T = 1)$

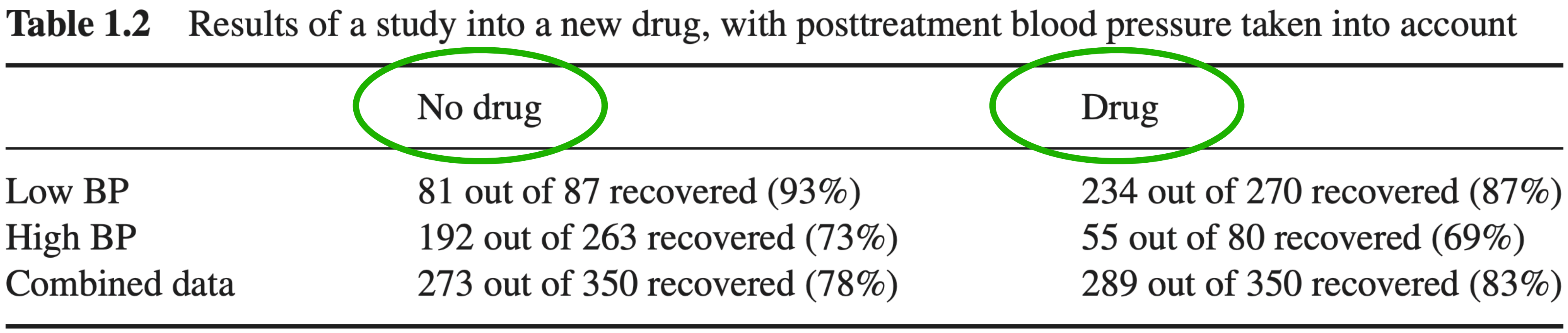
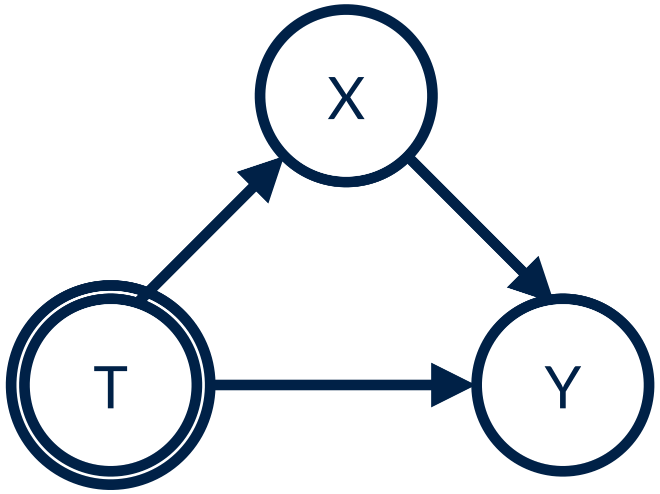


Table 1.2 Results of a study into a new drug, with posttreatment blood pressure taken into account

	No drug	Drug
Low BP	81 out of 87 recovered (93%)	234 out of 270 recovered (87%)
High BP	192 out of 263 recovered (73%)	55 out of 80 recovered (69%)
Combined data	273 out of 350 recovered (78%)	289 out of 350 recovered (83%)



Pearl, Causal Inference in Statistics (2016)

Pearl: To adjust or not to adjust

Pearls algorithmic approach tells us to adjust or not. Starting with: $p(Y=1|do(T=1))$, intervene on T. But since no arrow is entering T, there will be no change in the graph: $p(Y=1|do(T=1)) = p(Y=1|T=1)$

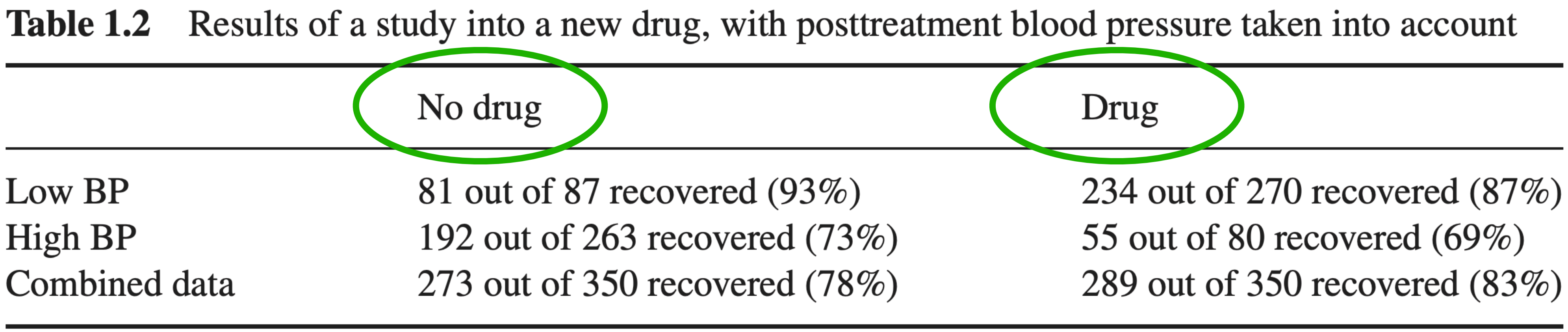
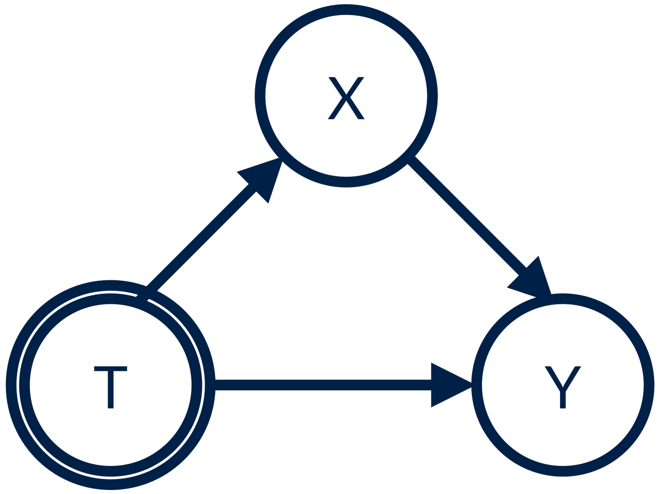


Table 1.2 Results of a study into a new drug, with posttreatment blood pressure taken into account

	No drug	Drug
Low BP	81 out of 87 recovered (93%)	234 out of 270 recovered (87%)
High BP	192 out of 263 recovered (73%)	55 out of 80 recovered (69%)
Combined data	273 out of 350 recovered (78%)	289 out of 350 recovered (83%)



The Causal Effect Rule: Given a graph G in which a set of variables PA are designated as the parents of T, the causal effect of T on Y is given by:

$$\boxed{p(Y=y|do(T=t)) = \sum_x p(Y=y|T=t, PA=X)p(PA=X)}$$

Pearl, Causal Inference in Statistics (2016)

The Backdoor Criterion

Under what conditions does a causal model permit computing the causal effect of one variable on another, from **data** obtained from **passive observations**, with **no intervention**? i.e.,
Under what conditions is the structure of a causal graph sufficient of computing a causal effect from a given data set? **Identifiability**

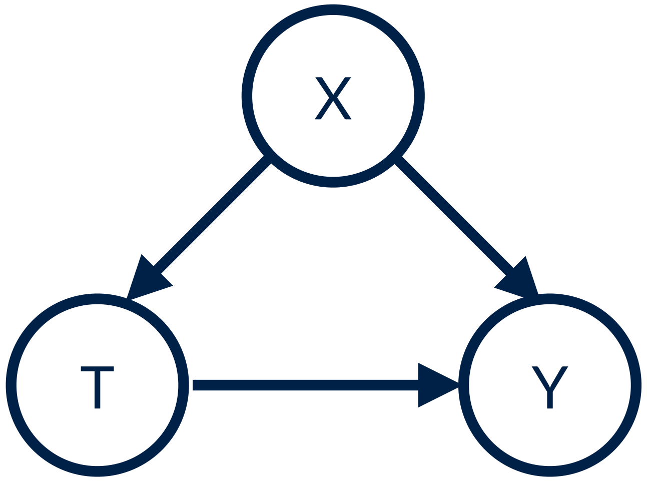
Backdoor Criterion: Given an ordered pair of variables (T,Y) in a DAG G, a set of variables X satisfies the backdoor criterion relative to (T,Y) if:

(i) no node in X is a descendent of T

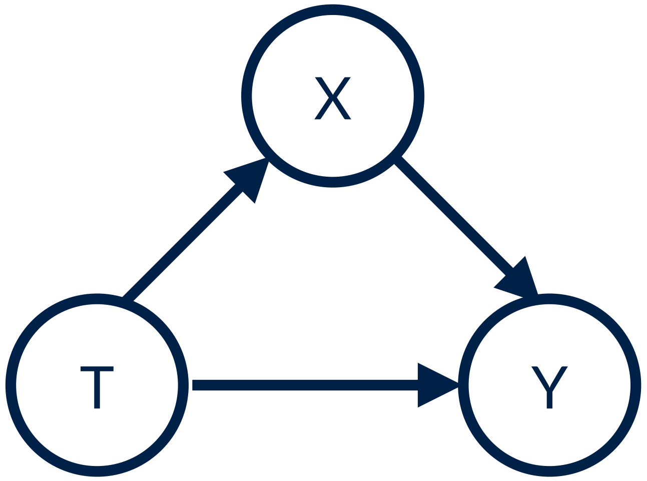
(ii) X block every path between T and Y that contains an arrow into T

If X satisfies the backdoor criterion then the causal effect of T on Y is given by:

$$p(Y=y|do(T=t)) = \sum_x p(Y=y|T=t, X=x)p(X=x)$$



✓



✗

Pearl, Causal Inference in Statistics (2016)

The Backdoor Criterion

Under what conditions does a causal model permit computing the causal effect of one variable on another, from **data** obtained from **passive observations**, with **no intervention**? i.e.,
Under what conditions is the structure of a causal graph sufficient of computing a causal effect from a given data set? **Identifiability**

Backdoor Criterion: Given an ordered pair of variables (T,Y) in a DAG G, a set of variables X satisfies the backdoor criterion relative to (T,Y) if:

(i) no node in X is a descendent of T

(ii) X block every path between T and Y that contains an arrow into T

If X satisfies the backdoor criterion then the causal effect of T on Y is given by:

$$p(Y=y|do(T=t)) = \sum_x p(Y=y|T=t, X=x)p(X=x)$$

In other words, condition on a set of nodes X such that:

(i) We block all spurious paths between T and Y

(ii) We leave all direct paths from T to Y unperturbed

(iii) We create no new spurious paths (do not unblock any new paths)

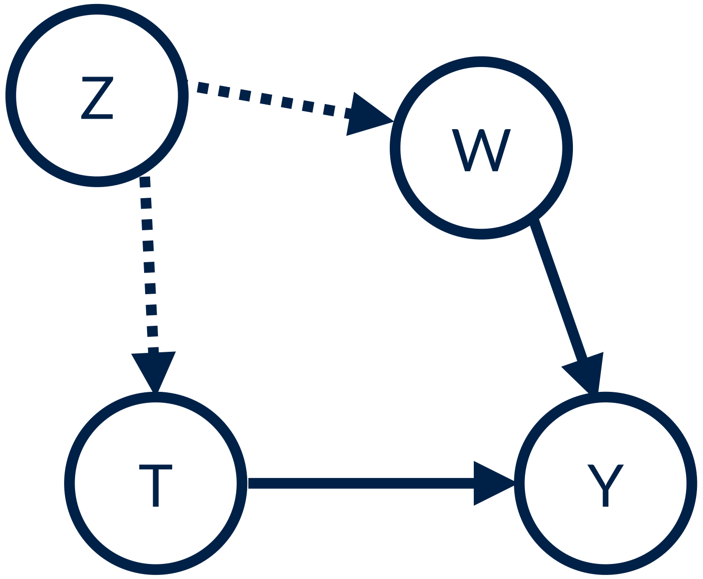
Pearl, Causal Inference in Statistics (2016)

The Backdoor Criterion: Example 1

T = Drug, Y = recovery, W = weight, Z = unmeasured socioeconomic status

Z affects both weight and choice to receive treatment (but Z data was not recorded)

Can we compute the causal effect of T on Y, using W only
(even though Z is not measured)?



Pearl, Causal Inference in Statistics (2016)

The Backdoor Criterion: Example 1

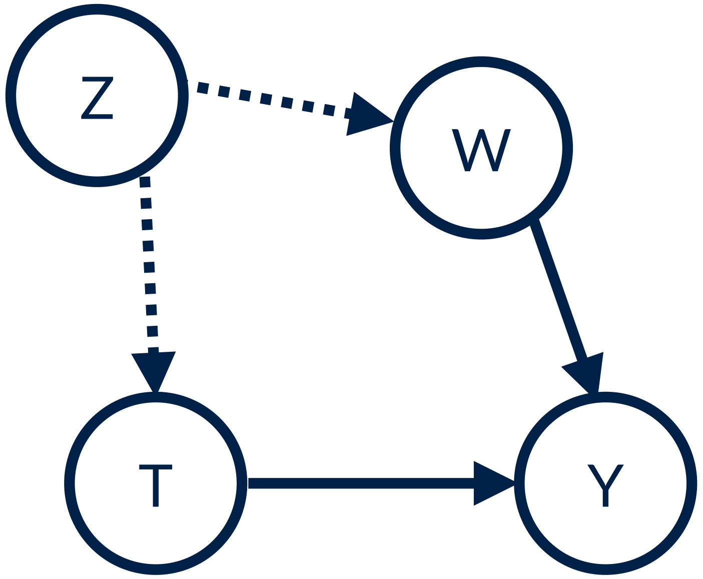
T = Drug, Y = recovery, W = weight, Z = unmeasured socioeconomic status
Z affects both weight and choice to receive treatment (but Z data was not recorded)

Can we compute the causal effect of T on Y, using W only (even though Z is not measured)?

Yes:, W satisfies the back-door path because:

(i) W blocks T <- Z -> W -> Y
(ii) W leaves the directed path from T to Y unperturbed
(iii) W is not a collider and is not a descendent of T

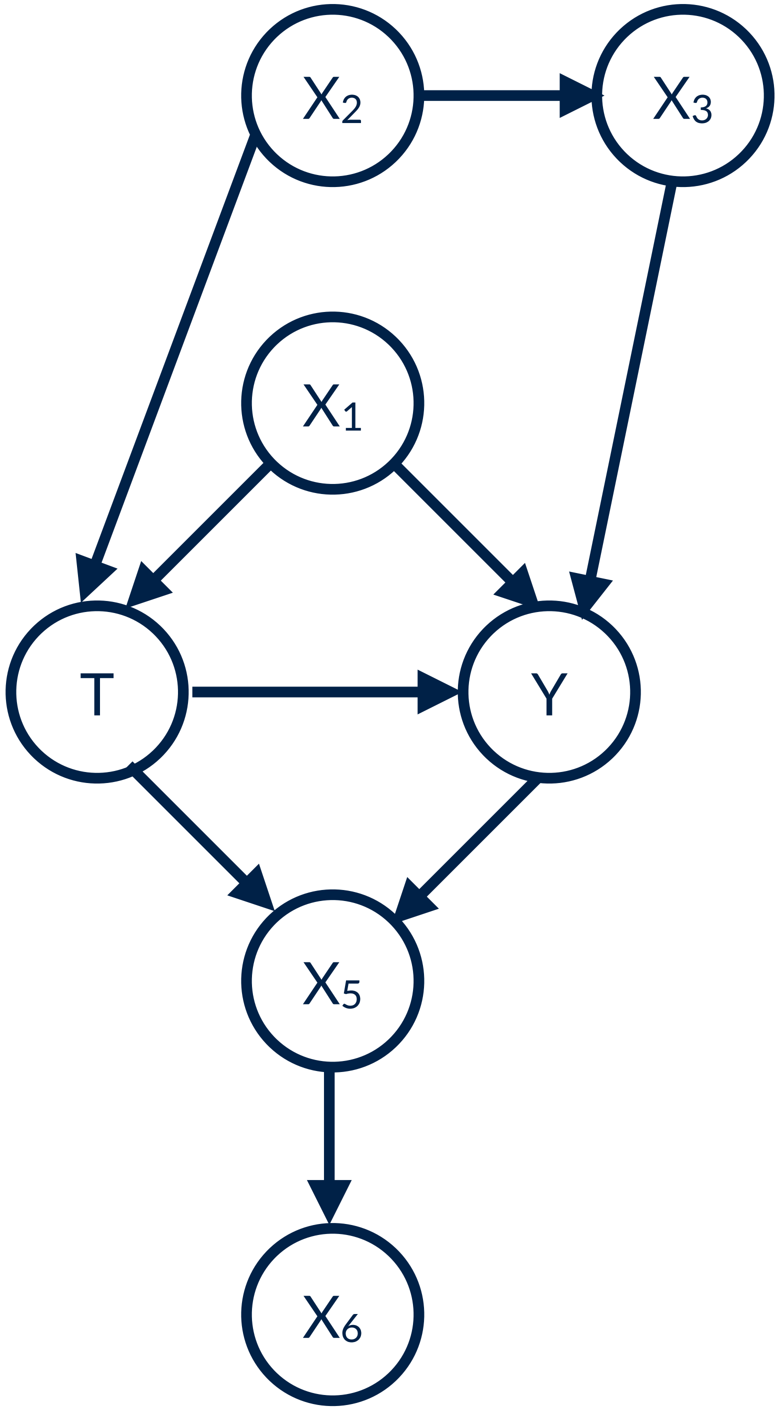
$$p(Y=y|do(T=t)) = \sum_w p(Y=y|T=t, W=w)p(W=w)$$



Pearl, Causal Inference in Statistics (2016)

The Backdoor Criterion: Example 2

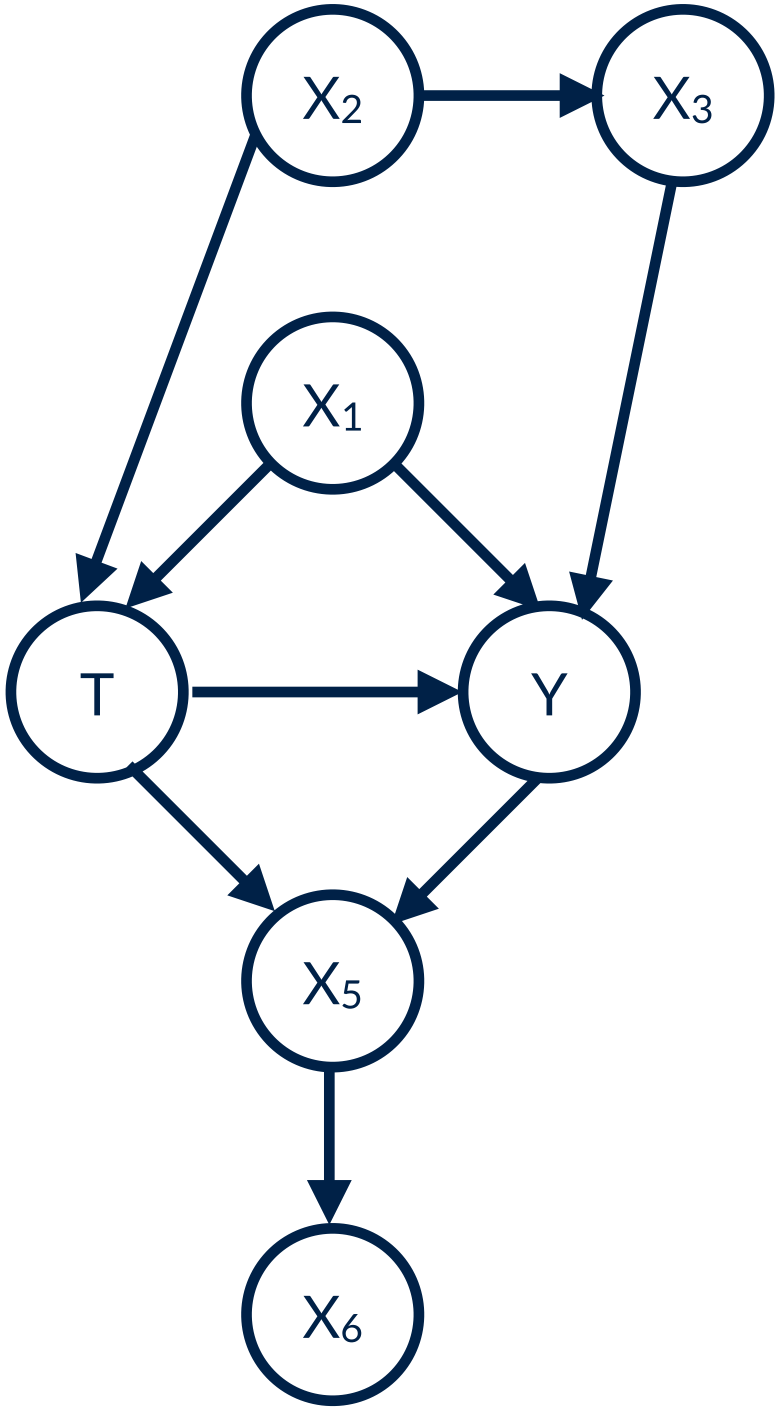
In computing the causal effect of T on Y, which variables should/not we condition on?



Pearl, Causal Inference in Statistics (2016)

The Backdoor Criterion: Example 2

In computing the causal effect of T on Y, which variables should/not we condition on?



Condition on X_1

Condition on either or both X_2, X_3

NOT X_5 and X_6

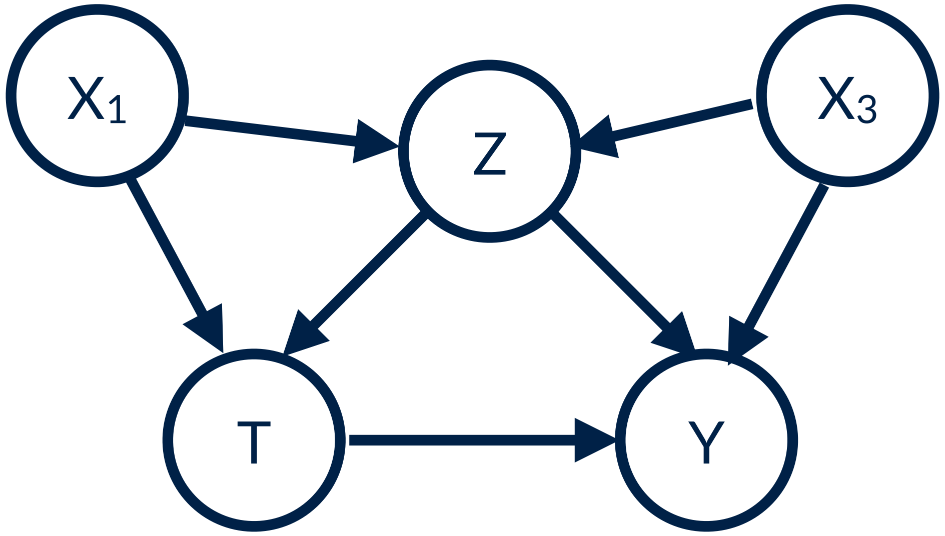
Because descendants of T and colliders, i.e.,

Conditioning opens a new path between T and X!

Pearl, Causal Inference in Statistics (2016)

The Backdoor Criterion: Example 3

Previous examples might have given the impression that
“We should never contain on colliders!”

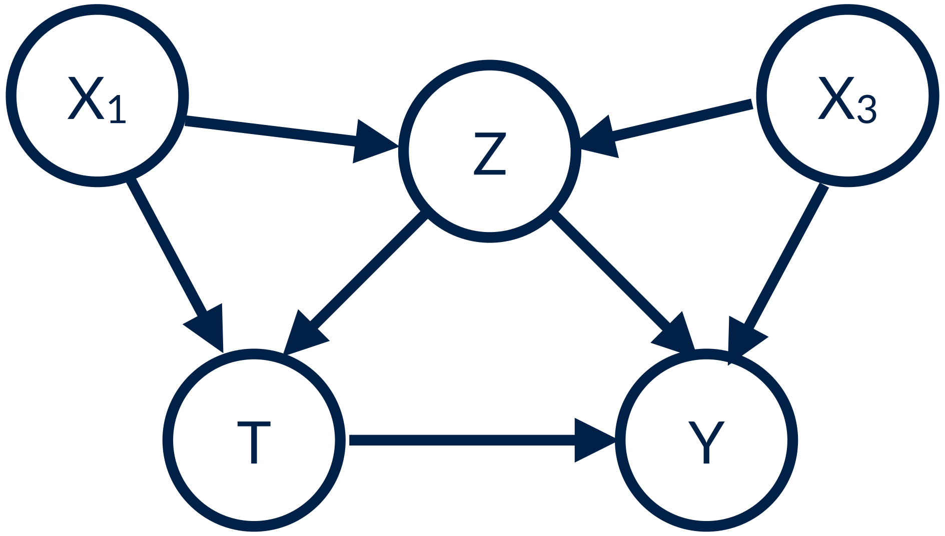


Pearl, Causal Inference in Statistics (2016)

The Backdoor Criterion: Example 3

Previous examples might have given the impression that
“We should never contain on colliders!”

This is not correct, because sometimes it’s unavoidable:
In this case, we need to condition on Z to stop the backdoor T <- Z -> Y
But then, this opens a new backdoor T <- X_1 -> Z <- X_2 -> Y
So we need to condition on {Z,X_1} or {Z,X_2} or {Z,X_1,X_2}
Therefore, even though Z is a collider, we managed to get causal identifiably



Pearl, Causal Inference in Statistics (2016)

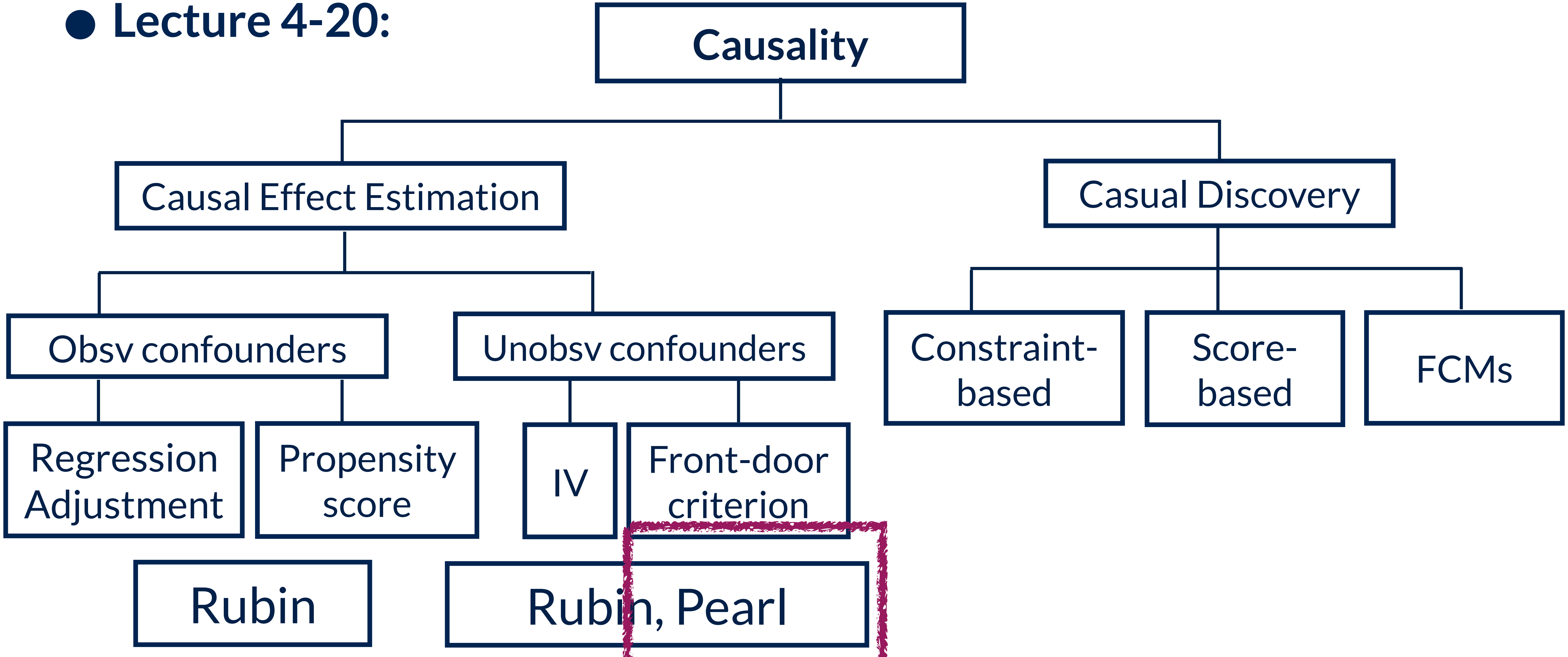
Rubin & Pearl

Rubin	Pearl
SUTVA	Implicit assumption of no interference between any pairs of individual
Unconfoundedness	Back-door criterion satisfied
Potential outcomes: $y_0^{(i)}$, $y_1^{(i)}$ Observed: $y_{\underline{0}}^{(i)}$, Unobserved: ${y^*}_1^{(i)}$	Counterfactuals are equivalent to individual unobserved outcomes in Rubin Do-operation

Pearl, Causal Inference in Statistics (2016)

Overview of the course

- **Lecture 1**: Introduction & Motivation, why do we care about causality? Why deriving causality from observational data is non-trivial.
- **Lecture 2:** Recap of probability theory, variables, events, conditional probabilities, independence, law of total probability, Bayes' rule
- **Lecture 3**: Recap of regression, multiple regression, graphs, SCM
- **Lecture 4-20:**



Causality

Causal Effect Estimation

Casual Discovery

Obsv confounders

Unobsv confounders

Constraint-based

Score-based

FCMs

Regression Adjustment

Propensity score

IV

Front-door criterion

Rubin

Rubin, Pearl