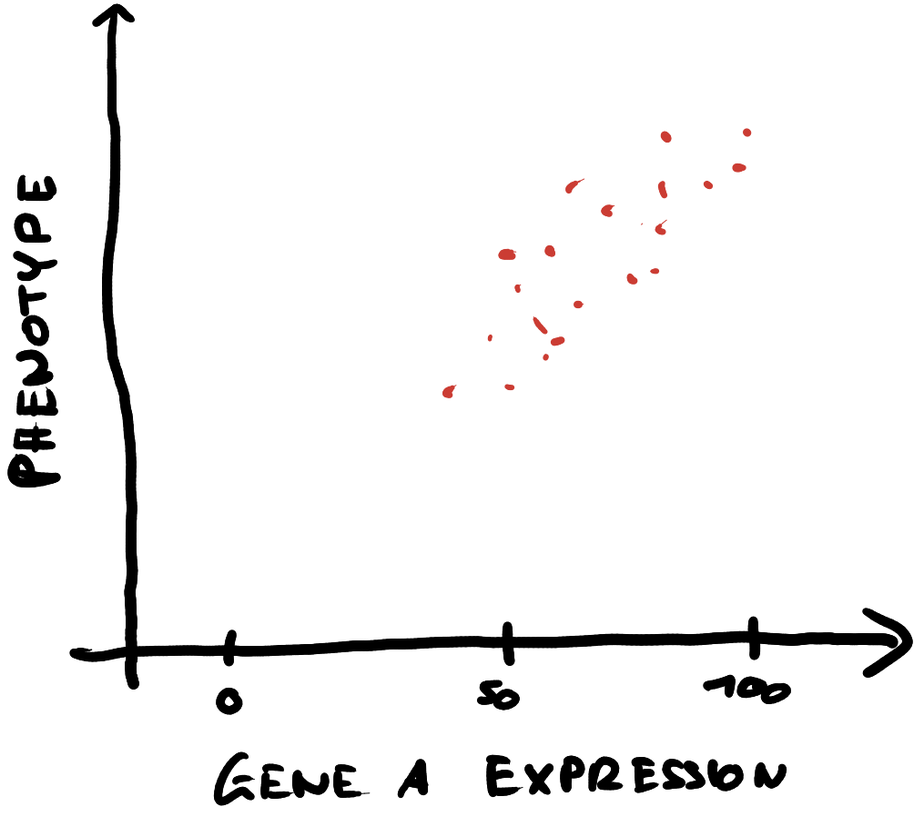
Causal Inference 101

Philipp Gabler

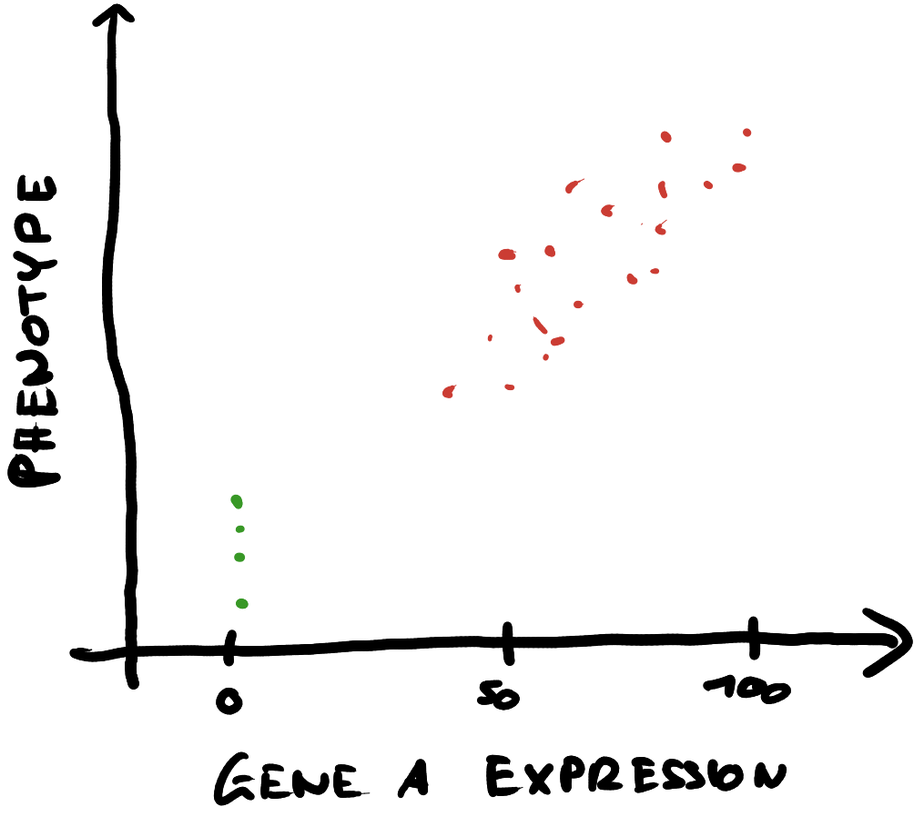
2021-03-17

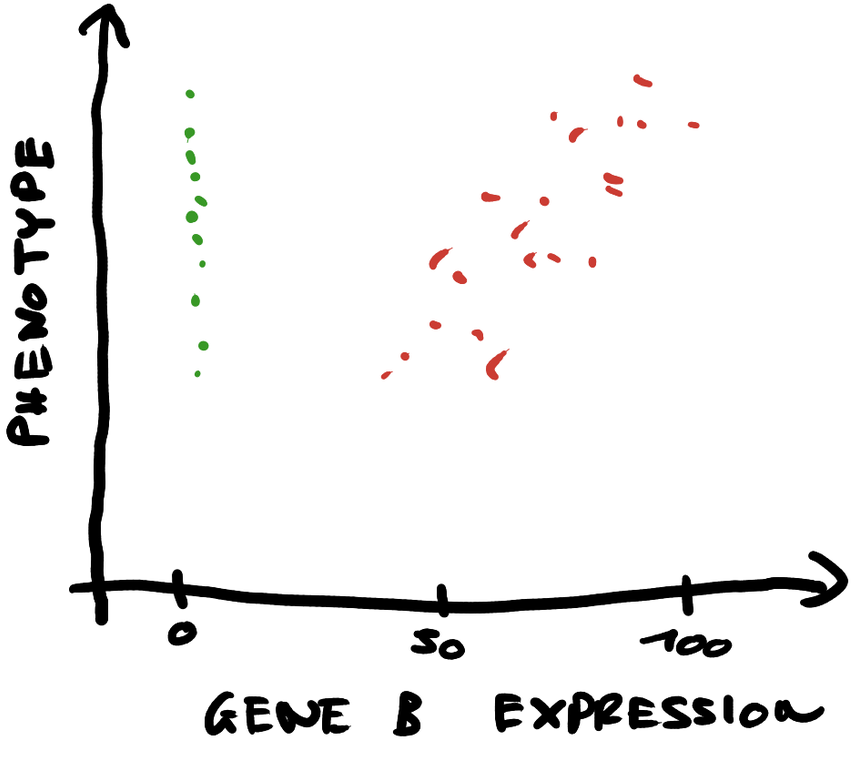
Two Genes


PHENOTYPE
0
50
100
GENE A EXPRESSION

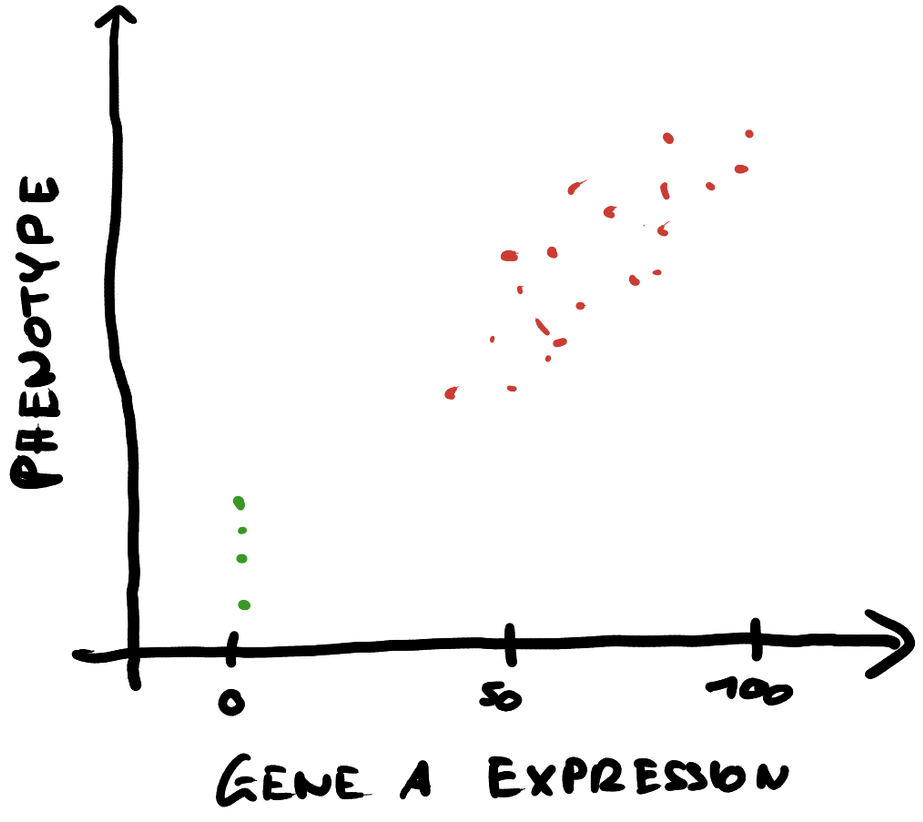

PHENOTYPE
0
50
100
GENE B EXPRESSION

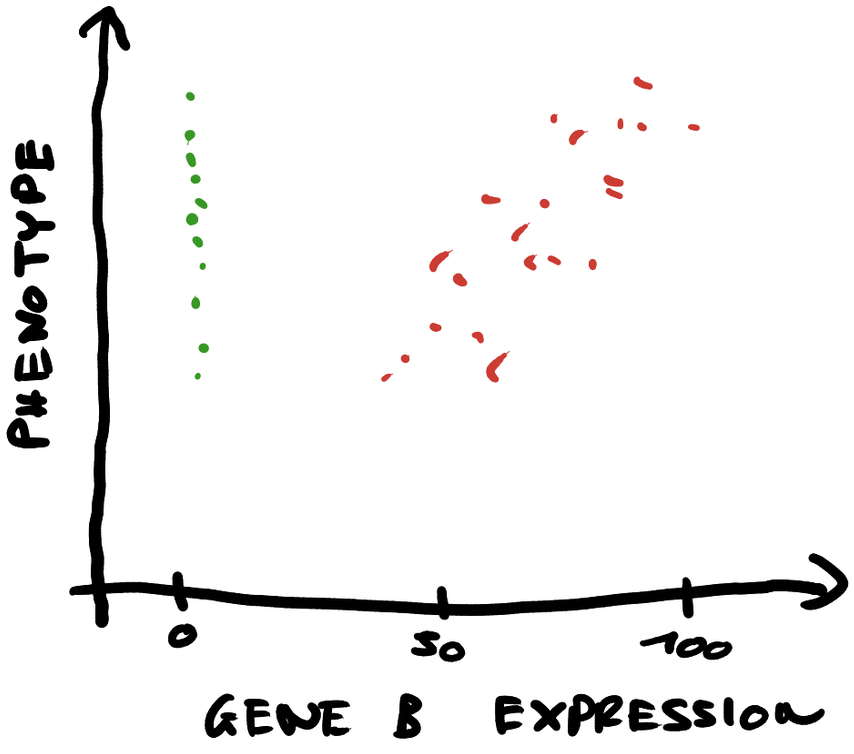
Two Genes


PHENOTYPE
0
50
100
GENE A EXPRESSION

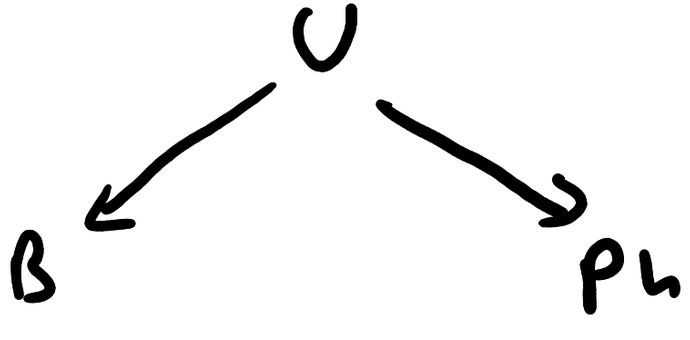

PHENOTYPE
0
50
100
GENE B EXPRESSION

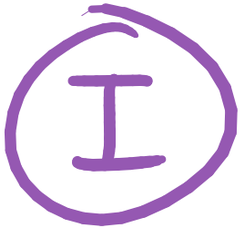
Two Genes


PHENOTYPE
0
50
100
GENE A EXPRESSION


PHENOTYPE
0
50
100
GENE B EXPRESSION


A
Ph


U
B
Ph



BASIC NOTIONS OF CAUSALITY

MEANING OF CAUSALITY

ASSOCIATION

↓

PREDICTION

OBSERVING X, WHAT CAN WE SAY ABOUT Y?

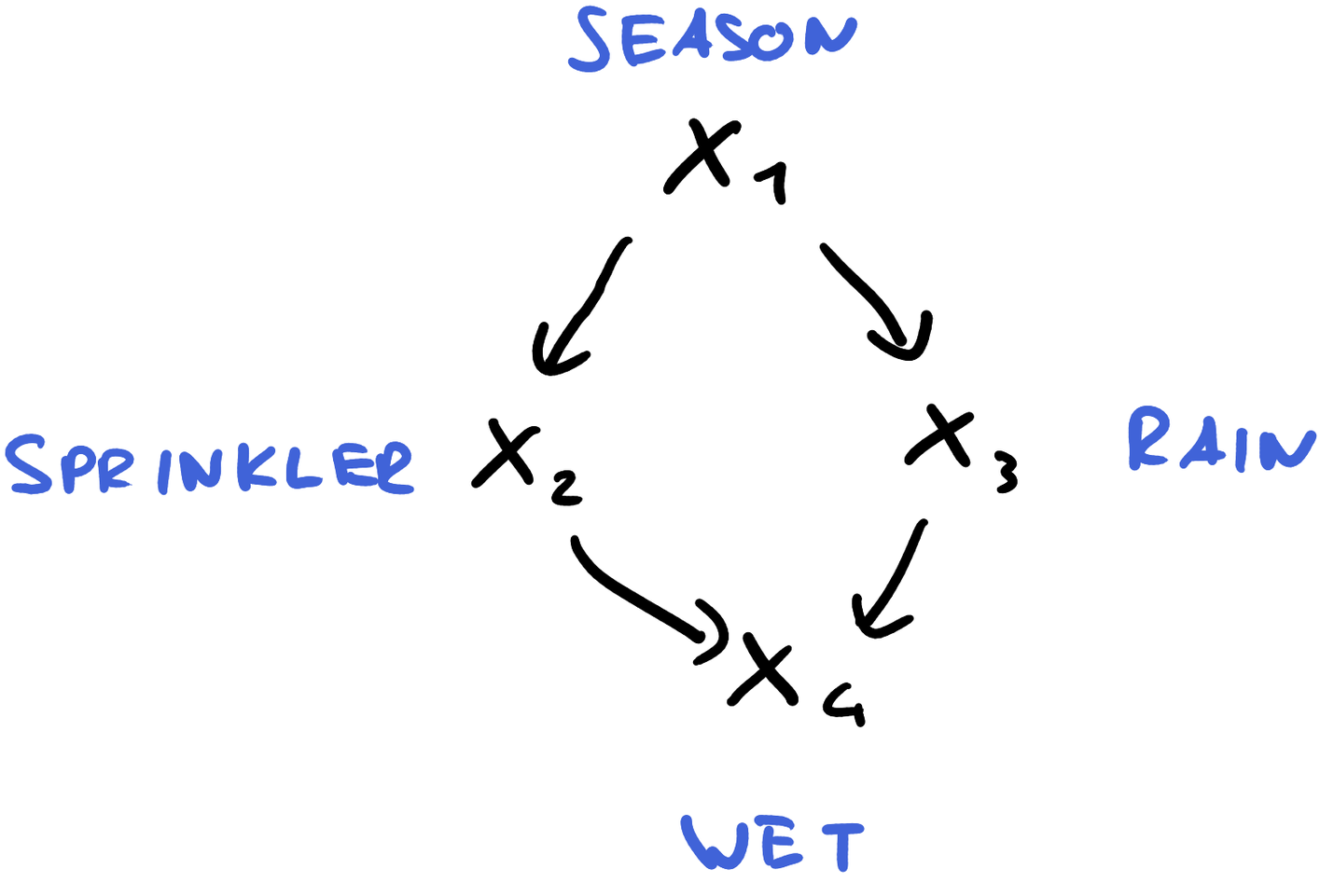
MECHANISM

↓

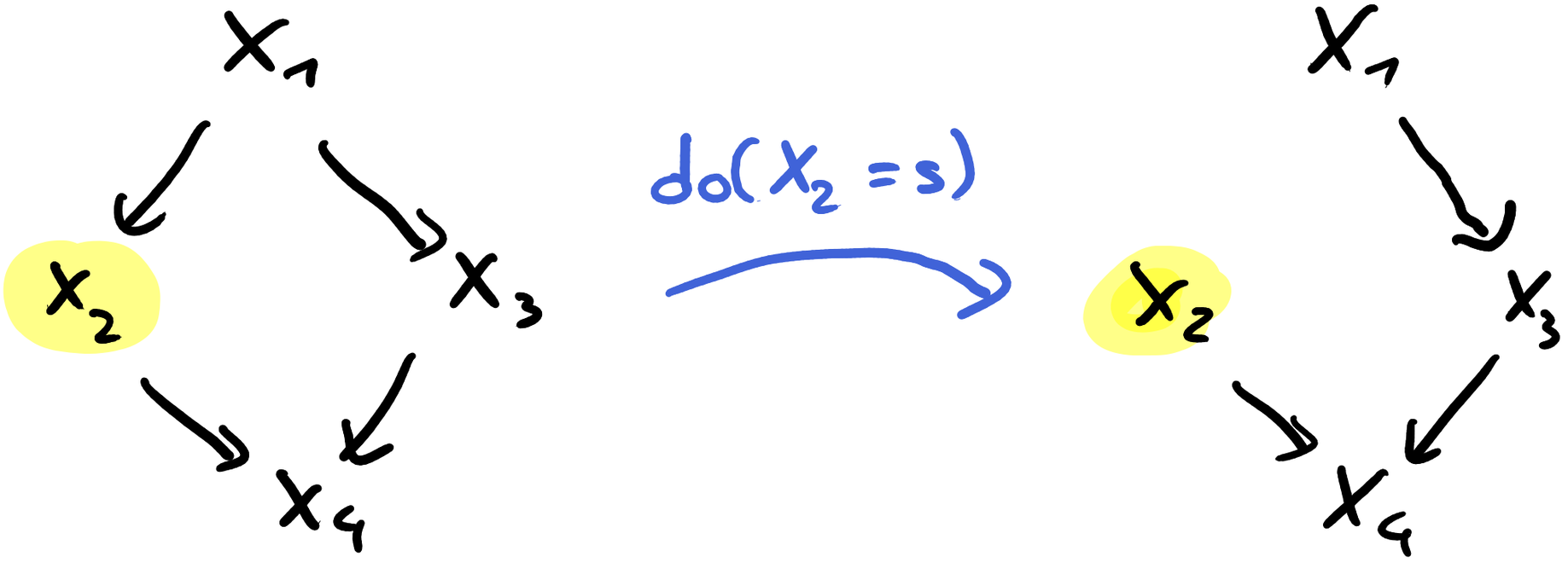
INTERVENTION

CHANGING X, WHAT WILL Y BE?

CAUSAL GRAPHS



INTERVENTIONS



$$P(X_1, X_2, X_3, X_4) =$$
$$P(X_1)\, P(X_2 \mid X_1)$$
$$P(X_3 \mid X_1)\, P(X_4 \mid X_2, X_3)$$

$$P(X_1, X_2, X_3, X_4) =$$
$$P(X_1)\, P(X_3 \mid X_1)$$
$$P'(X_2)\, P(X_4 \mid X_3, X_2)$$

INDEPENDENCE VS. MECHANISM

$$P(X_2, X_4) = P(X_2)\,P(X_4 \mid X_2)$$

SPRINKLER $\longrightarrow$ WET

$$= P(X_4)\,P(X_2 \mid X_4)$$

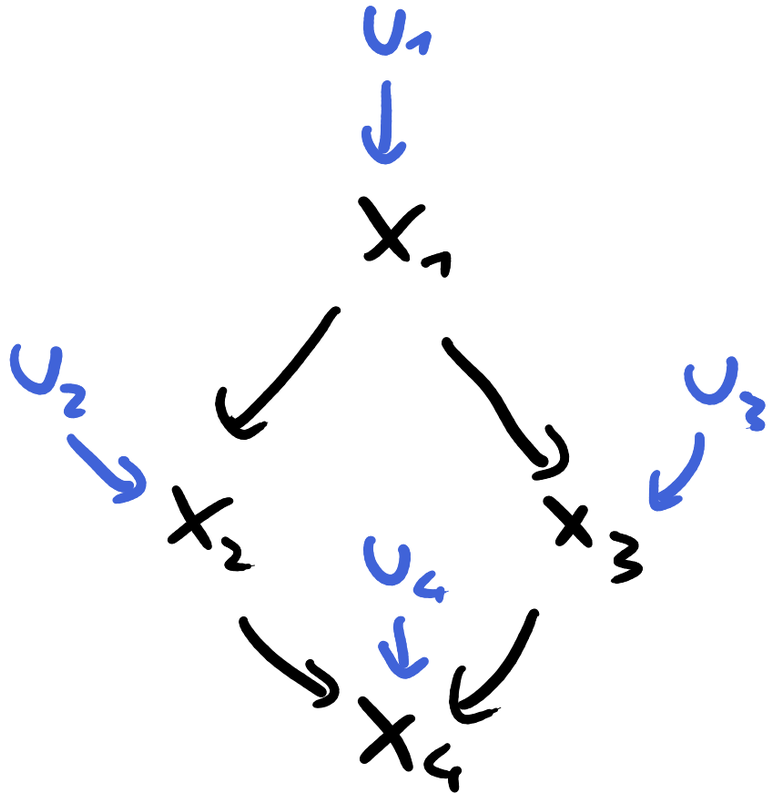
WET $\longrightarrow$ SPRINKLER

$$P(X_2 \mid do(X_4 = w)) = P(X_2)$$

$$P(X_4 \mid do(X_2 = s)) = P(X_4 \mid X_2 = s)$$

$$P(X_2 \mid X_4 = w) = \frac{P(X_2, X_4 = w)}{P(X_4 = w)}$$

STRUCTURAL EQUATIONS



$$x_1 = f_1(*, U_1)$$

$$x_2 = f_2(x_1, U_2)$$

$$\downarrow do(X_2 = s)$$

$$x_2 = s$$

$$x_3 = f_3(x_1, U_3)$$

$$x_4 = f_4(x_2, x_3, U_4)$$

U_i : INDEPENDENT RVs

f_i : DETERMINISTIC

LINEAR MODELS

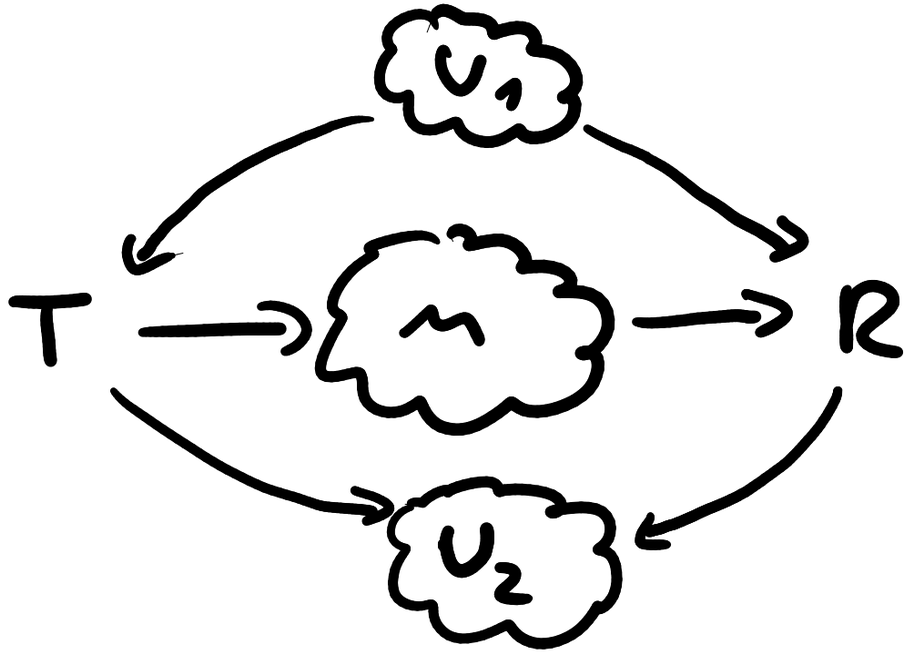
$$y = f(x, u) = \beta_0 + \beta_1 x + u$$

$$u \sim \mathcal{N}(0, \sigma)$$

$$\beta_1 = \frac{\partial}{\partial x} \mathbb{E}[Y = y \mid do(X = x)]$$

$$X \xrightarrow{\beta_1} Y$$

EFFECTS

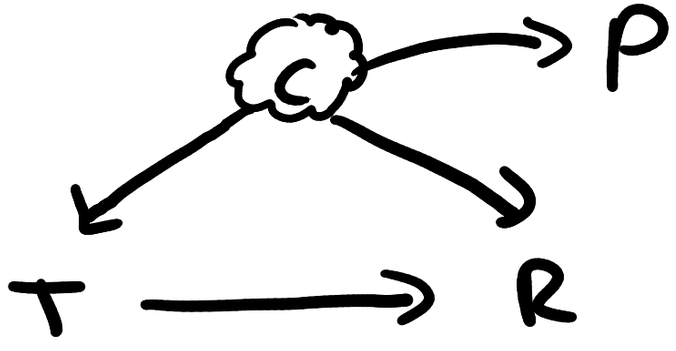


$$\mathbb{E}[R=v \mid do(T=1)] - \mathbb{E}[R=v \mid do(T=0)]$$

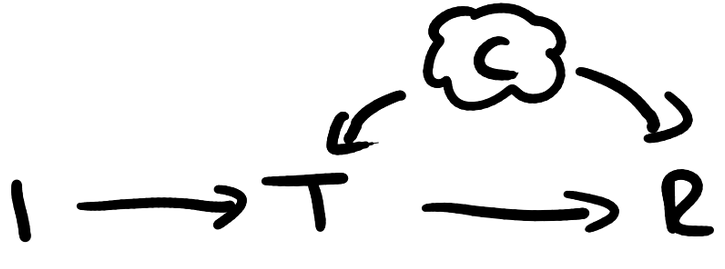
AVERAGE TREATMENT EFFECT

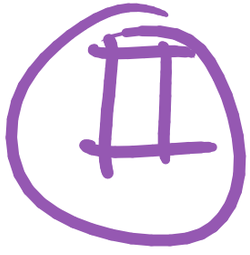
HELPFUL OBSERVATIONS

CONFOUNDER
PROXY



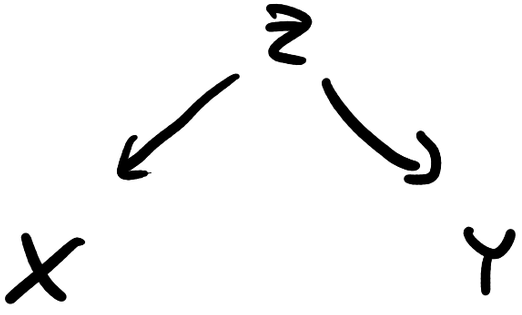
INSTRUMENTAL
VARIABLE



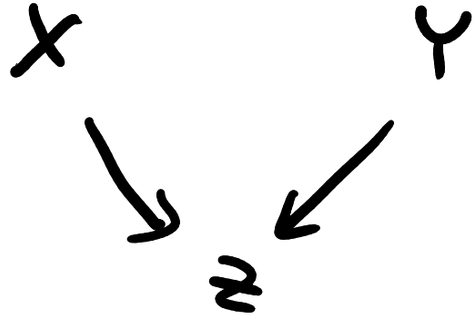


MODEL STRUCTURE

Common Patterns



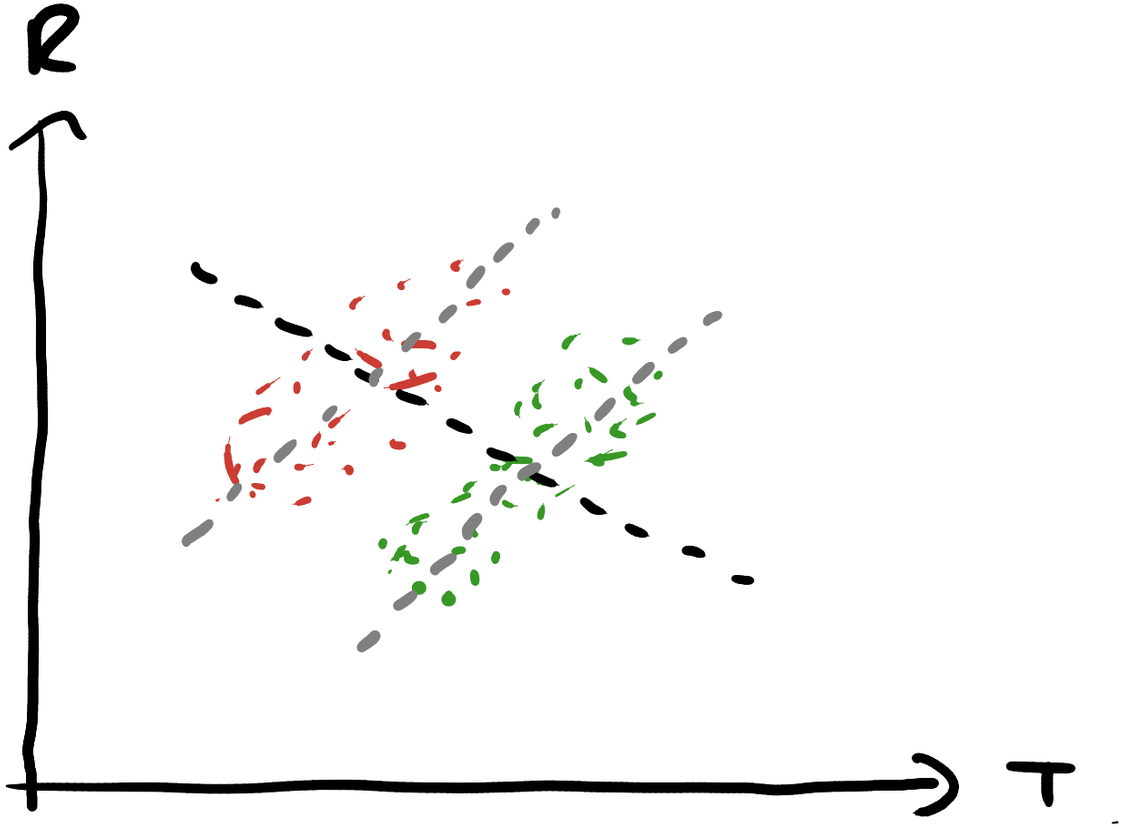
Confounder

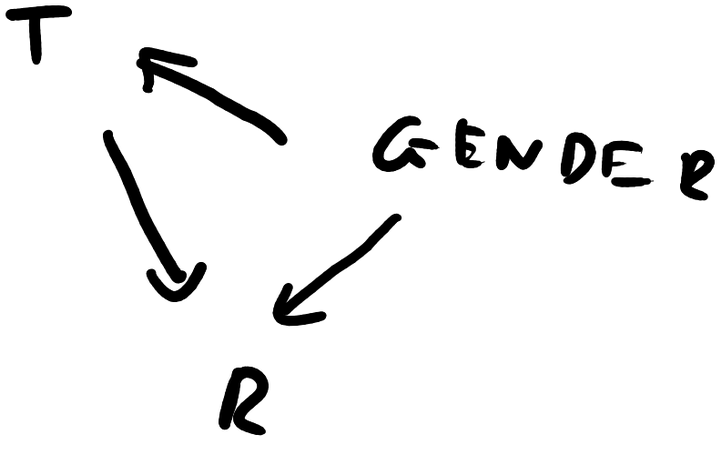


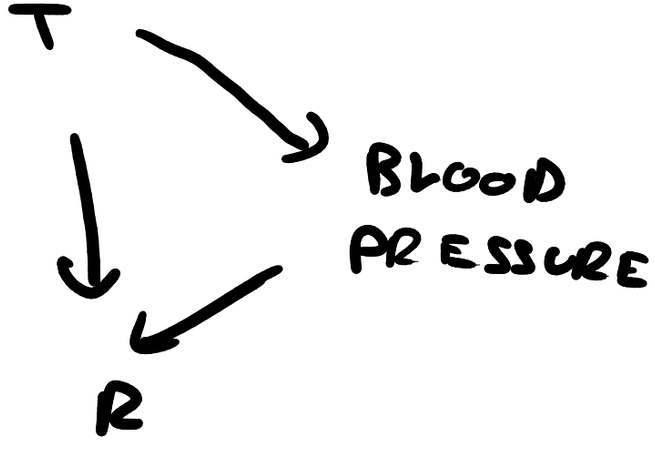
Collider

Mediator

SIMPSON'S PARADOX

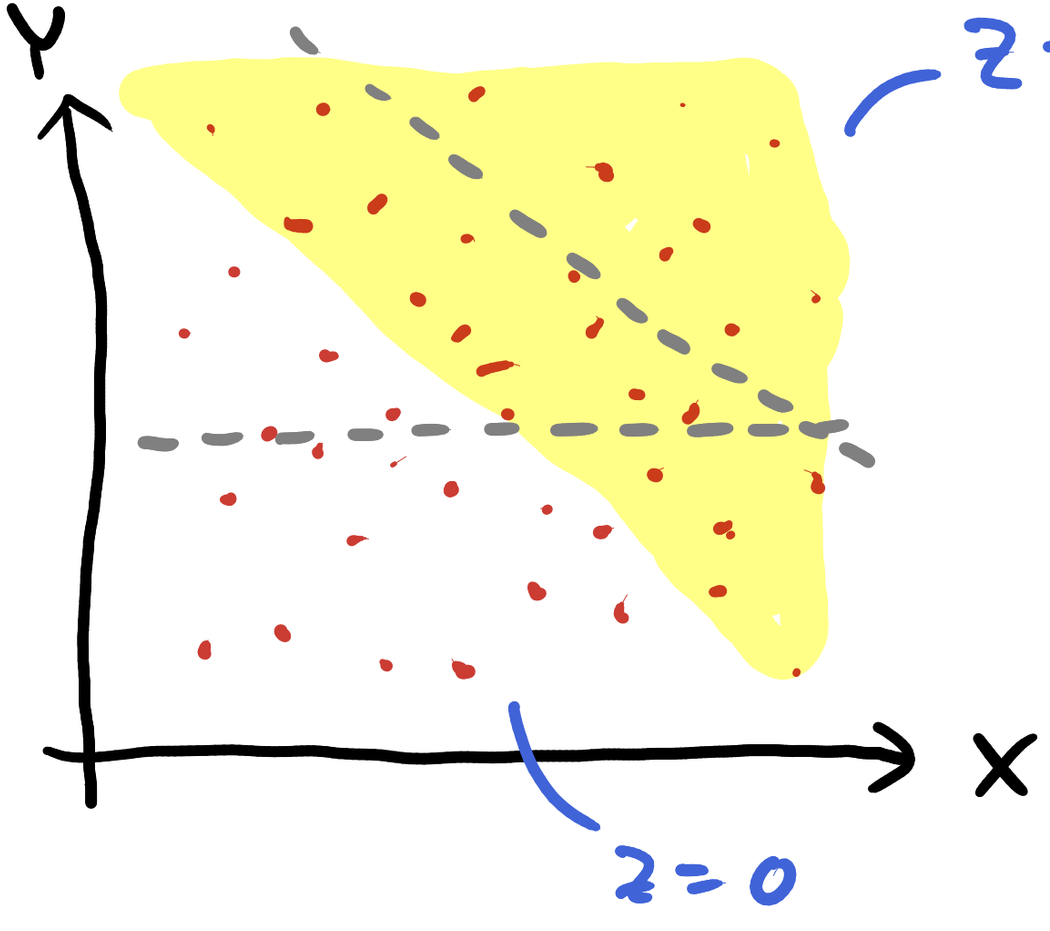


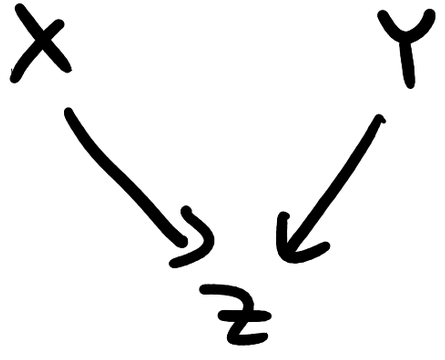


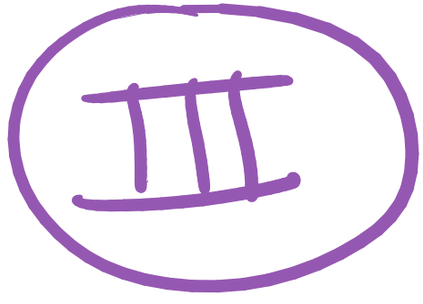


BERKSON'S PARADOX

SELECTION BIAS → SPURIOUS CORRELATION







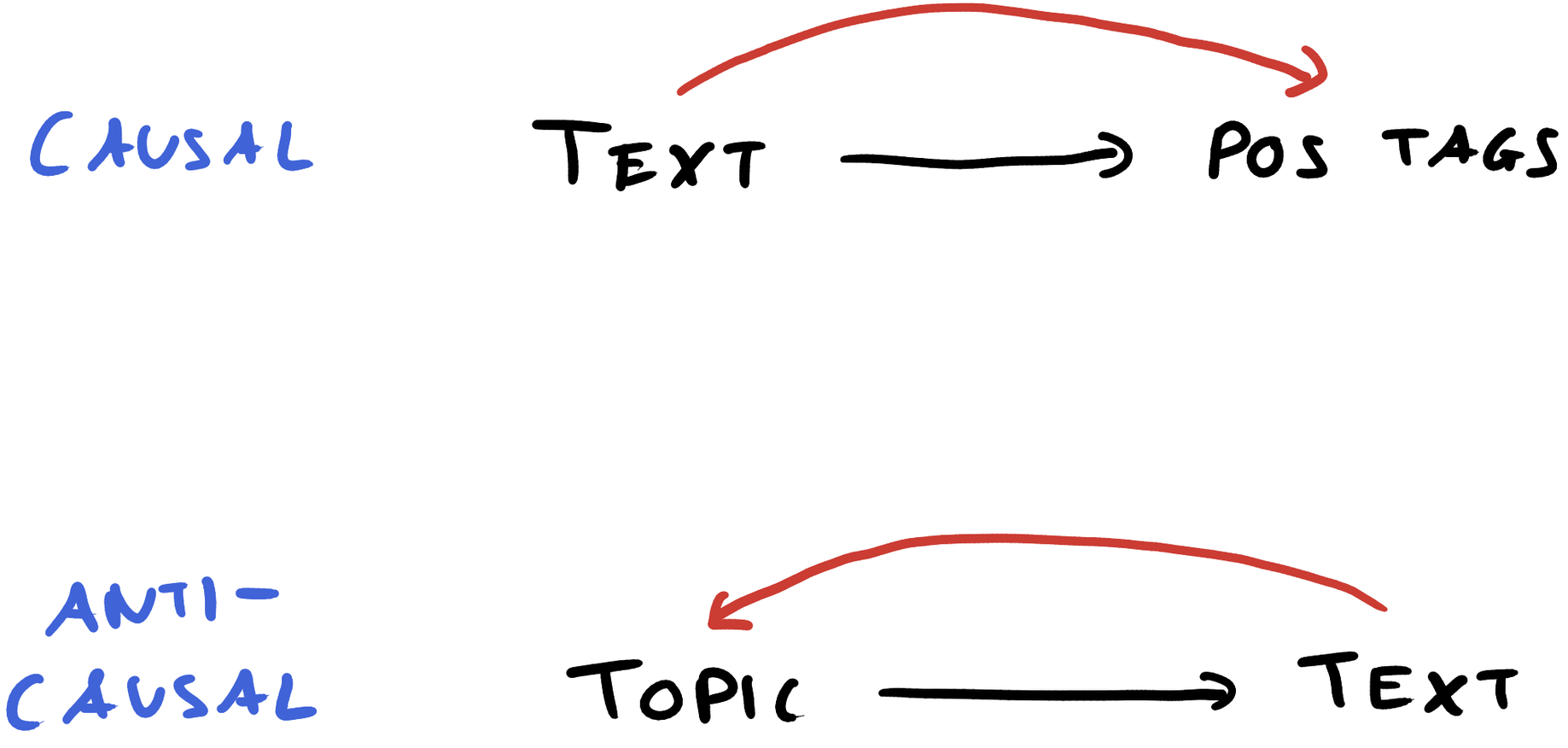
ADVANTAGES OF CAUSAL THINKING

KINDS OF DATA

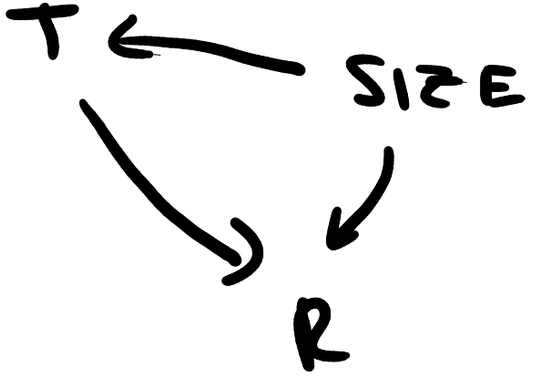
- EXPERIMENTAL DATA:
  - EXCHANGEABILITY
  - INVARIANCE

- OBSERVATIONAL DATA:
  - IDENTIFYABILITY?

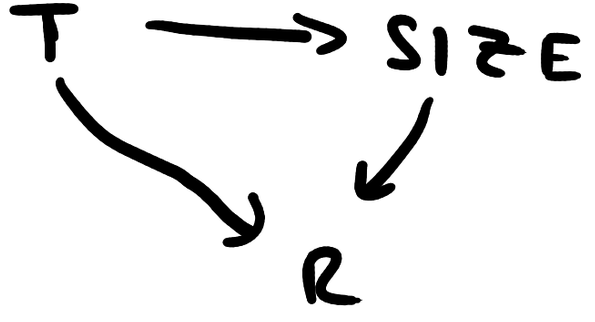
Inference Directions



When to Control?



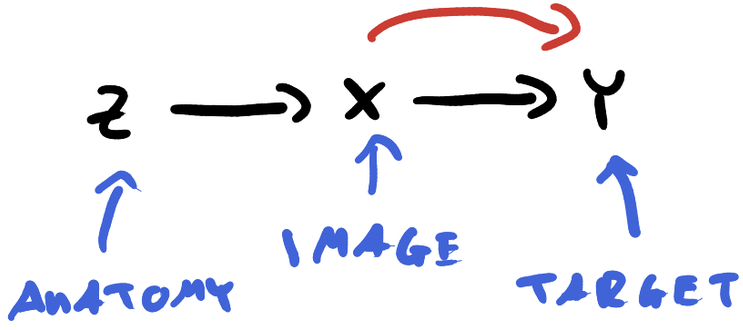
Yes

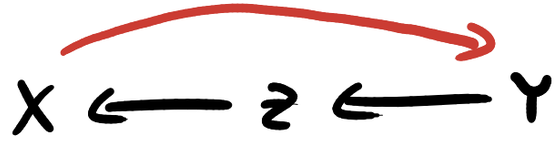


No

CLASSIFY WHAT'S WRONG

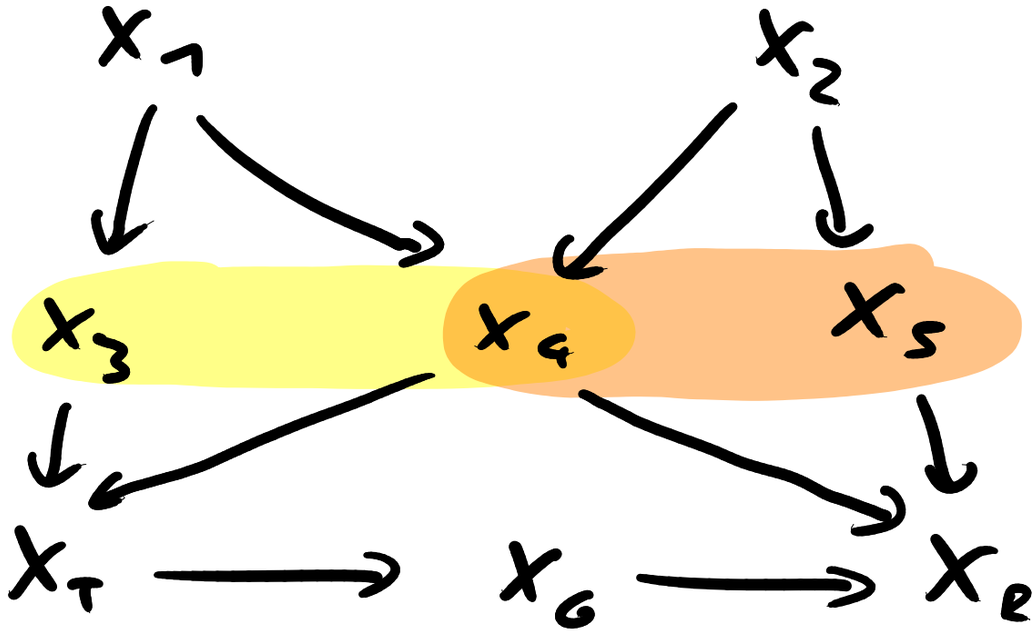
- POPULATION SHIFT
- ACQUISITION SHIFT
- ANNOTATION SHIFT
- PREVALENCE SHIFT
- MANIFESTATION SHIFT





SEE CASTRO & AL., 2020

IDENTIFIABILITY



ADJUSTMENT FOR $\{X_3, X_4\}$ OR FOR $\{X_4, X_5\}$ BY BACK-DOOR CRITERION