

Quantitative Methods

Human Sciences, 2020–21

Elias Nosrati

Lecture 8: 3 December 2020

Today

Today

- ▶ A third theory of inference.

Today

- ▶ A third theory of inference.
- ▶ Causal graphs.

Today

- ▶ A third theory of inference.
- ▶ Causal graphs.
- ▶ Three forms of systematic bias.

Neyman-Pearson hypothesis testing

Neyman-Pearson hypothesis testing

- ▶ Strict (but arbitrary) distinction between null hypothesis (H_0) and alternative hypothesis (H_1).

Neyman-Pearson hypothesis testing

- ▶ Strict (but arbitrary) distinction between null hypothesis (H_0) and alternative hypothesis (H_1).
- ▶ For instance: test whether X has causal effect β on Y .

Neyman-Pearson hypothesis testing

- ▶ Strict (but arbitrary) distinction between null hypothesis (H_0) and alternative hypothesis (H_1).
- ▶ For instance: test whether X has causal effect β on Y .
- ▶ $H_0 : \beta = 0$; $H_1 : \beta \neq 0$.

Neyman-Pearson hypothesis testing

- ▶ Strict (but arbitrary) distinction between null hypothesis (H_0) and alternative hypothesis (H_1).
- ▶ For instance: test whether X has causal effect β on Y .
- ▶ $H_0 : \beta = 0$; $H_1 : \beta \neq 0$.
- ▶ Choose *Type I error rate*: the probability of accepting H_1 when H_0 is actually true, α (not to be confused with model intercept!).

Neyman-Pearson hypothesis testing

- ▶ Strict (but arbitrary) distinction between null hypothesis (H_0) and alternative hypothesis (H_1).
- ▶ For instance: test whether X has causal effect β on Y .
- ▶ $H_0 : \beta = 0$; $H_1 : \beta \neq 0$.
- ▶ Choose *Type I error rate*: the probability of accepting H_1 when H_0 is actually true, α (not to be confused with model intercept!).
- ▶ Typical choices: $\alpha = 0.05$, $\alpha = 0.01$, $\alpha = 0.001$.

Neyman-Pearson hypothesis testing

- ▶ Strict (but arbitrary) distinction between null hypothesis (H_0) and alternative hypothesis (H_1).
- ▶ For instance: test whether X has causal effect β on Y .
- ▶ $H_0 : \beta = 0$; $H_1 : \beta \neq 0$.
- ▶ Choose *Type I error rate*: the probability of accepting H_1 when H_0 is actually true, α (not to be confused with model intercept!).
- ▶ Typical choices: $\alpha = 0.05$, $\alpha = 0.01$, $\alpha = 0.001$.
- ▶ Choose an estimator (e.g., OLS).

Neyman-Pearson hypothesis testing

- ▶ Strict (but arbitrary) distinction between null hypothesis (H_0) and alternative hypothesis (H_1).
- ▶ For instance: test whether X has causal effect β on Y .
- ▶ $H_0 : \beta = 0$; $H_1 : \beta \neq 0$.
- ▶ Choose *Type I error rate*: the probability of accepting H_1 when H_0 is actually true, α (not to be confused with model intercept!).
- ▶ Typical choices: $\alpha = 0.05$, $\alpha = 0.01$, $\alpha = 0.001$.
- ▶ Choose an estimator (e.g., OLS).
- ▶ Calculate $\hat{\beta}$.

Neyman-Pearson hypothesis testing

- ▶ Strict (but arbitrary) distinction between null hypothesis (H_0) and alternative hypothesis (H_1).
- ▶ For instance: test whether X has causal effect β on Y .
- ▶ $H_0 : \beta = 0$; $H_1 : \beta \neq 0$.
- ▶ Choose *Type I error rate*: the probability of accepting H_1 when H_0 is actually true, α (not to be confused with model intercept!).
- ▶ Typical choices: $\alpha = 0.05$, $\alpha = 0.01$, $\alpha = 0.001$.
- ▶ Choose an estimator (e.g., OLS).
- ▶ Calculate $\hat{\beta}$.
- ▶ Assume sample is large enough for Central Limit Theorem to kick in (otherwise use t -distribution).

Neyman-Pearson hypothesis testing

- ▶ Strict (but arbitrary) distinction between null hypothesis (H_0) and alternative hypothesis (H_1).
- ▶ For instance: test whether X has causal effect β on Y .
- ▶ $H_0 : \beta = 0$; $H_1 : \beta \neq 0$.
- ▶ Choose *Type I error rate*: the probability of accepting H_1 when H_0 is actually true, α (not to be confused with model intercept!).
- ▶ Typical choices: $\alpha = 0.05$, $\alpha = 0.01$, $\alpha = 0.001$.
- ▶ Choose an estimator (e.g., OLS).
- ▶ Calculate $\hat{\beta}$.
- ▶ Assume sample is large enough for Central Limit Theorem to kick in (otherwise use t -distribution).
- ▶ Then

$$\hat{\beta} \mid (\beta = 0) \sim \mathcal{N}(0, \sigma_{\hat{\beta}}^2),$$

Neyman-Pearson hypothesis testing

- ▶ Strict (but arbitrary) distinction between null hypothesis (H_0) and alternative hypothesis (H_1).
- ▶ For instance: test whether X has causal effect β on Y .
- ▶ $H_0 : \beta = 0$; $H_1 : \beta \neq 0$.
- ▶ Choose *Type I error rate*: the probability of accepting H_1 when H_0 is actually true, α (not to be confused with model intercept!).
- ▶ Typical choices: $\alpha = 0.05$, $\alpha = 0.01$, $\alpha = 0.001$.
- ▶ Choose an estimator (e.g., OLS).
- ▶ Calculate $\hat{\beta}$.
- ▶ Assume sample is large enough for Central Limit Theorem to kick in (otherwise use t -distribution).
- ▶ Then

$$\hat{\beta} \mid (\beta = 0) \sim \mathcal{N}(0, \sigma_{\hat{\beta}}^2), \quad z = \frac{\hat{\beta} - \beta}{\hat{\sigma}(\hat{\beta})} \sim \mathcal{N}(0, 1).$$

Neyman-Pearson hypothesis testing

- ▶ Strict (but arbitrary) distinction between null hypothesis (H_0) and alternative hypothesis (H_1).
- ▶ For instance: test whether X has causal effect β on Y .
- ▶ $H_0 : \beta = 0$; $H_1 : \beta \neq 0$.
- ▶ Choose *Type I error rate*: the probability of accepting H_1 when H_0 is actually true, α (not to be confused with model intercept!).
- ▶ Typical choices: $\alpha = 0.05$, $\alpha = 0.01$, $\alpha = 0.001$.
- ▶ Choose an estimator (e.g., OLS).
- ▶ Calculate $\hat{\beta}$.
- ▶ Assume sample is large enough for Central Limit Theorem to kick in (otherwise use t -distribution).
- ▶ Then

$$\hat{\beta} \mid (\beta = 0) \sim \mathcal{N}(0, \sigma_{\hat{\beta}}^2), \quad z = \frac{\hat{\beta} - \beta}{\hat{\sigma}(\hat{\beta})} \sim \mathcal{N}(0, 1).$$

- ▶ Note: all tests are done “under” (i.e., assuming) H_0 .

Neyman-Pearson hypothesis testing (cont.)

Neyman-Pearson hypothesis testing (cont.)

- ▶ Let $\alpha = 0.05$ (“significance level”).

Neyman-Pearson hypothesis testing (cont.)

- ▶ Let $\alpha = 0.05$ (“significance level”).
- ▶ Under the standard Normal distribution, $\mathbb{P}(|z| \geq 1.96) = \alpha$.

Neyman-Pearson hypothesis testing (cont.)

- ▶ Let $\alpha = 0.05$ (“significance level”).
- ▶ Under the standard Normal distribution, $\mathbb{P}(|z| \geq 1.96) = \alpha$.
- ▶ 1.96 is called the “critical value”.

Neyman-Pearson hypothesis testing (cont.)

- ▶ Let $\alpha = 0.05$ (“significance level”).
- ▶ Under the standard Normal distribution, $\mathbb{P}(|z| \geq 1.96) = \alpha$.
- ▶ 1.96 is called the “critical value”.
- ▶ Decision rule:

Neyman-Pearson hypothesis testing (cont.)

- ▶ Let $\alpha = 0.05$ (“significance level”).
- ▶ Under the standard Normal distribution, $\mathbb{P}(|z| \geq 1.96) = \alpha$.
- ▶ 1.96 is called the “critical value”.
- ▶ Decision rule:
 - ★ Reject $H_0: \beta \neq 0$ if $z > 1.96$.

Neyman-Pearson hypothesis testing (cont.)

- ▶ Let $\alpha = 0.05$ (“significance level”).
- ▶ Under the standard Normal distribution, $\mathbb{P}(|z| \geq 1.96) = \alpha$.
- ▶ 1.96 is called the “critical value”.
- ▶ Decision rule:
 - ★ Reject $H_0: \beta \neq 0$ if $z > 1.96$.
 - ★ Fail to reject $H_0: \beta = 0$ if $z \leq 1.96$.

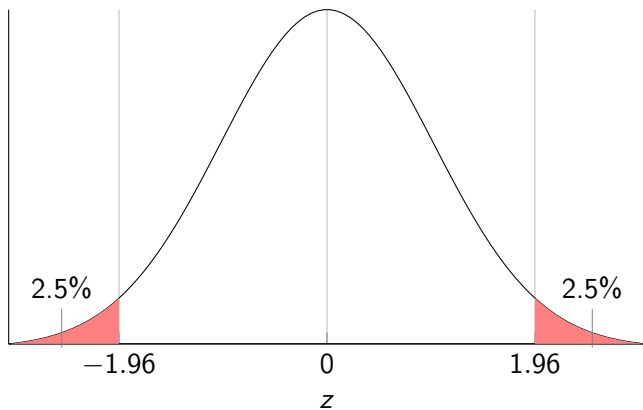
Neyman-Pearson hypothesis testing (cont.)

- ▶ Let $\alpha = 0.05$ (“significance level”).
- ▶ Under the standard Normal distribution, $\mathbb{P}(|z| \geq 1.96) = \alpha$.
- ▶ 1.96 is called the “critical value”.
- ▶ Decision rule:
 - ★ Reject $H_0: \beta \neq 0$ if $z > 1.96$.
 - ★ Fail to reject $H_0: \beta = 0$ if $z \leq 1.96$.
- ▶ p -value: the probability of obtaining something at least as extreme as the observed z -statistic, assuming H_0 is true.

Neyman-Pearson hypothesis testing (cont.)

- ▶ Let $\alpha = 0.05$ (“significance level”).
- ▶ Under the standard Normal distribution, $\mathbb{P}(|z| \geq 1.96) = \alpha$.
- ▶ 1.96 is called the “critical value”.
- ▶ Decision rule:
 - ★ Reject $H_0: \beta \neq 0$ if $z > 1.96$.
 - ★ Fail to reject $H_0: \beta = 0$ if $z \leq 1.96$.
- ▶ p -value: the probability of obtaining something at least as extreme as the observed z -statistic, assuming H_0 is true.
- ▶ Common mistake: the p -value is *not* the probability that H_1 is false. (Why not?)

Neyman-Pearson hypothesis testing (cont.)



Neyman-Pearson hypothesis testing: example

```
##
## Call:
## lm(formula = y ~ x, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -14.8208  -3.1487   0.0409   3.1336  13.1631
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.23379    0.50486   0.463   0.644
## x            5.11564    0.08244  62.052 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.755 on 498 degrees of freedom
## Multiple R-squared:  0.8855, Adjusted R-squared:  0.8852
## F-statistic: 3851 on 1 and 498 DF, p-value: < 2.2e-16
```

Causal graphs



Causal graphs

$$X \longrightarrow Y$$



- Causal graphs visually encode causal relationships between random variables.

Causal graphs

$$X \longrightarrow Y$$



- ▶ Causal graphs visually encode causal relationships between random variables.
- ▶ A graph is composed of two objects:

Causal graphs



- ▶ Causal graphs visually encode causal relationships between random variables.
- ▶ A graph is composed of two objects:
 1. Nodes, representing random variables.

Causal graphs



- ▶ Causal graphs visually encode causal relationships between random variables.
- ▶ A graph is composed of two objects:
 1. Nodes, representing random variables.
 2. Arrows, representing the causal relationship between the nodes.

Causal graphs



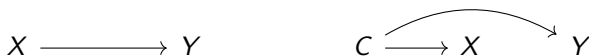
- ▶ Causal graphs visually encode causal relationships between random variables.
- ▶ A graph is composed of two objects:
 1. Nodes, representing random variables.
 2. Arrows, representing the causal relationship between the nodes.
- ▶ The presence of an arrow indicates the presence of a (known) causal effect.

Causal graphs



- ▶ Causal graphs visually encode causal relationships between random variables.
- ▶ A graph is composed of two objects:
 1. Nodes, representing random variables.
 2. Arrows, representing the causal relationship between the nodes.
- ▶ The presence of an arrow indicates the presence of a (known) causal effect.
- ▶ The lack of an arrow indicates the (known) absence of a causal effect.

Causal graphs



- ▶ Causal graphs visually encode causal relationships between random variables.
- ▶ A graph is composed of two objects:
 1. Nodes, representing random variables.
 2. Arrows, representing the causal relationship between the nodes.
- ▶ The presence of an arrow indicates the presence of a (known) causal effect.
- ▶ The lack of an arrow indicates the (known) absence of a causal effect.
- ▶ A standard causal graph does not express the magnitude or sign of a causal effect.

Causal graphs



- ▶ Causal graphs visually encode causal relationships between random variables.
- ▶ A graph is composed of two objects:
 1. Nodes, representing random variables.
 2. Arrows, representing the causal relationship between the nodes.
- ▶ The presence of an arrow indicates the presence of a (known) causal effect.
- ▶ The lack of an arrow indicates the (known) absence of a causal effect.
- ▶ A standard causal graph does not express the magnitude or sign of a causal effect.
- ▶ Conditional on its direct causes, any variable in a causal graph is independent of any other variable for which it is not a cause.

Causal graphs



- ▶ Causal graphs visually encode causal relationships between random variables.
- ▶ A graph is composed of two objects:
 1. Nodes, representing random variables.
 2. Arrows, representing the causal relationship between the nodes.
- ▶ The presence of an arrow indicates the presence of a (known) causal effect.
- ▶ The lack of an arrow indicates the (known) absence of a causal effect.
- ▶ A standard causal graph does not express the magnitude or sign of a causal effect.
- ▶ Conditional on its direct causes, any variable in a causal graph is independent of any other variable for which it is not a cause.
- ▶ Implication: the common causes of any pair of variables in the graph must also be in the graph.

A randomised experiment

$$X \longrightarrow Y$$

A randomised experiment

$$X \longrightarrow Y$$

- ▶ Smoking (X) causes lung cancer (Y):

$$\mathbb{E}(Y_1) \neq \mathbb{E}(Y_0).$$

A randomised experiment

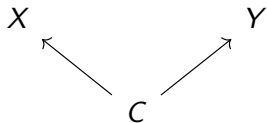
$$X \longrightarrow Y$$

- ▶ Smoking (X) causes lung cancer (Y):

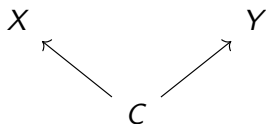
$$\mathbb{E}(Y_1) \neq \mathbb{E}(Y_0).$$

- ▶ The causal graph depicts this knowledge for an experiment in which cigarettes are randomly assigned to a treatment group.

Confounding bias



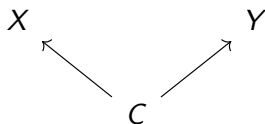
Confounding bias



- Carrying a lighter (X) does not cause lung cancer (Y):

$$\mathbb{E}(Y_1) = \mathbb{E}(Y_0).$$

Confounding bias

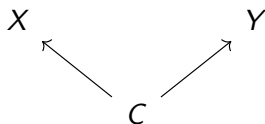


- ▶ Carrying a lighter (X) does not cause lung cancer (Y):

$$\mathbb{E}(Y_1) = \mathbb{E}(Y_0).$$

- ▶ The causal graph depicts this knowledge through the lack of an arrow between X and Y .

Confounding bias

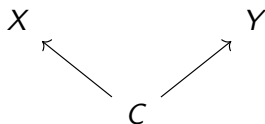


- ▶ Carrying a lighter (X) does not cause lung cancer (Y):

$$\mathbb{E}(Y_1) = \mathbb{E}(Y_0).$$

- ▶ The causal graph depicts this knowledge through the lack of an arrow between X and Y .
- ▶ But smoking (C) is a common cause of both carrying a lighter and lung cancer, hence the arrow from C to both X and Y .

Confounding bias

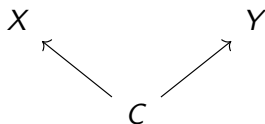


- ▶ Carrying a lighter (X) does not cause lung cancer (Y):

$$\mathbb{E}(Y_1) = \mathbb{E}(Y_0).$$

- ▶ The causal graph depicts this knowledge through the lack of an arrow between X and Y .
- ▶ But smoking (C) is a common cause of both carrying a lighter and lung cancer, hence the arrow from C to both X and Y .
- ▶ Without conditioning on C , we observe an association between X and Y that flows through their common cause C .

Confounding bias

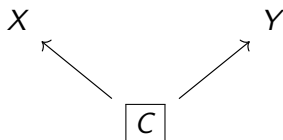


- ▶ Carrying a lighter (X) does not cause lung cancer (Y):

$$\mathbb{E}(Y_1) = \mathbb{E}(Y_0).$$

- ▶ The causal graph depicts this knowledge through the lack of an arrow between X and Y .
- ▶ But smoking (C) is a common cause of both carrying a lighter and lung cancer, hence the arrow from C to both X and Y .
- ▶ Without conditioning on C , we observe an association between X and Y that flows through their common cause C .
- ▶ C is a *confounder* of the association between X and Y .

Confounding bias

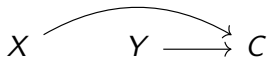


- ▶ Carrying a lighter (X) does not cause lung cancer (Y):

$$\mathbb{E}(Y_1) = \mathbb{E}(Y_0).$$

- ▶ The causal graph depicts this knowledge through the lack of an arrow between X and Y .
- ▶ But smoking (C) is a common cause of both carrying a lighter and lung cancer, hence the arrow from C to both X and Y .
- ▶ Without conditioning on C , we observe an association between X and Y that flows through their common cause C .
- ▶ C is a *confounder* of the association between X and Y .
- ▶ If we do condition on C , we block the association between X and Y (via conditional independence).

Selection bias



Selection bias



- ▶ A certain genetic haplotype (X) has no causal effect on the risk of neurodegenerative disease (Y):

$$\mathbb{P}(Y_1 = 1) = \mathbb{P}(Y_0 = 1).$$

Selection bias



- ▶ A certain genetic haplotype (X) has no causal effect on the risk of neurodegenerative disease (Y):

$$\mathbb{P}(Y_1 = 1) = \mathbb{P}(Y_0 = 1).$$

- ▶ Both X and Y cause depression (C).

Selection bias



- ▶ A certain genetic haplotype (X) has no causal effect on the risk of neurodegenerative disease (Y):

$$\mathbb{P}(Y_1 = 1) = \mathbb{P}(Y_0 = 1).$$

- ▶ Both X and Y cause depression (C).
- ▶ The common effect C is called a *collider* because of the colliding arrows on the path $X \rightarrow C \leftarrow Y$.

Selection bias

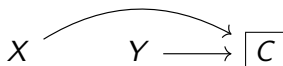


- ▶ A certain genetic haplotype (X) has no causal effect on the risk of neurodegenerative disease (Y):

$$\mathbb{P}(Y_1 = 1) = \mathbb{P}(Y_0 = 1).$$

- ▶ Both X and Y cause depression (C).
- ▶ The common effect C is called a *collider* because of the colliding arrows on the path $X \rightarrow C \leftarrow Y$.
- ▶ Colliders block the flow of association along the path on which they lie: X and Y are independent because the only path between them is blocked by C .

Selection bias

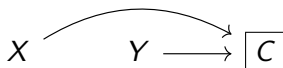


- ▶ A certain genetic haplotype (X) has no causal effect on the risk of neurodegenerative disease (Y):

$$\mathbb{P}(Y_1 = 1) = \mathbb{P}(Y_0 = 1).$$

- ▶ Both X and Y cause depression (C).
- ▶ The common effect C is called a *collider* because of the colliding arrows on the path $X \rightarrow C \leftarrow Y$.
- ▶ Colliders block the flow of association along the path on which they lie: X and Y are independent because the only path between them is blocked by C .
- ▶ If we do condition on C , we induce an association between X and Y (via conditional dependence).

Selection bias

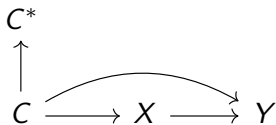


- ▶ A certain genetic haplotype (X) has no causal effect on the risk of neurodegenerative disease (Y):

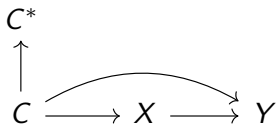
$$\mathbb{P}(Y_1 = 1) = \mathbb{P}(Y_0 = 1).$$

- ▶ Both X and Y cause depression (C).
- ▶ The common effect C is called a *collider* because of the colliding arrows on the path $X \rightarrow C \leftarrow Y$.
- ▶ Colliders block the flow of association along the path on which they lie: X and Y are independent because the only path between them is blocked by C .
- ▶ If we do condition on C , we induce an association between X and Y (via conditional dependence).
- ▶ This is known as *selection bias*.

Measurement bias

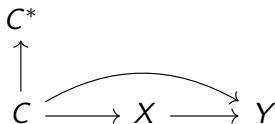


Measurement bias



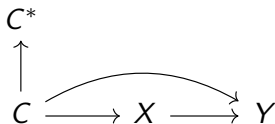
- Thus far, we have assumed that all variables are perfectly measured.

Measurement bias



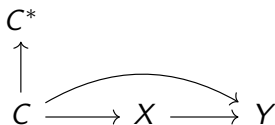
- ▶ Thus far, we have assumed that all variables are perfectly measured.
- ▶ *Measurement bias* occurs when the association between treatment and outcome is affected by the process by which the study data are measured.

Measurement bias



- ▶ Thus far, we have assumed that all variables are perfectly measured.
- ▶ *Measurement bias* occurs when the association between treatment and outcome is affected by the process by which the study data are measured.
- ▶ For instance, mismeasurement of confounders induce bias even if both treatment and outcome are perfectly measured.

Measurement bias



- ▶ Thus far, we have assumed that all variables are perfectly measured.
- ▶ *Measurement bias* occurs when the association between treatment and outcome is affected by the process by which the study data are measured.
- ▶ For instance, mismeasurement of confounders induce bias even if both treatment and outcome are perfectly measured.
- ▶ In general, the path $X \leftarrow C \rightarrow Y$ cannot be blocked by conditioning on mismeasured confounder C^* .

Sources of association

Four possible reasons why two variables may be associated:

Sources of association

Four possible reasons why two variables may be associated:

1. One causes the other.

Sources of association

Four possible reasons why two variables may be associated:

1. One causes the other.
2. They share common causes.

Sources of association

Four possible reasons why two variables may be associated:

1. One causes the other.
2. They share common causes.
3. They share a common effect and the analysis is restricted to a certain level of that common effect.

Sources of association

Four possible reasons why two variables may be associated:

1. One causes the other.
2. They share common causes.
3. They share a common effect and the analysis is restricted to a certain level of that common effect.
4. Random variability (disappears as $n \rightarrow \infty$).

Three forms of systematic bias

Three forms of systematic bias

- ▶ A study suffers from *systematic bias* when the data are insufficient to identify (i.e., compute) a causal effect even with an infinite sample size.

Three forms of systematic bias

- ▶ A study suffers from *systematic bias* when the data are insufficient to identify (i.e., compute) a causal effect even with an infinite sample size.
- ▶ Any structural association between treatment and outcome that does not arise from the causal effect of treatment on outcome.

Three forms of systematic bias

- ▶ A study suffers from *systematic bias* when the data are insufficient to identify (i.e., compute) a causal effect even with an infinite sample size.
- ▶ Any structural association between treatment and outcome that does not arise from the causal effect of treatment on outcome.
- ▶ Three forms of systematic bias:

Three forms of systematic bias

- ▶ A study suffers from *systematic bias* when the data are insufficient to identify (i.e., compute) a causal effect even with an infinite sample size.
- ▶ Any structural association between treatment and outcome that does not arise from the causal effect of treatment on outcome.
- ▶ Three forms of systematic bias:
 1. Confounding: not conditioning on common causes.

Three forms of systematic bias

- ▶ A study suffers from *systematic bias* when the data are insufficient to identify (i.e., compute) a causal effect even with an infinite sample size.
- ▶ Any structural association between treatment and outcome that does not arise from the causal effect of treatment on outcome.
- ▶ Three forms of systematic bias:
 1. Confounding: not conditioning on common causes.
 2. Selection: conditioning on common effects.

Three forms of systematic bias

- ▶ A study suffers from *systematic bias* when the data are insufficient to identify (i.e., compute) a causal effect even with an infinite sample size.
- ▶ Any structural association between treatment and outcome that does not arise from the causal effect of treatment on outcome.
- ▶ Three forms of systematic bias:
 1. Confounding: not conditioning on common causes.
 2. Selection: conditioning on common effects.
 3. Measurement: information error.

Three forms of systematic bias

- ▶ A study suffers from *systematic bias* when the data are insufficient to identify (i.e., compute) a causal effect even with an infinite sample size.
- ▶ Any structural association between treatment and outcome that does not arise from the causal effect of treatment on outcome.
- ▶ Three forms of systematic bias:
 1. Confounding: not conditioning on common causes.
 2. Selection: conditioning on common effects.
 3. Measurement: information error.
- ▶ These biases may arise in observational studies *and* in randomised experiments (how?).