CRG 17/18 Meeting 3

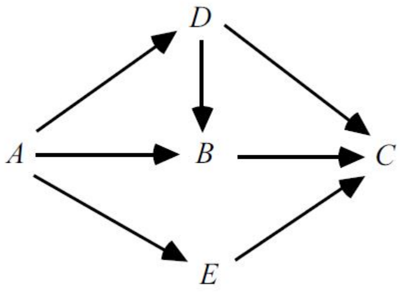
Discussant: Oisín Ryan

November 22, 2017

Todays readings

- Chapter 2, sections 2.3.2 to 2.6 (inclusive)
- Chapter 3, up to and including section 3.3 (p.19-29)

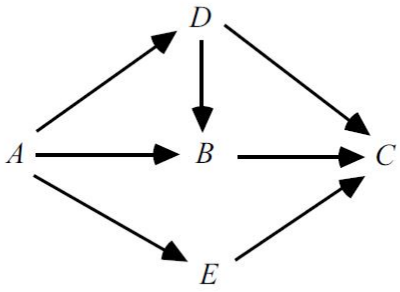
Identify that term!



Consider directed paths from A to C

- **Source** $=$

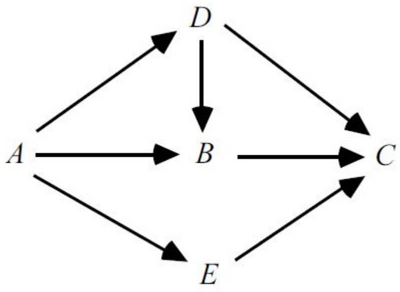
Identify that term!



Consider directed paths from A to C

- **Source** $= A$

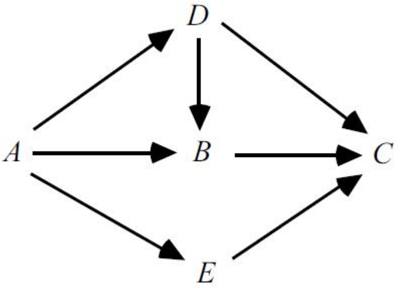
Identify that term!



Consider directed paths from A to C

- **Source** $= A$
- **Sink** $=$

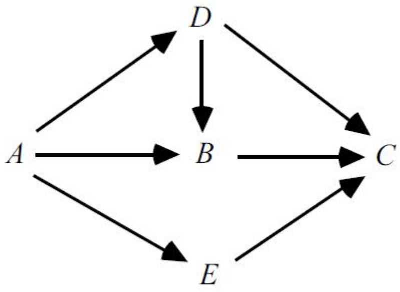
Identify that term!



Consider directed paths from A to C

- **Source** $= A$
- **Sink** $= C$

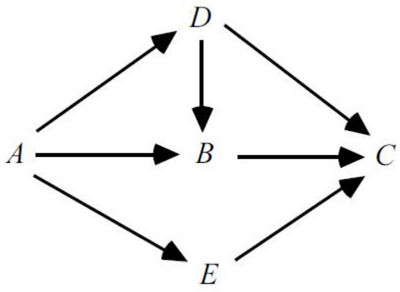
Identify that term!



Consider directed paths from A to C

- **Source** $= A$
- **Sink** $= C$
- **Trek** $=$

Identify that term!



Consider directed paths from A to C

- **Source** $= A$
- **Sink** $= C$
- **Trek** $= A \to D \to C$ and $A \to E \to C$

A trek between distinct vertices A and B is an unordered pair of directed paths between A and B that have the same source, and intersect only at the source. One of the paths in a trek may be an empty path

Minimality

- If G is a directed acyclic graph over $\mathbf{V}$ and P a probability distribution over $\mathbf{V}$, $< G, P >$ satisfies the Minimality Condition if and only if for every proper subgraph H of G with vertex set $\mathbf{V}$, $< H, P >$ does not satisfy the Markov condition
- Minimality **and** the Markov condition are met if all of the edges in G are necessary to fully describe the dependency structure in the joint density. I cannot build a simplified version of the graph which also satisfies the markov condition
- If a distribution $P(\boldsymbol{V}$ satisfies the Markov and Minimality conditions for a directed acyclic graph G, then G is a **minimal I-map** of P

2.3.2 Directed Independence Graphs (DIG)

- Almost equivalent way of representing conditional independence relations
- We can say that there is an *ordering* of vertices respected in the graph
  - Edges point from lower ordered to higher ordered vertices
- The DAG G is a DIG of $P(\boldsymbol{V})$ for an ordering $>$ of the vertices in G if an only if $A \rightarrow B$ occurs in G if A is conditionally dependent on B, conditioned on the set of vertices V such that $V \neq A$ and $V > B$.
- There should be an arrow from A to B only if they are dependent conditional on anything "downstream" of B
- Not equivalent when the probability distribution is not positive

2.3.3 Faithfulness

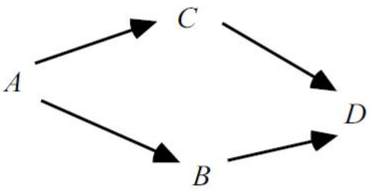


Figure 2.9

- Markov Con says nodes should be independent of their non-parents and non-descendants, given their parents
- A distribution may have other independence relations besides those given by Markov Con
- A might be independent of D - the paths trough C and B might exactly cancel one another out
- We assume this cannot be the case

2.3.3 Faithfulness

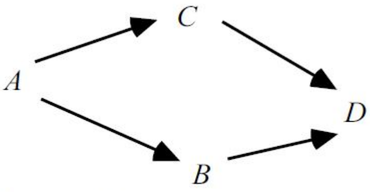


Figure 2.9

- If all and only the CI relations true in P are entailed by the Markov Con applied to G, P and G are **faithful**
- G is a perfect map of P
- P is a DAG-Isomorph of G

2.3.4 d-seperation

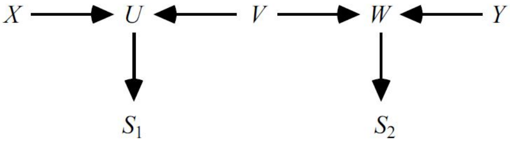


Figure 2.10

X and Y are d-seperated given $\boldsymbol{W}$ if and only if there is no undirected path U between X and Y such that:

1. Every collider on U has a descendant in $\boldsymbol{W}$
   AND
2. No other vertex on U is in $\boldsymbol{W}$

X and Y are d-connected if they are not d-seperated

2.3.4 d-seperation

Rule 1: X and Y are *d-connected* if there is an unblocked path between them

- A path is any sequence of edges, disregarding their direction
- A path is “unblocked” if it doesn't pass through a **collider**, e.g. $\rightarrow U \leftarrow$

2.3.4 d-seperation

Rule 1: X and Y are *d-connected* if there is an unblocked path between them

- A path is any sequence of edges, disregarding their direction
- A path is "unblocked" if it doesn't pass through a **collider**, e.g. $\rightarrow U \leftarrow$

Rule 2: X and Y are *d-connected* conditional on $\boldsymbol{Q}$ if there is a collider-free path between X and Y that traverses no member of $\boldsymbol{Q}$.

- If no such path exists, they are *d-seperated* by $\boldsymbol{Q}$

2.3.4 d-seperation

Rule 1: X and Y are *d-connected* if there is an unblocked path between them

- A path is any sequence of edges, disregarding their direction
- A path is "unblocked" if it doesn't pass through a **collider**, e.g. $\rightarrow U \leftarrow$

Rule 2: X and Y are *d-connected* conditional on $\boldsymbol{Q}$ if there is a collider-free path between X and Y that traverses no member of $\boldsymbol{Q}$.

- If no such path exists, they are *d-seperated* by $\boldsymbol{Q}$

Rule 3: If a collider is a member of the conditioning set $\boldsymbol{Q}$, **or has a descendant** in $\boldsymbol{Q}$, then it no longer blocks any path that traces this collider

2.3.4 d-seperation

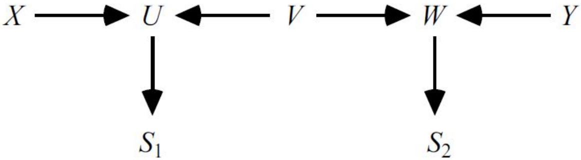


Figure 2.10

X and Y are d-seperated given the empty set

Rule 1: X and Y are d-connected if there is an unblocked path between them

2.3.4 d-seperation

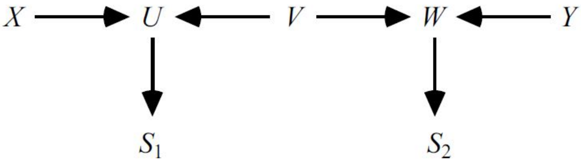


Figure 2.10

X and Y are d-connected given the set $\{S_1, S_2\}$

- Rule 2: X and Y are d-connected conditional on $\{S_1, S_2\}$ if there is a collider-free path between them that does not go through $\{S_1, S_2\}$
- Rule 3: If a collider is a member of $\{S_1, S_2\}$ or has a descendant in $\{S_1, S_2\}$, then it no longer blocks any path through it
  - U and W are colliders, but since we condition on their children, they no longer block any paths

2.3.4 d-seperation

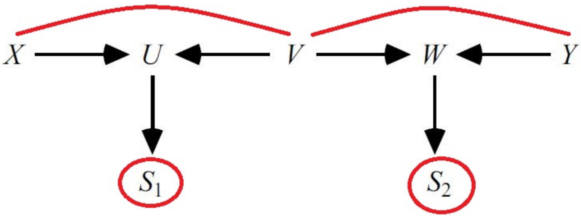


Figure 2.10

X and Y are d-connected given the set $\{S_1, S_2\}$

- Rule 2: X and Y are d-connected conditional on $\{S_1, S_2\}$ if there is a collider-free path between them that does not go through $\{S_1, S_2\}$
- Rule 3: If a collider is a member of $\{S_1, S_2\}$ or has a descendant in $\{S_1, S_2\}$, then it no longer blocks any path through it
  - U and W are colliders, but since we condition on their children, they no longer block any paths

2.3.4 d-seperation

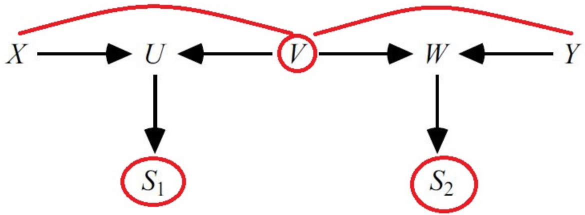


Figure 2.10

X and Y are d-seperated given the set $\{S_1, S_2, V\}$

- Rule 2: X and Y are d-connected conditional on $\{S_1, S_2, V\}$ if there is a collider-free path between them that does not go through $\{S_1, S_2, V\}$
- All newly-open paths between X and Y (through U and W) still must go through V. So conditioning additionally on V blocks these

2.3.5 Linear Structures

We can say our DAG G is a linear representation of representing an expanded graph G' over a superset $\boldsymbol{V'}$ and distribution $P''(\boldsymbol{V})$

- Every endogenous variable in V, has a unique error variable in V'
- Each endogenous variable is a linear function of its parents in G'
- In $P''(\boldsymbol{V})$ the correlation between exogenous variables in G' is zero $P(\boldsymbol{V}$ is the marginal of $P''(\boldsymbol{V})$ over $\boldsymbol{V}$

If G linearly represents $P(\boldsymbol{V})$ we say $G, P(\boldsymbol{V})$ is a linear model with DAG G

2.4 Undirected Independence Graphs

- Another graphical representation of CI
- A and B are connected only if they are conditionally independent given $\setminus\{A, B\}$
- In general the UIG **is not** the same as an undirected version of a DAG, but will be a subgraph of this

2.5 Deterministic and Pseudoindeterministic Systems

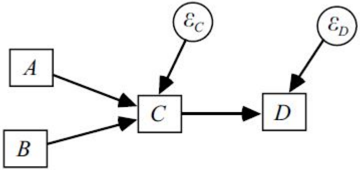


Figure 2.11

- Related to the SEM representation/idea of models
- If C is *uniquely* determined by A and ϵ_c it is deterministic
- If we do not have ϵ_c and ϵ_d, but we think of our observed variables as embedded in this deterministic graph, we can say the graph is pseudoindeterministic

Group Discussion: d-seperation

Two useful links for further reading on d-seperation

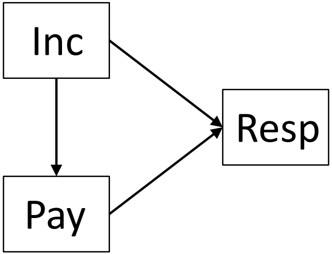
1. d-seperation without tears
   - A more didactical description of d-seperation rules with graph examples
2. Endogenous Selection Bias: The Problem of Conditioning on a Collider Variable
   - A nice sociology article giving lots of applied examples of colliders and the problem of conditioning on them

Group Discussion: Conditioning on a collider

- We tend to think about *conditioning* on a third variable in a regression sense; for example, we want to know the relationship between X and Y. In building a regression model predicting Y from X, we add C as a predictor, thus estimating the conditional relationship between X and Y, conditioned on C.
- More generally conditioning on a third variable entails estimating relationship between X and Y at a particular value of C. Let C be gender, X be intelligence and Y be income. Take it that we wanted to know the relationship between intelligence and income; we conduct a survey asking about income and intelligence, however we *sample only males*. Then, estimating the correlation between income and intelligence in our dataset would amount to estimating the relationship between intelligence and income *conditional on* gender being male.
- As such, if we wish to estimate the relationship between X and Y, and we *sample* based on a variable C, which is itself caused by X and Y, then we will introduce bias due to conditioning on a collider

Group Discussion: Example of collider conditioning

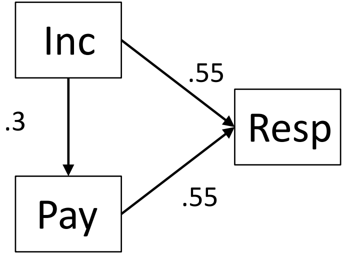
We can view an example of **nonresponse bias** (in our exmaple a MNAR mechanism) as a problem of conditioning on a collider



- Take it that we are interested in investigating the influence of Income (Inc) on Child-Support Payments (Pay) in divorced fathers
- Say we survey a sample of divorced fathers, but not all of them respond. Let Resp$= 0$ for non-responders, and Resp$= 1$ for responders.
- Suppose that in reality, fathers who are have a higher income, and fathers who pay more child support, are more likely to respond to the survey.
- In this case when we estimate the relationship between Inc and Pay in our sample, we will be conditioning on Resp$= 1$, introducing bias

Group Discussion: Example of collider conditioning

We can view (at least certain instances) of **nonresponse bias** as a problem of conditioning on a collider

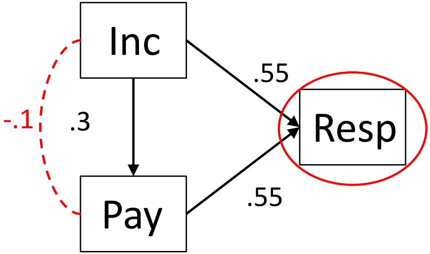


We can see this bias with a simple numerical example

- For simplicity lets assume linearity and continuous variables so we can use correlations
- Let's say that the marginal correlation between Income and Payment in the population (so the true relationship) is $r_{IP} = .3$. This is the relationship **without** conditioning on a collider
- Let's say we can express the effect of Inc on Resp with the partial correlation of Inc and Resp, conditional on Pay, $r_{IR.P} = .55$, and vice versa for Pay on Resp, $r_{PR.I} = .55$
- These numbers are assigned to their corresponding arrows in the DAG

Group Discussion: Example of collider conditioning

We can view (at least certain instances) of **nonresponse bias** as a problem of conditioning on a collider

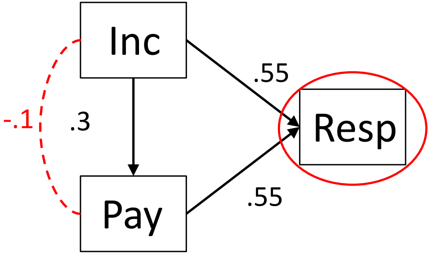


We can see this bias with a simple numerical example

- When we **condition** on Resp, as we would do in our sample, we find a small negative partial correlation between Inc and Pay, $r_{IP.R} = -.1$

Group Discussion: Example of collider conditioning

We can view (at least certain instances) of **nonresponse bias** as a problem of conditioning on a collider



Marginal Correlation Matrix

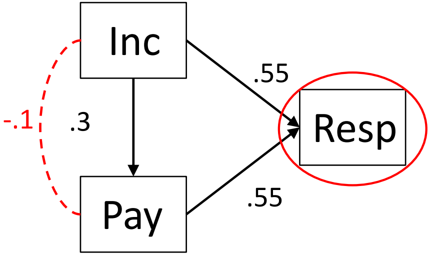
$$\begin{bmatrix} 1 & 0.3 & 0.6 \\ 0.3 & 1 & 0.6 \\ 0.6 & 0.6 & 1 \end{bmatrix}$$

Partial Correlation matrix

$$\begin{bmatrix} 1 & -.094 & 0.55 \\ -.094 & 1 & 0.55 \\ 0.55 & 0.55 & 1 \end{bmatrix}$$

Group Discussion: Example of collider conditioning

We can view (at least certain instances) of **nonresponse bias** as a problem of conditioning on a collider




```
library("corpcor")
c12<-.3
c13<-.6
c23<-.6
m<-matrix(c(1,c12,c13,
            c12,1,c23,
            c13,c23,1),3,3,byrow=TRUE)
cor2pcor(m)
```