CRG 17/18 Meeting 6

Discussant: Oisín Ryan

February 22, 2018

Topics

- d-seperation (the return)
- Finding faithful DAGs from data
- The Manipulation Theorem

d-seperation: who cares?

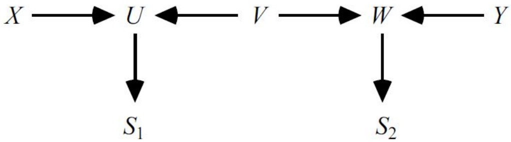


Figure 2.10

- Markov Con: nodes should be independent of their **non-parents** and **non-descendants**, conditional on their **parents**
  - $W \perp\!\!\!\perp \{U, X, S_1\} | \{V, Y\}$
- MC is not sufficient to describe **all** CIs/CDs
  - $X \perp\!\!\!\perp Y$
  - $X \not\perp\!\!\!\perp Y | \{S_1, S_2\}$
  - $X \perp\!\!\!\perp Y | \{S_1, S_2, V\}$
- d-seperated $=$ (conditionally) independent

d-seperation rules

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

d-seperation in action

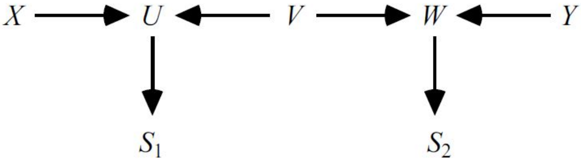


Figure 2.10

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

d-seperation in action

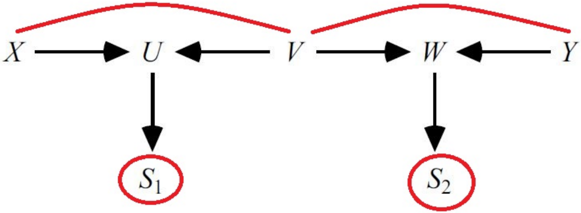


Figure 2.10

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

d-seperation in action

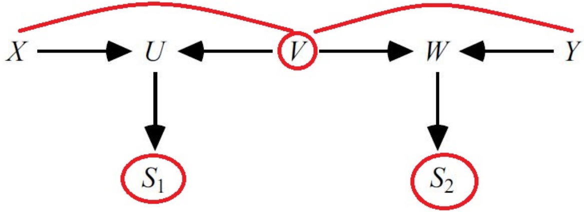


Figure 2.10

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

d-seperation and faithfulness

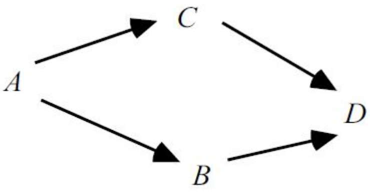


Figure 2.9

1. If all and only the CI relations true in P are entailed by the Markov Con applied to G, $P(\boldsymbol{V})$ and G are **faithful**
2. $P(\boldsymbol{V})$ is faithful to graph G if and only if, for all disjoint sets $\boldsymbol{X}$, $\boldsymbol{Y}$, $\boldsymbol{Z}$, the variables in $\boldsymbol{X}$ and $\boldsymbol{Y}$ are independent conditional on $\boldsymbol{Z}$ only if they are d-seperated given $\boldsymbol{Z}$
   - The only CIs that are present between variables follow the d-seperation rules

Towards a DAG-finding algorithm

Theorem 3.4: $P(\boldsymbol{V}$ is faithful to a DAG G if and only if

1. for all vertices X, Y, of G, X and Y are adjacent if and only if X and Y are dependent conditional on **every set of vertices** of G that does not include X and Y
2. for all vertices X, Y, Z, such that X is adjacent to Y, Y is adjacent to Z, and X and Z are **not** adjacent, we can orientate $X - Y - Z$ as $X \rightarrow Y \leftarrow Z$ only if X and Z are dependent conditional on every set containing Y but not X or Z

faithful_finder.R

The Manipulation Theorem

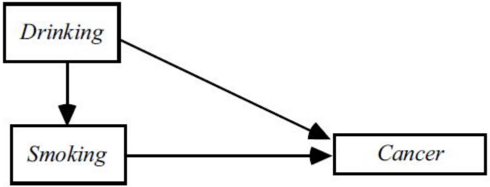
- If the causal structure is known, and the direct effects of the manipulation are known, the the joint distribution under the manipulation an be estimated from the **unmanipulated population**

The Manipulation Theorem

- If the causal structure is known, and the direct effects of the manipulation are known, the the joint distribution under the manipulation an be estimated from the **unmanipulated population**
- If we can figure out the causal structure from observational data, and we can define an intervention well, we can find out what the effect of that intervention on the whole system would be

Manipulation in action: Step 1

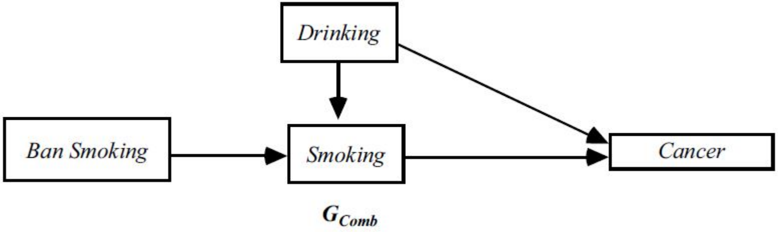
Unmanipulated (observational/natural) causal system



G_{Unman}

Step 2

Represent the intervention we want to know about as a variable in an expanded DAG

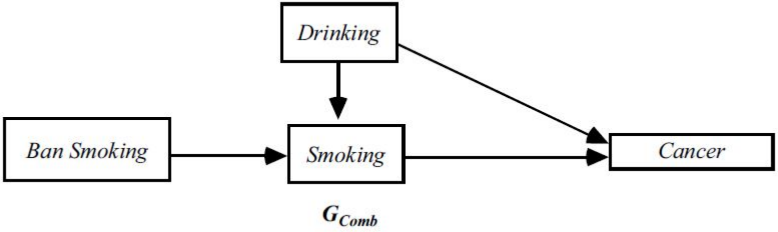


What are our theoretical assumptions here?

1. Banning smoking will not affect drinking or cancer directly

Step 2

Represent the intervention we want to know about as a variable in an expanded DAG

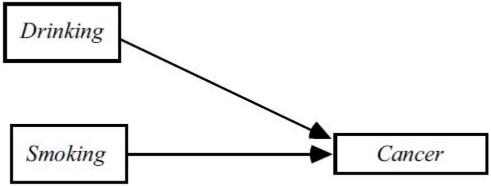


What are our theoretical assumptions here?

1. Banning smoking will not affect drinking or cancer directly
2. Banning smoking will be completely effective

Step 3: the manipulated graph

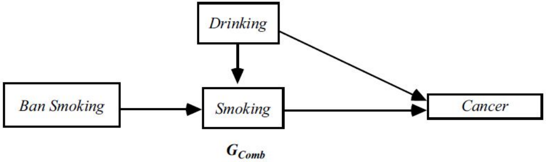
What would the graph look like in the hypothetitcal manipulated population?



G_{Man}

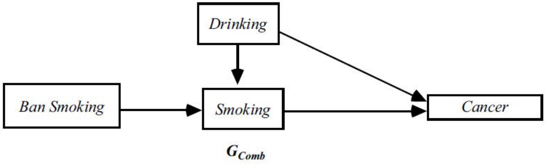
- Because the ban is completely effective, drinking doesn't influence smoking anymore
- Note: this looks a lot like a **mediation** scenario to me!

Making inferences about the manipulation



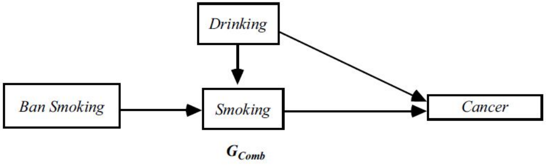
- Let $\boldsymbol{X} = \{Drinking, Smoking, Cancer\}$
- We only **observe** the population in which there is no smoking ban, $P(\boldsymbol{X}|BS = 0) = P_{unman}(\boldsymbol{X})$
- We want to know about $P(\boldsymbol{X}|BS = 1) = P_{man}(\boldsymbol{X})$.
- In $P_{man}(\boldsymbol{X})$, $Smoking = 0$ for everyone

Making inferences about the manipulation



- Query: Is $P_{unman}(C|S = 0)$ the same as $P_{man}(C|S = 0)$?
- Is *BS* d-seperated from *Cancer* conditional on *Smoking*? **No**

Making inferences about the manipulation

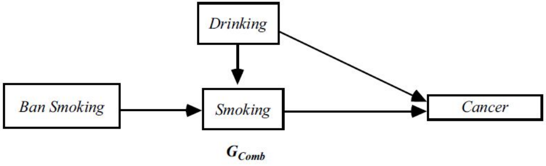


- Query: Is $P_{unman}(C|S=0)$ the same as $P_{man}(C|S=0)$?
- Is *BS* d-seperated from *Cancer* conditional on *Smoking*? **No**

Why?

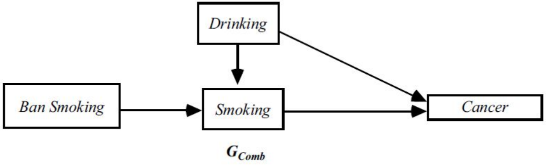
- *Smoking* is a collider - conditioning on it induces a path from $BS - Drinking \rightarrow Cancer$
- Say drinking positively predicts smoking and cancer. In the unmanipulated population, those who have $Smoking = 0$ also have low values for *Drinking*
- In the manipulated population, everyone has $Smoking = 0$, including those who drink a lot

Making inferences about the manipulation



- Query: Is $P_{unman}(C|S=0,D)$ the same as $P_{man}(Cancer|S=0,D)$?
- Is *BS* d-seperated from *Cancer* conditional on *Smoking* and *Drinking*? **Yes**

Making inferences about the manipulation

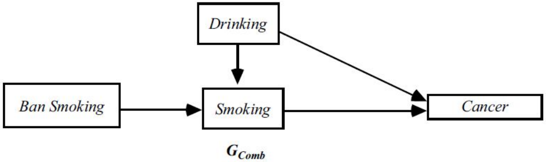


- Query: Is $P_{unman}(C|S = 0, D)$ the same as $P_{man}(Cancer|S = 0, D)$?
- Is *BS* d-seperated from *Cancer* conditional on *Smoking* and *Drinking*? **Yes**

Why?

- *Smoking* is a collider - conditioning on it induces a path from $BS - Drinking \rightarrow Cancer$
- Conditioning on *Drinking* blocks this path

Making inferences about the manipulation



- Query: Is $P_{unman}(C|S = 0, D)$ the same as $P_{man}(Cancer|S = 0, D)$?
- Is *BS* d-seperated from *Cancer* conditional on *Smoking* and *Drinking*? **Yes**

My interpretation

- We observe in our population some people who drink a little or a lot, and happen not to smoke by choice
- The distribution of cancer as a function of drinking, will be the same whether the population chooses not to smoke or whether they are forced not to smoke
- From this, we can infer the distribution of cancer as a function of drinking if everyone was forced not to smoke

Manipulation Theorem

- **Manipulation**: if $\boldsymbol{W}$ is exogenous to $\boldsymbol{V}$ (i.e. $\boldsymbol{V} \nrightarrow W$ then changing the value of $\boldsymbol{W}$ from $\boldsymbol{w_1}$ to $\boldsymbol{w_2}$ is a **manipulation** wrt to $\boldsymbol{V}$ if and only if $P(\boldsymbol{V}|\boldsymbol{W} = \boldsymbol{w_1}) \neq P(\boldsymbol{V}|\boldsymbol{W} = \boldsymbol{w_2})$
- **Manipulated (W)**: the children of $\boldsymbol{W}$ that are also in $\boldsymbol{V}$
- G_{unman} is a subgraph of G_{comb}. G_{man} is a subgraph of G_{unman}. Exactly what subgraph depends on the intervention (deleting edges not necessary!).
- If
  1. We know the unmanipulated joint density
  2. The unmanipulated joint density P_{unman} can be factorised according to the Markov Con
  3. We know the **direct effects** of the maniplation on **Manipulated(W)** (e.g. banning smoking makes all $Smoking = 0$)

  Then we can get the joint density under the manipulation by taking the unmanipulated density, and for every variable in **Manipulated(W)**, replace its density with its density under the manipulation.

Manipulation and discovery: open questions

- Here we have the DAG in the observational setting, and we know how an experimental setting would change it
  - From this, we can infer the effect of an experimental manipulation **without doing it**
  - You can always test this by doing the experiment!
- What about if we don't know exactly what the manipulation does?
  - If we had data from a) unmanipulated population and b) manipulated population
  - By fitting and comparing the DAGs in both populations, can we infer the full DAG, G_comb?
  - Might be useful for instance in micro-randonmised behavioural interventions