

Identifying Direct and Indirect effects

S.Geneletti

Imperial College Department of Epidemiology and Public Health

25/01/2008

Outline

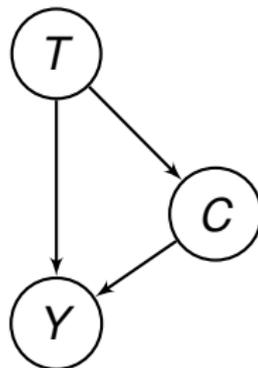
- Basics
- DAGs and intervention
- Examples
- Causal?
- Regimes
 - Heuristic
 - Formal
- Definition of effects
- Identification of effects
- Extensions
- Conclusions

Conditional Independence

$$p(A, C|B) = p(A|B) \times p(C|B) \Rightarrow A \perp\!\!\!\perp C|B$$

DAGs

- Directed Acyclic Graphs
- Encode conditional independence
- More than one can encode same CI - like intuitive DAG for direct and indirect effects.



Basics continued

Notation

- T treatment: 0,1
- F_T treatment assignment: 0,1, \emptyset
- C intermediate variable: 0,1
- Y outcome/response variable: y

Basics continued

Notation

- T treatment: 0,1
- F_T treatment assignment: 0,1, \emptyset
- C intermediate variable: 0,1
- Y outcome/response variable: y

F_T : the intervention

- F_T is such that
- $p(T = t|X, F_T = t) = 1$ when $t \neq \emptyset$ and
- $p(T = t|X, F_T = \emptyset) = p(T = t|X)$
i.e. T is drawn from its “natural” distribution
- F_T is an intervention variable, treatment assignment indicator, regime indicator.
- DAGs with intervention variables = **Augmented DAGs**

Causality and ADs

Causes as interventions

- In the framework proposed here, **causes are** equated to **intervention** e.g. randomization
- i.e when $F_T \neq \emptyset$ then we are talking cause!

Causality and ADs

Causes as interventions

- In the framework proposed here, **causes are** equated to **intervention** e.g. randomization
- i.e when $F_T \neq \emptyset$ then we are talking cause!



same as

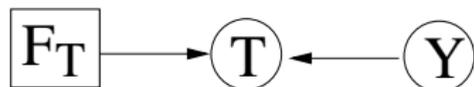


w.r.t. conditional indep

- LHS - T and Y are dependent



not the same as



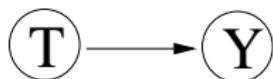
w.r.t. conditional indep

BUT

Causality and ADs

Causes as interventions

- In the framework proposed here, **causes are** equated to **intervention** e.g. randomization
- i.e when $F_T \neq \emptyset$ then we are talking cause!



same as

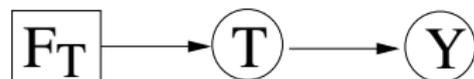


w.r.t. conditional indep

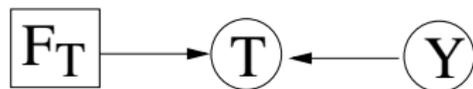
- LHS - T and Y are dependent

- RHS - **top** says $Y \perp\!\!\!\perp F_T | T$ so **T causes Y**

BUT



not the same as



w.r.t. conditional indep

Causality and ADs

Causes as interventions

- In the framework proposed here, **causes are** equated to **intervention** e.g. randomization
- i.e when $F_T \neq \emptyset$ then we are talking cause!



same as

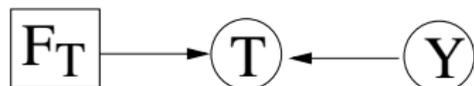
BUT



not the same as



w.r.t. conditional indep

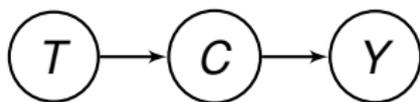


w.r.t. conditional indep

- LHS - T and Y are dependent
- RHS - **top** says $Y \perp\!\!\!\perp F_T | T$ so **T causes Y**
- RHS - **bottom** says $Y \perp\!\!\!\perp F_T$ so **T does not cause Y**

DAGs for direct and indirect effects

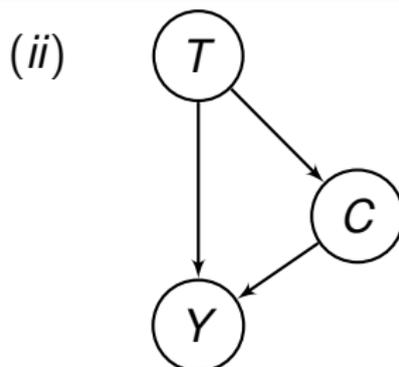
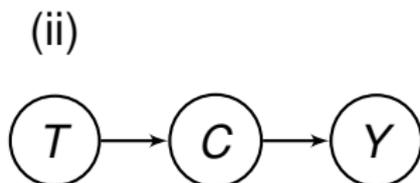
(ii)



The DAGs encode conditional independence

- (i) says $T \perp\!\!\!\perp Y \mid C$ i.e. if we know C , T cannot tell us anything more about Y

DAGs for direct and indirect effects



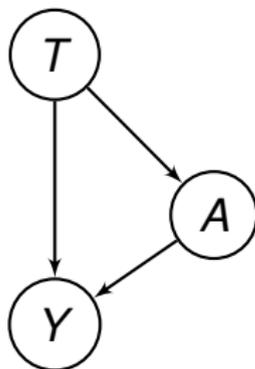
The DAGs encode conditional independence

- (i) says $T \perp\!\!\!\perp Y | C$ i.e. if we know C , T cannot tell us anything more about Y
- (ii) says nothing ... but that silence speaks volumes! We can extract more out of it if we add more variables - see how later

Example 1: side-effects

Story

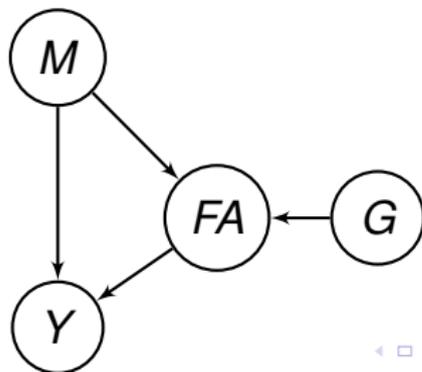
- A drug has strong head-aches as a side-effect
- Patients take aspirin to alleviate this
- It is thought that the aspirin might affect the outcome
- Drug company wants to know what are direct/indirect effects?



Example 2: meat and colon cancer

Story

- Meat consumption thought to be associated to colon cancer
- Also associated to colon cancer are genes that are responsible for breaking down fatty acids (CD36) via complex chemistry involving proteins etc
- Can we change/emulate gene behaviour (injections?)
- for those with genes that predispose them to cancer?



Causal?

Remember Causes as interventions

- i.e. sex, age, race are not causes - we cannot change them!
- treatments and other things we can intervene on (education, weight) are potential causes
- genes can be causes if we can conceive of intervening on them or their function

Causal?

Remember Causes as interventions

- i.e. sex, age, race are not causes - we cannot change them!
- treatments and other things we can intervene on (education, weight) are potential causes
- genes can be causes if we can conceive of intervening on them or their function

C can be intervened upon

- For there to be a direct/indirect effect, it must be possible to intervene on C
- Ex 1: aspirin intake of patients can in principle be controlled
- Ex 2: gene function changed or emulated

Identifiable?

Problem

- OK C can be intervened upon **in principle**....
- but **in practice** often not possible:
- cannot deny a patient with head-ache pain relief
- cannot intervene on a human knowledge not solid enough

Identifiable?

Problem

- OK C can be intervened upon **in principle**....
- but **in practice** often not possible:
- cannot deny a patient with head-ache pain relief
- cannot intervene on a human knowledge not solid enough

Solution

- It all depends on the **conditional independences!**
- Using these and some cunning, we can identify the direct and indirect effects

Identifiable?

Problem

- OK C can be intervened upon **in principle**....
- but **in practice** often not possible:
- cannot deny a patient with head-ache pain relief
- cannot intervene on a human knowledge not solid enough

Solution

- It all depends on the **conditional independences!**
- Using these and some cunning, we can identify the direct and indirect effects
- and often estimate them using data that are purely observational!

Direct & indirect effects: a heuristic overview

The type of direct and indirect effect you get depends on **what you do to C**

Ways of manipulating C

- Fix C

Direct & indirect effects: a heuristic overview

The type of direct and indirect effect you get depends on **what you do to C**

Ways of manipulating C

- Fix C
- Draw C from $p(C|T)$

Direct & indirect effects: a heuristic overview

The type of direct and indirect effect you get depends on **what you do to C**

Ways of manipulating C

- Fix C
- Draw C from $p(C|T)$
- Draw C from an appropriate distribution

Each one of these manipulations can be described by a **regime**.

Fixing regime: heuristic

Ex 1: imaginary experiment I

- Imagine you can fix aspirin intake e.g. everyone has to take no aspirins

Fixing regime: heuristic

Ex 1: imaginary experiment I

- Imagine you can fix aspirin intake e.g. everyone has to take no aspirins
- then can estimate the direct effect of treatment for no aspirins.

Fixing regime: heuristic

Ex 1: imaginary experiment I

- Imagine you can fix aspirin intake e.g. everyone has to take no aspirins
- then can estimate the direct effect of treatment for no aspirins.
- Imagine you can randomise aspirins as well as treatment

Fixing regime: heuristic

Ex 1: imaginary experiment I

- Imagine you can fix aspirin intake e.g. everyone has to take no aspirins
- then can estimate the direct effect of treatment for no aspirins.
- Imagine you can randomise aspirins as well as treatment
- then can estimate the direct effect of treatment for 2 levels of aspirin and some kind of indirect effect by comparing the two direct effects.

Fixing regime: heuristic

Ex 1: imaginary experiment I

- Imagine you can fix aspirin intake e.g. everyone has to take no aspirins
- then can estimate the direct effect of treatment for no aspirins.
- Imagine you can randomise aspirins as well as treatment
- then can estimate the direct effect of treatment for 2 levels of aspirin and some kind of indirect effect by comparing the two direct effects.

Problems

- Not ethical!
- Not realistic – people “out there” in the population do not behave like this.

Drawing C from $p(C|T)$ regime: heuristic

Ex 1: imaginary experiment II

- One random group of patients are given the control $F_T = 0$ and we record their “natural” aspirin intake $p(C|F_T = 0)$

Drawing C from $p(C|T)$ regime: heuristic

Ex 1: imaginary experiment II

- One random group of patients are given the control $F_T = 0$ and we record their “natural” aspirin intake $p(C|F_T = 0)$
- next we take the rest of the patients and
 - If a patient gets $F_T = 1$, they are administered aspirin according to $p(C|F_T = 0)$

Drawing C from $p(C|T)$ regime: heuristic

Ex 1: imaginary experiment II

- One random group of patients are given the control $F_T = 0$ and we record their “natural” aspirin intake $p(C|F_T = 0)$
- next we take the rest of the patients and
 - If a patient gets $F_T = 1$, they are administered aspirin according to $p(C|F_T = 0)$
 - If a patient gets $F_T = 0$, we leave them alone, and assume that they take aspirin according to $p(C|F_T = 0)$

Drawing C from $p(C|T)$ regime: heuristic

Ex 1: imaginary experiment II

- One random group of patients are given the control $F_T = 0$ and we record their “natural” aspirin intake $p(C|F_T = 0)$
- next we take the rest of the patients and
 - If a patient gets $F_T = 1$, they are administered aspirin according to $p(C|F_T = 0)$
 - If a patient gets $F_T = 0$, we leave them alone, and assume that they take aspirin according to $p(C|F_T = 0)$
- This way we can compare direct effects fixing not the aspirin intake, rather its distribution to that of the control treatment

Problem

- Not ethical!

Drawing C from $p(C^*)$ regime: heuristic

Ex 1: imaginary experiment III

- We cannot estimate $p(C|T)$ - lack of funds!
- So we pick a suitable distribution $p(D)$ such that domain of C is same as D .
 - If a patient gets $F_T = 1$, they are administered aspirin according to $p(D)$
 - If a patient gets $F_T = 0$, they are also administered aspirin according to $p(D)$
- In this case, the comparison is not as easily interpretable, but we can still get a handle on the direct and indirect effects.

Drawing C from $p(C^*)$ regime: heuristic

Ex 1: imaginary experiment III

- We cannot estimate $p(C|T)$ - lack of funds!
- So we pick a suitable distribution $p(D)$ such that domain of C is same as D .
 - If a patient gets $F_T = 1$, they are administered aspirin according to $p(D)$
 - If a patient gets $F_T = 0$, they are also administered aspirin according to $p(D)$
- In this case, the comparison is not as easily interpretable, but we can still get a handle on the direct and indirect effects.

Standardisation

- The 2nd and 3rd regimes are forms of direct standardisation.

Formal definition of Regimes

Manipulating C

- Define M_C the variable of manipulations of C
- M_C has 4 regimes:
 - 1 $\mathbf{M}_C = \mathbf{c}$: C is left alone and comes from
$$p(C|T = t, F_T = \emptyset, M_C = \emptyset) = P_t$$

Formal definition of Regimes

Manipulating C

- Define M_C the variable of manipulations of C
- M_C has 4 regimes:
 - 1 $\mathbf{M}_C = \mathbf{c}$: C is left alone and comes from $p(C|T = t, F_T = \emptyset, M_C = \emptyset) = P_t$
 - 2 $\mathbf{M}_C = \mathbf{c}$: C is set to c with no uncertainty, distribution of C is δ_c

Formal definition of Regimes

Manipulating C

- Define M_C the variable of manipulations of C
- M_C has 4 regimes:
 - 1 $\mathbf{M}_C = \mathbf{c}$: C is left alone and comes from $p(C|T = t, F_T = \emptyset, M_C = \emptyset) = P_t$
 - 2 $\mathbf{M}_C = \mathbf{c}$: C is set to c with no uncertainty, distribution of C is δ_c
 - 3 $\mathbf{M}_C = \mathbf{t}^*$: C is drawn from $p(C|F_T = \emptyset, M_C = t^*) = P_t$

Formal definition of Regimes

Manipulating C

- Define M_C the variable of manipulations of C
- M_C has 4 regimes:
 - 1 $\mathbf{M}_C = \mathbf{c}$: C is left alone and comes from $p(C|T = t, F_T = \emptyset, M_C = \emptyset) = P_t$
 - 2 $\mathbf{M}_C = \mathbf{c}$: C is set to c with no uncertainty, distribution of C is δ_c
 - 3 $\mathbf{M}_C = \mathbf{t}^*$: C is drawn from $p(C|F_T = \emptyset, M_C = t^*) = P_t$
 - 4 $\mathbf{M}_C = \mathbf{D}$: C is drawn from $p(D) = P_D$
- Note that Regime 2 is a special case of regime 3 which in turn is a special case of regime 4. Regime 1 is also a special case of 4.

Formal definition of Regimes continued

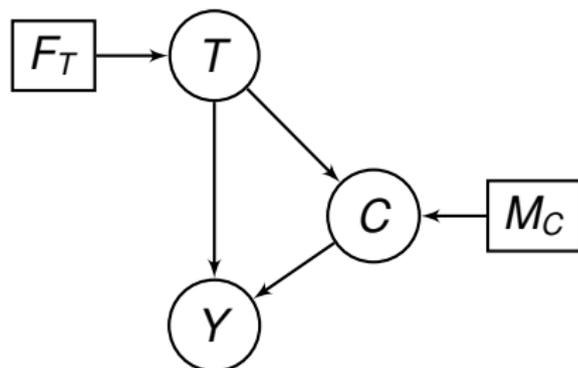
Conditional independences

$$(Y, C) \perp\!\!\!\perp F_T | T; \quad (1)$$

$$Y \perp\!\!\!\perp M_C | (C, T); \quad (2)$$

$$C \perp\!\!\!\perp (F_T, T) | M_C \neq \emptyset. \quad (3)$$

Equation 3 only for regimes 2-4



Some points

- The framework of these regimes is determined by M_C as defined

Some points

- The framework of these regimes is determined by M_C as defined
- Adding other variables might require a redefinition of M_C

Some points

- The framework of these regimes is determined by M_C as defined
- Adding other variables might require a redefinition of M_C
- As there are only 3 variables we consider here, we call this the **3DI framework**

Some points

- The framework of these regimes is determined by M_C as defined
- Adding other variables might require a redefinition of M_C
- As there are only 3 variables we consider here, we call this the **3DI framework**
- Extensions are quite straightforward, might show some later

Some simplifying notation

$$E(Y|F_T = a, M_C = b) = E(a, b)$$

Some simplifying notation

$$E(Y|F_T = a, M_C = b) = E(a, b)$$

Some simplifying notation

$$\begin{aligned} E(Y|F_T = a, M_C = b) &= E(a, b) \\ E(Y|F_T = a, M_C = b, C = c) &= E(a, b, c) \\ Pr(C = c|M_C = b) &= p(c|b) \\ Pr(C = c|F_T = a, M_C = b) &= p(c|a, b) \end{aligned}$$

Definition of direct & indirect effects

DE for C set at c - Set effect (SE_c)

The direct effect of $F_T = t$ with respect to baseline $F_T = t^*$ on response Y for C set at c is given by

$$\begin{aligned} E(Y|F_T = t, M_C = c) - E(Y|F_T = t^*, M_C = c) & \quad (4) \\ & = E(t, c) - E(t^*, c). \end{aligned}$$

Definition of direct & indirect effects

DE for C set at c - Set effect (SE_c)

The direct effect of $F_T = t$ with respect to baseline $F_T = t^*$ on response Y for C set at c is given by

$$\begin{aligned} E(Y|F_T = t, M_C = c) - E(Y|F_T = t^*, M_C = c) & \quad (4) \\ & = E(t, c) - E(t^*, c). \end{aligned}$$

DE for C drawn from P_{t^*} : Generated effect (GDE_{t^*})

The direct effect of $F_T = t$ with respect to the baseline $F_T = t^*$ on response for C generated from P_{t^*} the distribution of C conditional on $F_T = t^*$ the baseline treatment is given by

$$E(t, t^*) - E(t^*, t^*)$$

Definition of direct & indirect effects

DE for C set at c - Set effect (SE_c)

The direct effect of $F_T = t$ with respect to baseline $F_T = t^*$ on response Y for C set at c is given by

$$\begin{aligned} E(Y|F_T = t, M_C = c) - E(Y|F_T = t^*, M_C = c) & \quad (4) \\ & = E(t, c) - E(t^*, c). \end{aligned}$$

DE for C drawn from P_{t^*} : Generated effect (GDE_{t^*})

The direct effect of $F_T = t$ with respect to the baseline $F_T = t^*$ on response for C generated from P_{t^*} the distribution of C conditional on $F_T = t^*$ the baseline treatment is given by

$$E(t, t^*) - E(t^*, t^*) \equiv E(t, t^*) - E(t^*, \emptyset). \quad (5)$$

Definition of direct & indirect effects cont

DE for C drawn from P_D : Generated effect (GDE_D)

The direct effect of $F_T = t$ with respect to the baseline $F_T = t^*$ on response for C drawn from a specified distribution over D is given by:

$$E(Y|t, D) - E(t^*, D) \quad (6)$$

Definition of direct & indirect effects cont

DE for C drawn from P_D : Generated effect (GDE_D)

The direct effect of $F_T = t$ with respect to the baseline $F_T = t^*$ on response for C drawn from a specified distribution over D is given by:

$$E(Y|t, D) - E(t^*, D) \quad (6)$$

Total effect: TE

The total effect of $F_T = t$ with respect to the baseline $F_T = t^*$ on response is given by

$$E(Y|F_T = t) - E(Y|F_T = t^*). \quad (7)$$

Definition of direct & indirect effects cont

DE for C drawn from P_D : Generated effect (GDE_D)

The direct effect of $F_T = t$ with respect to the baseline $F_T = t^*$ on response for C drawn from a specified distribution over D is given by:

$$E(Y|t, D) - E(t^*, D) \quad (6)$$

Total effect: TE

The total effect of $F_T = t$ with respect to the baseline $F_T = t^*$ on response is given by

$$E(Y|F_T = t) - E(Y|F_T = t^*). \quad (7)$$

IDE for C drawn from P ($P \in \{\delta_C, P_{t^*}, P_D\}$)

$$IDE = DE - TE$$

Identifying effects in 3DI

- The TE is always identifiable, even when $F_T = \emptyset$ if there are no confounders
- If we can identify the DE then it follows we can identify the IDE, so focus is on DE
- We look first at scenario where there are no confounders and $F_T \neq \emptyset$
- Then consider what happens if there are confounders i.e. when

$$Y \perp\!\!\!\perp M_C | (C, T)$$

does not hold

- AND $F_T = \emptyset$.

Identifying GDE_{t^*}

$$GDE_{t^*} = \underbrace{E(t, t^*)}_{(a)} - \overbrace{E(t^*, \emptyset)}^{(b)}$$

Identifying GDE_{t^*}

$$GDE_{t^*} = \underbrace{E(t, t^*)}_{(a)} - \overbrace{E(t^*, \emptyset)}^{(b)}$$

(b) is easy = $E(Y|T = t^*, F_T = \emptyset, M_C = \emptyset)$ by (1)

(a) =

$$= \sum_c E(t, t^*, c) \times p(c|t, t^*) \quad (8)$$

Identifying GDE_{t^*}

$$GDE_{t^*} = \underbrace{E(t, t^*)}_{(a)} - \overbrace{E(t^*, \emptyset)}^{(b)}$$

(b) is easy = $E(Y|T = t^*, F_T = \emptyset, M_C = \emptyset)$ by (1)

(a) =

$$= \sum_c E(t, t^*, c) \times p(c|t, t^*) \quad (8)$$

$$= \sum_c E(t, \emptyset, c) \times p(c|t^*) \quad (9)$$

■ from (8)-(9) by (1)

Identifying GDE_{t^*}

$$GDE_{t^*} = \underbrace{E(t, t^*)}_{(a)} - \overbrace{E(t^*, \emptyset)}^{(b)}$$

(b) is easy = $E(Y|T = t^*, F_T = \emptyset, M_C = \emptyset)$ by (1)

(a) =

$$= \sum_c E(t, t^*, c) \times p(c|t, t^*) \quad (8)$$

$$= \sum_c E(t, \emptyset, c) \times p(c|t^*) \quad (9)$$

$$= \sum_c E(Y|T = t, F_T = \emptyset, M_C = \emptyset, C = c) \\ \times p(C = c|T = t^*, M_C = \emptyset) \quad (10)$$

■ from (8)-(9) by (1)

■ from (9)-(10) by (2)

Identifying continued

GDE_D is identified in same way as GDE_{t^*} with P_D replacing P_{t^*}

Pros

- So it is possible to identify and hence estimate the 3DI effects from data that are purely observational for **both** T and C as F_T and M_C are both \emptyset .

Identifying continued

GDE_D is identified in same way as GDE_{t^*} with P_D replacing P_{t^*}

Pros

- So it is possible to identify and hence estimate the 3DI effects from data that are purely observational for **both** T and C as F_T and M_C are both \emptyset .
- Simple formulae, easy to understand

Identifying continued

GDE_D is identified in same way as GDE_{t^*} with P_D replacing P_{t^*}

Pros

- So it is possible to identify and hence estimate the 3DI effects from data that are purely observational for **both** T and C as F_T and M_C are both \emptyset .
- Simple formulae, easy to understand - once you get the hang of it

Identifying continued

GDE_D is identified in same way as GDE_{t^*} with P_D replacing P_{t^*}

Pros

- So it is possible to identify and hence estimate the 3DI effects from data that are purely observational for **both** T and C as F_T and M_C are both \emptyset .
- Simple formulae, easy to understand - once you get the hang of it

Cons

- No counfounder assumption is unlikely to hold in practice

Identifying with confounders

Confounder between T and C or T and Y

- This is only really an issue if $F_T = \emptyset$

Identifying with confounders

Confounder between T and C or T and Y

- This is only really an issue if $F_T = \emptyset$
- In this situation it is necessary to observe the confounders - if they are unobserved, we're in trouble

Identifying with confounders

Confounder between T and C or T and Y

- This is only really an issue if $F_T = \emptyset$
- In this situation it is necessary to observe the confounders - if they are unobserved, we're in trouble
- If observed, we treat them as usual - condition and sum over

Identifying with confounders

Confounder between T and C or T and Y

- This is only really an issue if $F_T = \emptyset$
- In this situation it is necessary to observe the confounders - if they are unobserved, we're in trouble
- If observed, we treat them as usual - condition and sum over

Confounder between C and Y

- This requires a **redefinition of M_C** as the distribution of C now depends on this confounder - call it W as well as T .

W confounder between C and Y

W is a nuisance

- Ex 1: Two doctors administer the treatment.

W confounder between C and Y

W is a nuisance

- Ex 1: Two doctors administer the treatment.
- One knows that the side-effect is headache, tells patients not to take

W confounder between C and Y

W is a nuisance

- Ex 1: Two doctors administer the treatment.
- One knows that the side-effect is headache, tells patients not to take
- The other knows nothing so his patients are ignorant

W confounder between C and Y

W is a nuisance

- Ex 1: Two doctors administer the treatment.
- One knows that the side-effect is headache, tells patients not to take
- The other knows nothing so his patients are ignorant
- In this situation we want to get rid of W

W is relevant

- Ex 1: W is sex: women suffer more side-effect than men.

W confounder between C and Y

W is a nuisance

- Ex 1: Two doctors administer the treatment.
- One knows that the side-effect is headache, tells patients not to take
- The other knows nothing so his patients are ignorant
- In this situation we want to get rid of W

W is relevant

- Ex 1: W is sex: women suffer more side-effect than men.
- We might want to intervene on W somehow

W confounder between C and Y

W is a nuisance

- Ex 1: Two doctors administer the treatment.
- One knows that the side-effect is headache, tells patients not to take
- The other knows nothing so his patients are ignorant
- In this situation we want to get rid of W

W is relevant

- Ex 1: W is sex: women suffer more side-effect than men.
- We might want to intervene on W somehow

The two roles of W need to be treated in slightly different ways.

W as sex

Just consider this - the other case in the paper

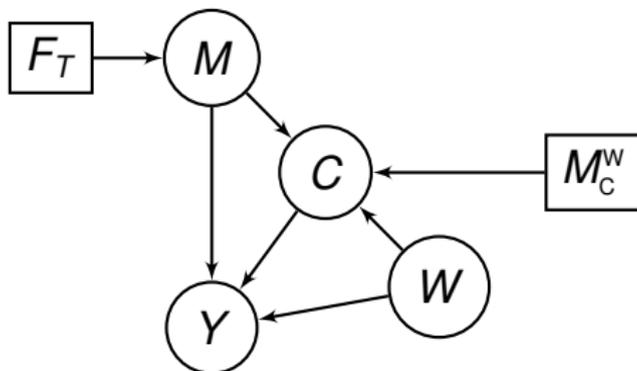
New variable and conditional independence - 4DI

Define M_C^W as a manipulation variable on C s.t.

$$(Y, C, W) \perp\!\!\!\perp F_T | T \quad (11)$$

$$W \perp\!\!\!\perp (F_T, M_C^W) \quad (12)$$

$$Y \perp\!\!\!\perp M_C^W | (T, C, W). \quad (13)$$



W as sex continued

Regimes of M_C^W

- 1 $M_C^W = \emptyset$: C is left alone
- 2 $M_C^W = t^*$: C is drawn from $p(C|T = t^*, W = w, F_T = \emptyset)$ where w corresponds to the realised value of W .

W as sex continued

Regimes of M_C^W

- 1 $M_C^W = \emptyset$: C is left alone
- 2 $M_C^W = t^*$: C is drawn from $p(C|T = t^*, W = w, F_T = \emptyset)$ where w corresponds to the realised value of W .

- No need for other regimes as they sever relationship between C and W and are equivalent to 3DI
- that means that there is only 1 extra DE
- Regime 2 induces

$$C \perp\!\!\!\perp (F_T, T) | M_C^W = t^* \quad (14)$$

Identifying DE's

We want the woman ($W = 1$) specific GDE_{t^*}

$$E(t, t^*, W = 1) - E(t^*, t^*, W = 1)$$

Identifying DE's

We want the woman ($W = 1$) specific GDE_{t^*}

$$\begin{aligned} E(t, t^*, W = 1) &- E(t^*, t^*, W = 1) \\ &= \sum_c [E(t, \emptyset, c, W = 1) \end{aligned}$$

Identifying DE's

We want the woman ($W = 1$) specific GDE_{t^*}

$$\begin{aligned} E(t, t^*, W = 1) &- E(t^*, t^*, W = 1) \\ &= \sum_c [E(t, \emptyset, c, W = 1) \times p(c|t^*, \emptyset, W = 1)] \end{aligned}$$

Identifying DE's

We want the woman ($W = 1$) specific GDE_{t^*}

$$\begin{aligned} E(t, t^*, W = 1) &- E(t^*, t^*, W = 1) \\ &= \sum_c [E(t, \emptyset, c, W = 1) \times p(c|t^*, \emptyset, W = 1)] - E(t^*, \emptyset, W = 1). \end{aligned}$$

Identifying DE's

We want the woman ($W = 1$) specific GDE_{t^*}

$$\begin{aligned} E(t, t^*, W = 1) &- E(t^*, t^*, W = 1) \\ &= \sum_c [E(t, \emptyset, c, W = 1) \times p(c|t^*, \emptyset, W = 1)] - E(t^*, \emptyset, W = 1). \end{aligned}$$

This can be further expressed with $F_T = \emptyset$ and can thus also be estimated from purely observational data.

Conclusions

- DAGs and conditional independences are useful for expressing D&I effects
- We need to be **clear** about what is going on though - what CI's hold?
- Provided we do this, we have a wealth of DEs we can estimate
- The reason that the idea can be confusing is because of the different types of DEs
- Also remember that the concepts don't make much sense if we cannot intervene on the intermediate at all.
- The 3DI and 4DI are tools that can be used to clarify what we mean.
- The framework can be extended further.

References

- Dawid, A.P. (2000). Causal Inference without Counterfactuals (with comments and rejoinder). *JASA* 95(450), 407-448.
- Dawid A.P. (2002). Influence diagrams for causal modelling and inference. *Intern. Statist. Rev.* 70:161-189.
- Geneletti S. (2007). Identifying direct and indirect effects in a non-counterfactual framework. *J ROY STAT SOC B.* 69:199-215
- Kuriki K. et al. (2005). Increased risk of colorectal cancer due to interactions between meat consumption and the CD36 gene A52C polymorphism among Japanese. *Nutr Cancer*;51(2):170-7
- Pearl, J. (2001), *Causality*