The decision theoretic approach to causal inference OR Rethinking the paradigms of causal modelling

A.P.Dawid[1] and S.Geneletti[2]

[1] University of Cambridge, Statistical Laboratory
[2] Imperial College Department of Epidemiology and Public Health

04/05/2009

Outline

- Issues
- The simple problem - RCT's
- The hard problem - Observational studies
- The statistical decision theoretic approach

Questions

- **Will** aspirin cure **my** headache?
- **Will** it help **those who are prescribed it**?
- **Did** it cure **my** headache?
- **Did** it help **those who were prescribed it**?
- **Would** I still have a headache if **I** hadn't taken it?

Distinctions

- Retrospective
- Cause of effect
- Counterfactual
- Deterministic
- Value
- Observation (passive)

- Prospective
- Effect of cause
- Hypothetical
- Stochastic
- Distribution
- Intervention (active)

Problems

Before data

- Meaning

Problems

Before data

- Meaning
- Interpretation

Problems

Before data

- Meaning
- Interpretation
- Inference

Problems

Before data

- Meaning
- Interpretation
- Inference

What data?

- Experimental studies
  - Randomisation
- Observational studies
  - Confounding
- Dynamic treatment regimes /Alternative treatment effects

Formal frameworks

Maths

- Potential responses
- Functional models
- Conditional independence

Formal frameworks

Maths

- Potential responses
- Functional models
- Conditional independence

Tools

- Structural equations
- Path diagrams
- Directed acyclic graphs

Which way to go?

- Are there any differences between the frameworks?

Which way to go?

- Are there any differences between the frameworks?
- What explicit/implicit assumptions do they make?

Which way to go?

- Are there any differences between the frameworks?
- What explicit/implicit assumptions do they make?
- How reasonable are they?

Which way to go?

- Are there any differences between the frameworks?
- What explicit/implicit assumptions do they make?
- How reasonable are they?
- How do they assist(impose) the way we
  - **pose,frame**

Which way to go?

- Are there any differences between the frameworks?
- What explicit/implicit assumptions do they make?
- How reasonable are they?
- How do they assist(impose) the way we
  - **pose,frame**

    and
  - **answer**

  causal queries

A simple problem

- Randomised experiment
- Binary treatment **decision** variable T
- Response random **variable** Y

Stats (101) model (Fisher)

- Specify conditional distribution of Y given $T = t$ ($t = 0, 1$)
  e.g.
  $$Y \sim N(\mu_t, \sigma^2)$$
- Sufficient to decide which decision is best
- Measure the effect of treatment by estimating
  $$\delta = \mu_1 - \mu_0 \tag{1}$$

Error model

$$Y = \mu_T + E_T$$
$$\mathbf{E} = (E_0, E_1)$$
$$\text{s.t. } E \sim N(\mathbf{0}, \mathbf{\Sigma})$$

- The values of **E** for any unit stay the same regardless of the T that unit receives.
- When $E = E_0 = E_1$ then this is a structural equation model

Potential responses model

- Imagine there are two Y's for each person (corresponding to the treatment T)

$$\begin{aligned} Y_0 &: \text{ response to } T = 0 \\ Y_1 &: \text{ response to } T = 1 \end{aligned}$$

- and these exist independently
- until the treatment you get **reveals** one of them and $Y = Y_T$
- unrevealed one becomes **counterfactual**

Potential responses model

- So for any unit there is a pair $\mathbf{Y} = (Y_1, Y_0)$ with some joint distribution
- The unit level (individual) **random** causal effect (ICA)

$$Y_1 - Y_0$$

is *unobservable*

Average Causal Effect

- This *is observable*

$$
\begin{aligned}
E(Y_1 - Y_0) &= E(Y_1) - E(Y_0) \\
&= E(Y|T = 1) - E(Y|T = 0) \\
&= \mu_1 - \mu_0
\end{aligned}
$$

General Functional Model

$$Y = f(T, U)$$
$$\text{(e.g. } U = \mathbf{Y}\text{)}$$

- Value of U would stay the same if we were to change T from 0 to 1

Connections

PR $\leftrightarrow$ GFM

Any functional model generates a potential responses model (and vice-versa as a PR model is a FM with $U = (Y_0, Y_1)$)

$$Y_t = f(t, U)$$

Stat $\subseteq$ PR

Any PR model generates a statistical model

$$Pr(Y_t) = Pr(Y|T = t)$$

and more than one PR model can correspond to the same stats model

Potential response models: Problems?

$$\begin{cases} Y_t \sim N(\mu_t, \sigma^2) & (t = 0, 1) \\ corr(Y_0, Y_1) = \rho \end{cases}$$

Corresponding stats model

$$Pr(Y|T = t) = \Phi_{\mu_t, \sigma^2}(Y)$$

where $\Phi(\cdot)$ is the cumulative distribution function for the $N(\mu_t, \sigma^2)$

NB: ρ does not feature $\Rightarrow$ it cannot be estimated!

Potential response models: Problems?

- Under the PR model

$$var(Y_1 - Y_0) = 2(1 - \rho)\sigma^2$$

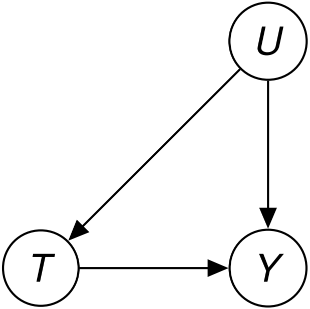
$\Rightarrow$ Cannot identify the pop^n variation in ICA

$$E(Y_1 - Y_0 | Y_1 = y_1) = (1 - \rho)y_1 + \rho\mu_1 - \mu_0$$

$\Rightarrow$ Cannot identify the counterfactual ICA having observed the response to the actual treatment (in this case $T = 1$)

Not so simple problem: Observational studies

- Treatment taken is associated to the patient's health (e.g. a confounder)
- What assumptions are required to make causal inferences?
- When and how can these assumptions be justified?



What are causal inferences?

- General consensus that they are about what happens when we intervene

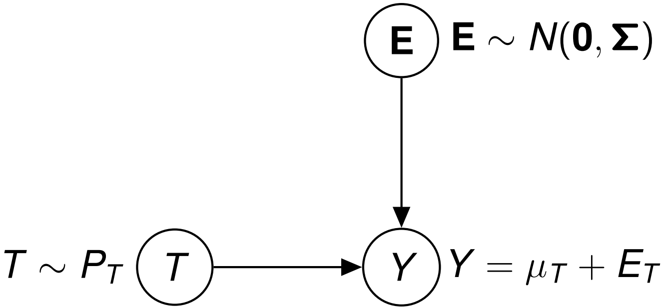
What are causal inferences?

- General consensus that they are about what happens when we intervene
- The big problem is that data are normally observational

What are causal inferences?

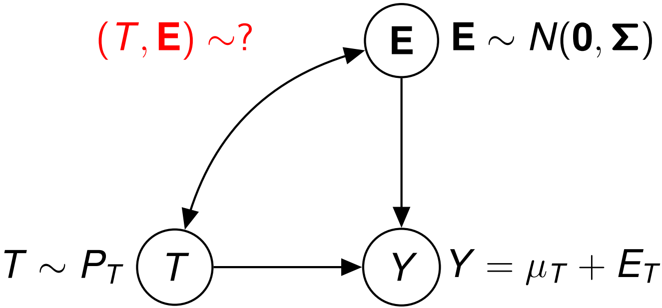
- General consensus that they are about what happens when we intervene
- The big problem is that data are normally observational
- Question then is, how do we make inferences about intervention from data that are observational?
- The different frameworks deal with this in different ways - more or less explicit

Error model



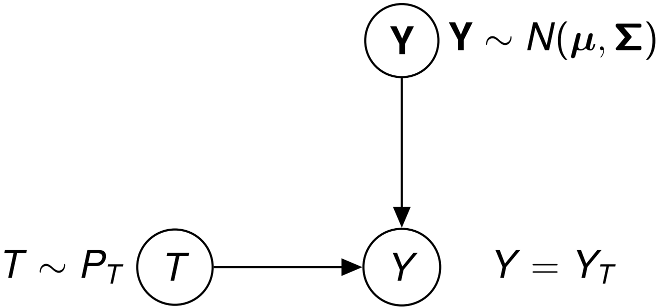
- No confounding $T \perp\!\!\!\perp \mathbf{E}$
  $\Rightarrow$ *treatment independent of errors*

Error model



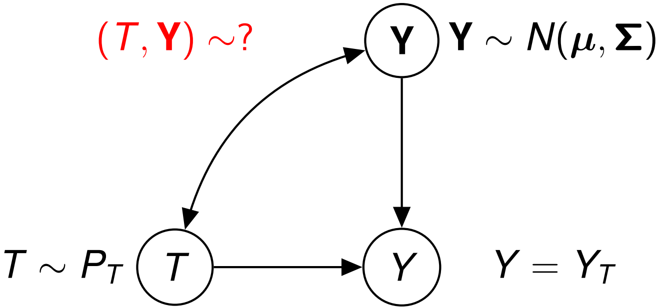
- No confounding $T \perp\!\!\!\perp \mathbf{E}$
  $\Rightarrow$ *treatment independent of errors*
- Otherwise what is joint of T and $\mathbf{E}$?

Potential responses model



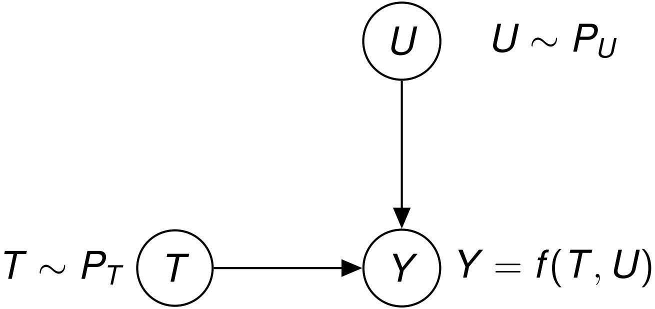
- Ignorable treatment assignment $T \perp\!\!\!\perp \mathbf{Y}$
  $\Rightarrow$ *treatment independent of PR's*

Potential responses model



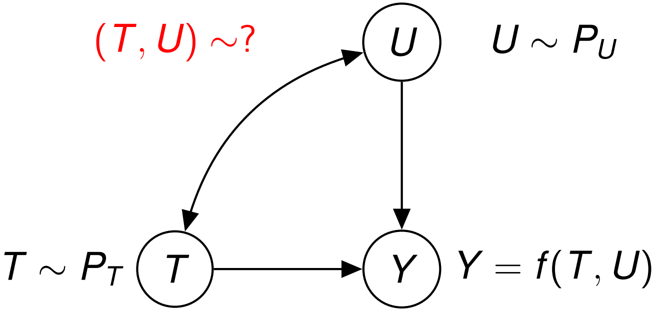
- Ignorable treatment assignment $T \perp\!\!\!\perp \mathbf{Y}$
  $\Rightarrow$ *treatment independent of PR's*
- Otherwise what is joint of T and $\mathbf{Y}$?

General Functional Model



- No confounding $T \perp\!\!\!\perp U$
  $\Rightarrow$ *treatment indep of unit characteristics*

General Functional Model



- No confounding $T \perp\!\!\!\perp U$
  $\Rightarrow$ *treatment indep of unit characteristics*
- Otherwise what is joint of T and U?

Potential response models: Problems?

- Value of $\mathbf{Y} = (Y_0, Y_1)$ for any unit the same for both experimental and observational cases
- as well as for either choice of T

Potential response models: Problems?

- Value of $\mathbf{Y} = (Y_0, Y_1)$ for any unit the same for both experimental and observational cases
- as well as for either choice of T
- So how are we to judge the independence of Y and T?

Potential response models: Problems?

- Value of $\mathbf{Y} = (Y_0, Y_1)$ for any unit the same for both experimental and observational cases
- as well as for either choice of T
- So how are we to judge the independence of Y and T?
- No reason to believe that responses the same under experiment and observation...

Statistical (Decision theoretic) Model

- Make the *regime* explicit with the variable F_T

Statistical (Decision theoretic) Model

- Make the *regime* explicit with the variable F_T
- $\mathbf{F_T = 1 \Rightarrow p(T = 1|F = 1) = 1}$ means set treatment 1 (as in RCT)

Statistical (Decision theoretic) Model

- Make the *regime* explicit with the variable F_T
- $\mathbf{F_T = 1 \Rightarrow p(T = 1|F = 1) = 1}$ means set treatment 1 (as in RCT)
- $\mathbf{F_T = 0 \Rightarrow p(T = 0|F = 0) = 1}$ means set treatment 0 (as in RCT)

Statistical (Decision theoretic) Model

- Make the *regime* explicit with the variable F_T
- $\mathbf{F_T = 1 \Rightarrow p(T = 1|F = 1) = 1}$ means set treatment 1 (as in RCT)
- $\mathbf{F_T = 0 \Rightarrow p(T = 0|F = 0) = 1}$ means set treatment 0 (as in RCT)
- $\mathbf{F_T = \emptyset \Rightarrow p(T = t|F = \emptyset) = p}$ means just observe, T arises “naturally" in the observational regime

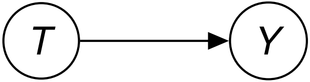
Statistical (Decision theoretic) Model

- Make the *regime* explicit with the variable F_T
- $\mathbf{F_T = 1 \Rightarrow p(T = 1|F = 1) = 1}$ means set treatment 1 (as in RCT)
- $\mathbf{F_T = 0 \Rightarrow p(T = 0|F = 0) = 1}$ means set treatment 0 (as in RCT)
- $\mathbf{F_T} = \emptyset \Rightarrow \mathbf{p(T = t|F} = \emptyset) = \mathbf{p}$ means just observe, T arises “naturally" in the observational regime
- Ignorable treatment assignment is $Y \perp\!\!\!\perp F_T | T$

Statistical (Decision theoretic) Model

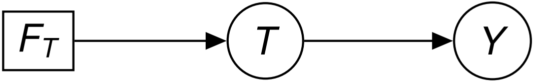
- Make the *regime* explicit with the variable F_T
- $\mathbf{F_T = 1 \Rightarrow p(T = 1|F = 1) = 1}$ means set treatment 1 (as in RCT)
- $\mathbf{F_T = 0 \Rightarrow p(T = 0|F = 0) = 1}$ means set treatment 0 (as in RCT)
- $\mathbf{F_T} = \emptyset \Rightarrow \mathbf{p(T = t|F} = \emptyset) = \mathbf{p}$ means just observe, T arises “naturally" in the observational regime
- Ignorable treatment assignment is $Y \perp\!\!\!\perp F_T | T$ simple

Influence diagrams



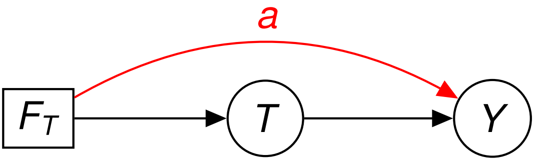
- Start simply

Influence diagrams



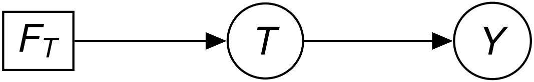
- Start simply
- Add regime indicator node – non random so in a box

Influence diagrams



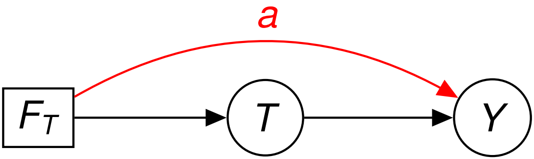
- Start simply
- Add regime indicator node – non random so in a box
- Absence of arrow *a* means $Y \perp\!\!\!\perp F_T \mid T$

Confounders



- $Y \perp\!\!\!\perp F_T | T$ simple case

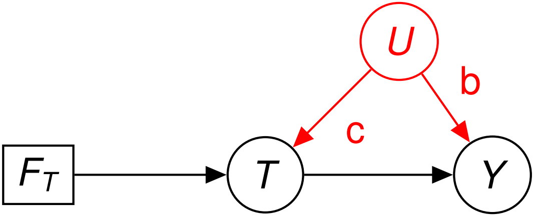
Confounders



- $Y \perp\!\!\!\perp F_T \,|\, T$ simple case
- If *a* then often

Confounders

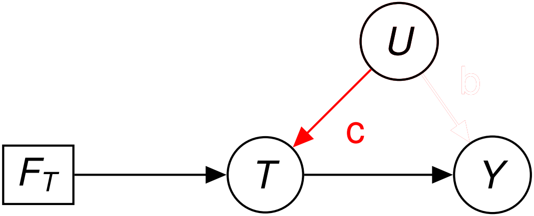
$$U \perp\!\!\!\perp F_T$$
$$Y \perp\!\!\!\perp F_T | (T, U)$$



- $Y \perp\!\!\!\perp F_T | T$ simple case
- If *a* then often U (un)confounder
- Treatment assignment is ignorable conditional on U

Confounders

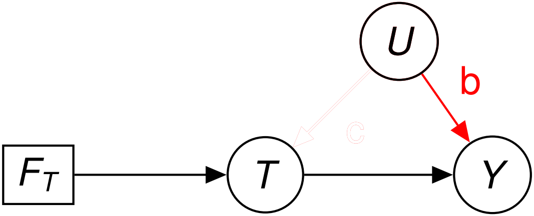
$$U \perp\!\!\!\perp F_T$$
$$Y \perp\!\!\!\perp F_T | (T, U)$$



- $Y \perp\!\!\!\perp F_T | T$ simple case
- If *a* then often U (un)confounder
- Treatment assignment is ignorable conditional on U
- If *b* absent ($T \perp\!\!\!\perp U | F_T$)

Confounders

$$U \perp\!\!\!\perp F_T$$
$$Y \perp\!\!\!\perp F_T | (T, U)$$



- $Y \perp\!\!\!\perp F_T | T$ simple case
- If *a* then often U (un)confounder
- Treatment assignment is ignorable conditional on U
- If *b* absent ($T \perp\!\!\!\perp U | F_T$)or *c* absent ($Y \perp\!\!\!\perp U | T$) then marginally ignorable

Causal Model

- Simply a more ambitious non-causal model
- expressing the invariance of certain modular structures across different regimes

Causal Model

- Simply a more ambitious non-causal model
- expressing the invariance of certain modular structures across different regimes
- E.g. something that behaves in the same way under observational and experimental regimes is a candidate for a stable relationship – causal

Causal Model

For a functional (e.g. PR) model

- invariant values of variables and functional relationships
- implicit, deterministic

Causal Model

For a functional (e.g. PR) model

- invariant values of variables and functional relationships
- implicit, deterministic

Statistical Model

- invariant conditional distributions
- explicit, stochastic

Brief word on estimation

PR model

- Expectation of responses over those we already treated
- Deals with what would have happened to Jack who we treated if he had not been treated?

Brief word on estimation

Statistical Model

- Bayesian predictive expectation of response for a new patient
- Deals with given we have observed Jack-like individuals, what decision should we recommend to a new patient exchangeable with Jack?
- Hence the name Decision theoretic

Advantages

- No *impossible to observe*-ables
- Stochastic not deterministic relationships
- Simple, explict and testable assumptions
- Focussed on "what is the best decision for the future" rather than "what would have happened if"

Issues tackled

- Compliance
- Dynamic treatment regimes
- Alternative treatment measures
- Direct and Indirect effects

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. Stats. Rev.* **70**, 161-189
- Dawid, A.P. (2003) Causal inference using influence diagrams: The problem of partial compliance (with Discussion). In *Highly Structured Stochastic System*, Eds P.J.Green, N.L.Hjort and S.Richardson. Oxford University Press
- Dawid, A.P. (2004) Probability, causality and the empirical world: A Bayes-de Finetti-Popper-Borel synthesis. *Statistical Science* **19**, 44-57
- Didelez,V and Dawid, A.P. (2008) Identifying optimal sequential decisions, In *Proceedings of the 24th Annual Conference on Uncertainty in Artifical Intelligence*, 113-120
- Geneletti, S. (2007). Identifying direct and indirect effects in a non-counterfactual framework. *J ROY STAT SOC B.* 69:199-215
- Geneltti, S and Dawid, A.P. (2009) Defining and Identifying the Effect of Treatment on the Treated, Technical Report, Imperial College, London
- Pearl, J. (2000) Causality, Cambridge University
- Rubin, D.B. (1974). Estimating causal effects of treatments in randomized and non-randomized studies. *Journal of Educational Psychology* **66**(5), 699-701.