Bristol, 27th June 2010

Introduction to causal inference via potential outcomes

Leonardo Grilli & Carla Rampichini
Dept. of Statistics, University of Florence

This introduction is a personal elaboration of slides and papers of Donald Rubin

Basic concepts /1

- Three key notions underlying the potential outcome approach (also called *Rubin Causal Model*):
  - **potential outcomes** corresponding to the various levels of a treatment or manipulation ("*no causation without manipulation*")
  - **multiple units** and the related stability assumption
  - **assignment mechanism**, which is crucial for inferring causal effects

2

Basic concepts /2

- ***Unit***: The person, place, or thing upon which a treatment will operate, at a particular time *Note: a single person, place, or thing at two different times comprises two different units.*
- ***Treatment***: An intervention, the effects of which (on some particular measurement of the units) the investigator wishes to assess relative to no intervention (i.e., the control)

3

Basic concepts /3

- ***Potential Outcomes***: The values of a unit's measurement of interest after (a) application of the treatment and (b) non-application of the treatment (i.e., under control)
- ***Causal Effect***: For each unit, the comparison of the potential outcome under treatment and the potential outcome under control

4

The Fundamental Problem of Causal Inference

- ***We can observe at most one of the potential outcomes for each unit***
- Let Y(not) denote the outcome given the control treatment, and Y(asp) the outcome given the active treatment

Unit	Initial Headache	Potential Outcomes	
	X	$Y(asp)$	$Y(not)$
you	80	25	75

Causal effect for you = Y(asp)-Y(not) = -50

unobservable

- We must rely on multiple units exposed to different treatments to make causal inferences:
  - observe the same physical object subject to different treatment levels at different points in time
  - observe different physical units at the same time

5

Counterfactuals

Rubin JASA 2005 wrote on page 325:

- Some authors call the potential outcomes **counterfactuals**, borrowing the term from philosophy
- I much prefer Neyman's implied term **potential outcomes**, because these values are not counterfactual until after treatments are assigned, and calling all potential outcomes "counterfactuals" certainly confuses quantities that can never be observed (e.g. your height at age 3 if you were born yesterday in the Arctic) and so are truly a priori counterfactual, with unobserved potential outcomes that are not a priori counterfactual

6

Need of a stability assumption

- For causal inference we need multiple units (i.e. ***replications***)
- But we also need a stability assumption to limit the number of cases to consider:
  - 2 units: CC, TC, CT, TT → $2^2 = 4$ cases
  - 3 units: CCC, TCC, CTC, CCT, … → $2^3 = 8$ cases
  - … ghosh!

All causal inference relies on assumptions that restrict the possible potential outcomes so that we can learn something about causal effects from observable data. Nothing is wrong with making assumptions; on the contrary, such assumptions are the strands that join the field of statistics to scientific disciplines. The quality of these assumptions and their precise explication, not their existence, is the issue (D. Rubin)

7

SUTVA (Stable Unit Treatment Value Assumption)

SUTVA has two parts:

a) ***No multiple versions of the treatment***: for each unit there is only one form of the treatment (and one form of the control)

b) ***No interference among units***: each unit's potential outcomes remain the same, no matter what treatment the other units receive

8

Violation of SUTVA part (a)

- Multiple versions of the treatment → more values of the treatment variable → more potential outcomes
- Example: suppose there are two types of aspirin tablets, strong (Asp+) and weak (Asp-), then
  - the treatment variable takes 3 values: Asp+, Asp-, Not
  - → 3 potential outcomes: Y(Asp+), Y(Asp-), Y(Not)

9

Violation of SUTVA part (b)

You take:	Asp	Not	Asp	Not
I take:	Asp	Not	Not	Asp
Unit				
1=you	$Y_1([A,A]) = 0$	$Y_1([N,N]) = 100$	$Y_1([A,N]) = 50$	$Y_1([N,A]) = 75$
2=me	$Y_2([A,A]) = 0$	$Y_2([N,N]) = 100$	$Y_2([A,N]) = 100$	$Y_2([N,A]) = 0$

- The effect of taking aspirin ***for me*** is -100 regardless of whether you take aspirin
- The effect of taking aspirin ***for you*** is
  - -75 if I take aspirin
  - -50 if I do not take aspirin

interference

10

Violation of SUTVA part (b) /cont

- Interferences can be ruled out in clinical trials but they arise in some settings, e.g.
  - The effect of vaccines on contagious diseases
  - Substitution effects in the labour market
- Solutions:
  - Change the unit of analysis (e.g. in case of a contagious disease the unit could be the village instead of the person)
  - Build a (parsimonious) model for the interactions (e.g. allow for interference among adjacent units)

11

Unit-level causal effects (N units & SUTVA)

Unit	X	Y(Asp)	Y(Not)	Unit Level Causal Effect
1	X_1	$Y_1(Asp)$	$Y_1(Not)$	$Y_1(Asp)-Y_1(Not)$
…	…	…	…	…
i	X_i	$Y_i(Asp)$	$Y_i(Not)$	$Y_i(Asp)-Y_i(Not)$
…	…	…	…	…
N	X_N	$Y_N(Asp)$	$Y_N(Not)$	$Y_N(Asp)-Y_N(Not)$

- This array represents the "science" (what we would like to know about the world)
- Causal effects are defined at the unit level
- Unit-level causal effects are not observable (the *Fundamental Problem of Causal Inference*)
- The definition does not involve the notion of probability

12

Population-level causal effects (N units & SUTVA)

- Population-level causal effects are usually obtained by summarizing unit-level causal effects, e.g.
  - ***Average Causal Effect*** (ACE) or ***Average Treatment Effect*** (ATE)

$$\begin{aligned} &Ave\left[Y_i(Asp) - Y_i(Not)\right] = \\ &= Ave\left[Y_i(Asp)\right] - Ave\left[Y_i(Not)\right] \\ &= \frac{1}{N}\sum_{i=1}^{N} Y_i(Asp) - \frac{1}{N}\sum_{i=1}^{N} Y_i(Not) \end{aligned}$$

  - The average can be restricted to subsets, such as males (ATE on male) or treated units (ATT: ***Average Treatment effect on the Treated***)

13

Population-level causal effects /cont

- Other summaries of unit-level causal effects:
  - ***Average ratio*** = Ave{Y_i(Asp)/Y_i(Not)}
  - ***Median causal effect*** of "Asp" vs. "Not" = Median{Y_i(Asp) - Y_i(Not)}
- Alternatively, a population-level causal effect can be defined as a comparison between the two distributions of the potential outcomes, e.g.
  - ***Difference of median of potential outcomes*** = Median{Y_i(Asp} - Median{Y_i(Not)}
- A population-level causal effect synthesizes the "science" … unfortunately it is unobservable (again the *Fundamental Problem of Causal Inference*)

14

Assignment mechanism

- To learn about causal effects, we must have replication (observe Y(Asp) for at least one unit and Y(Not) for at least one unit)
- The assignment mechanism determines which potential outcome we will observe for each unit: it is a (usually *probabilistic*) rule for selecting some units to receive control and other units to receive treatment
- The assignment mechanism is critical, even when SUTVA holds: we must know or posit a rule for how each unit received treatment or control

15

Some notation

- W_i → treatment indicator of unit i (usually 0 stands for control and 1 for treatment)
- $Y_i(0)$ and $Y_i(1)$ → *potential* outcomes of unit i
- Y_i^{obs} → *observed* outcome of unit i
- Y_i^{mis} → *missing* outcome of unit i

$$Y_i^{obs} \equiv Y_i(W_i) = W_i \cdot Y_i(1) + (1 - W_i) \cdot Y_i(0)$$
$$Y_i^{mis} \equiv Y_i(1 - W_i) = (1 - W_i) \cdot Y_i(1) + W_i \cdot Y_i(0)$$

Vector notation: $\mathbf{W} = (W_1,\ldots,W_n)$ is the vector of treatment indicators for all units, $\mathbf{Y}(0) = (Y_1(0),\ldots,Y_n(0))$ etc.

16

Types of Assignment Mechanism /1

- ***Probabilistic AM***: for every unit the probability of assignment is always strictly between 0 and 1

$$0 < Pr\left(W_i = 1 \mid \mathbf{X}, \mathbf{Y}(0), \mathbf{Y}(1)\right) < 1$$

- ***Unconfounded AM***: the assignment of treatment or control for all units is independent (conditionally on the covariates) of all the potential outcomes, observed or unobserved

$$Pr\left(\mathbf{W} \mid \mathbf{X}, \mathbf{Y}(0), \mathbf{Y}(1)\right) = Pr\left(\mathbf{W} \mid \mathbf{X}\right)$$

17

Types of Assignment Mechanism /2

- An AM is ***confounded*** when there are individuals with probability of being assigned to treatment (i.e. $W_i=1$) depending on the potential outcomes $Y_i(0), Y_i(1)$ (even conditionally on the covariates)
- This is *a serious issue in observational studies*, where the assignment is not enforced by the experimenter but chosen by the individual itself
- Usually the dependence of W_i on $Y_i(0), Y_i(1)$ is due to common unobserved variables (***confounders***)
  - Example: some unemployed persons are given the chance to attend a training program ($W_i=1$ if attended) and we wish to evaluate the effect on the time to get job ($Y_i(0)$ if not attended, $Y_i(1)$ if attended); *motivation* may be an unobserved variable affecting both W_i and $Y_i(0), Y_i(1)$, implying a confounded AM

18

Types of Assignment Mechanism /3

- An unconfounded AM is the key for unbiased estimation of causal effects – but for the purpose of unbiased estimation even the weaker condition of an ignorable AM is sufficient
- ***Ignorable AM***: the assignment of treatment or control for all units is independent (conditionally on the covariates) of the unobserved potential outcomes

$$Pr\left(\mathbf{W} \mid \mathbf{X}, \mathbf{Y}(0), \mathbf{Y}(1)\right) = Pr\left(\mathbf{W} \mid \mathbf{X}, \mathbf{Y}^{obs}\right)$$

- Example: in a sequential clinical trial the AM is ignorable though confounded

Unconfounded → Ignorable
Non-ignorable → Confounded

19

Example: the Perfect Doctor

- The hypothetical data given below show all potential outcomes under two different treatments:
  - Y(0) = years lived after standard surgery
  - Y(1) = years lived after new surgery

Unit	Y(0)	Y(1)
1	13	14
2	6	0
3	4	1
4	5	2
5	6	3
6	6	1
7	8	10
8	8	9
True aver.	7	5

$$true\ ACE = \\ = Ave\left[Y_i(1)\right] - Ave\left[Y_i(0)\right] \\ = 5 - 7 \\ = -2$$

The treatment, on average, is deleterious (it subtracts two years of life)

20

Example: the Perfect Doctor /cont

- The perfect doctor chooses the best treatment for each patient (i.e. the treatment under which the patient will live longer)
- What we would actually observe?

Unit	W	Y(0)	Y(1)
1	1	?	14
2	0	6	?
3	0	4	?
4	0	5	?
5	0	6	?
6	0	6	?
7	1	?	10
8	1	?	9
Observed averages		5.4	11

W denotes the treatment received (1=new, 0=standard)

$$observed\ ACE = \\ = Ave\left[Y_i(1) \mid W_i = 1\right] - Ave\left[Y_i(0) \mid W_i = 0\right] \\ = 11 - 5.4 \\ = 5.6$$

On the basis of the sample means we would WRONGLY conclude that the treatment, on average, is beneficial

21

Example: the Perfect Doctor /cont

- In the Perfect Doctor example, the treatment each unit receives depends on that unit's potential outcomes (both missing and observed) → the assignment mechanism is confounded → the ACE estimator is biased
- How to get an unconfounded assignment mechanism? Simply flip a coin!
- In fact, if the units are assigned to treatment or control on the basis of random draws, the assignment does not depend on potential outcomes (this is the reason why randomization is so popular!)

22

Example: the Perfect Doctor /cont

- The true ACE is a parameter (causal estimand)
- The observed ACE is an estimator
- ***Unbiased estimator***: if treatment assignment is repeated again and again, the average estimate is equal to the parameter
- In the Perfect Doctor example there are 56 ways of assigning 3 units to treatment and 5 units to control

Perfect Doctor assignment

#1 → 11100000 (observed ACE = -1.6)
…….
#21 → 10000011 (observed ACE = 5.6)
…….
#56 → 00000111 (observed ACE = -0.1)

The average of the 56 observed ACEs is just -2 (the true ACE)

Perfect Doctor AM: always select #21 → observed ACE is biased
Randomized AM: select # at random → observed ACE is unbiased

23

Assignment mechanism and science

- The assignment mechanism does not change the potential outcomes (the "science"): it simply reveals one outcome per unit
- The ***observed outcome*** Y_i^{obs} mixes the "science" and the assignment mechanism: this harmful blend affects regression models

$$Y_i^{obs} = \alpha + \beta W_i + \gamma X_i + e_i$$

24

Experiment vs Observation

- A **randomized experiment** is characterized by an assignment mechanism which is
  - Probabilistic
  - Ignorable
  - Known (chosen by the experimenter)
- An **observational study** is characterized by an assignment mechanism which is
  - Unknown (typically, the probability to take treatment or control depends on unknown parameters)

25

Introductory lectures

- Rubin D.B. (2006). Statistical Inference for Causal Effects, with Emphasis on Applications in Psychometrics and Education. *Handbook of Statistics, Vol. 26 Ch. 24.*
- Rubin D.B. (2008). Statistical Inference for Causal Effects, with Emphasis on Applications in Epidemiology and Medical Statistics. *Handbook of Statistics, Vol. 27 Ch. 2.*
- Mealli F., Pacini B., Rubin D.B. (2011). Statistical Inference for Causal Effects. *In: R. Kenneth, S. Salini (eds.) Modern Analysis of Customer Satisfaction Surveys*. Wiley (forthcoming)
- Rubin D.B. (2005). Causal inference using potential outcomes: design, modeling, decisions. 2004 Fisher Lecture. *Journal of the American Statistical Association* 100, 322-331.
- … still waiting for *the* book - Imbens G. & Rubin D. *Causal inference: Statistical methods for estimating causal effects in biomedical, social, and behavioral sciences (?)*

26