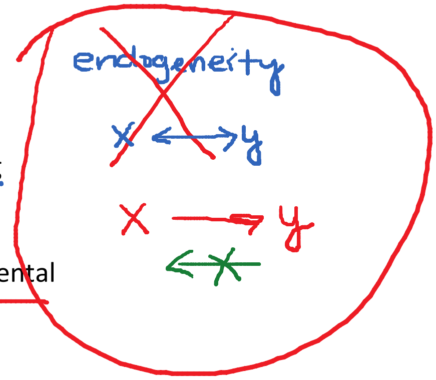
Introduction to matching

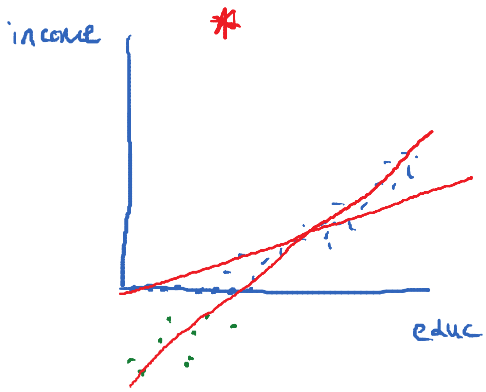
Tuesday, November 13, 2012
9:09 PM

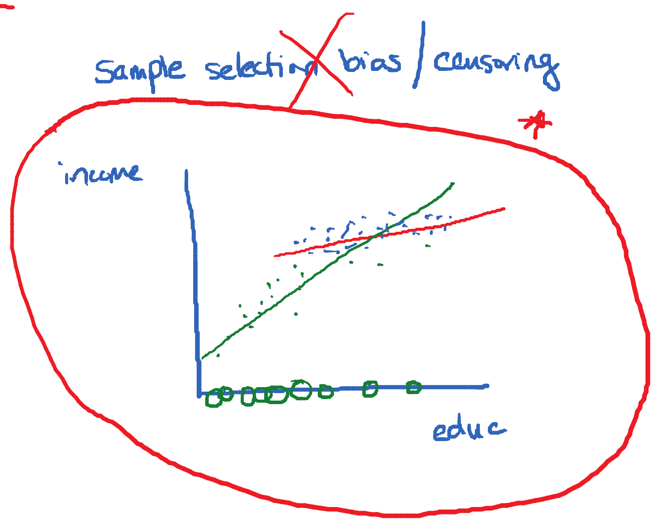


- In brief, matching is a way of non-parametrically controlling for confounding variables

- Not an inference procedure per se, but a method of reconstructing experimental conditions via observable contextual variables

- Sometimes characterized as a "causal inference" procedure…
  - ...which it is, but only under certain conditions
  - (so is linear regression, under certain conditions)





Causal inference

Tuesday, November 13, 2012
9:23 PM

- What is causality? The Neyman-Rubin causal model

for some set of treatment conditions

$$T \in \{T_1, T_2, \ldots, T_k\}$$

$y_i(T_j) \equiv$ state (of DV) of object i under treatment j

$$\delta_i^{jk} = y_i(T_j) - y_i(T_k)$$

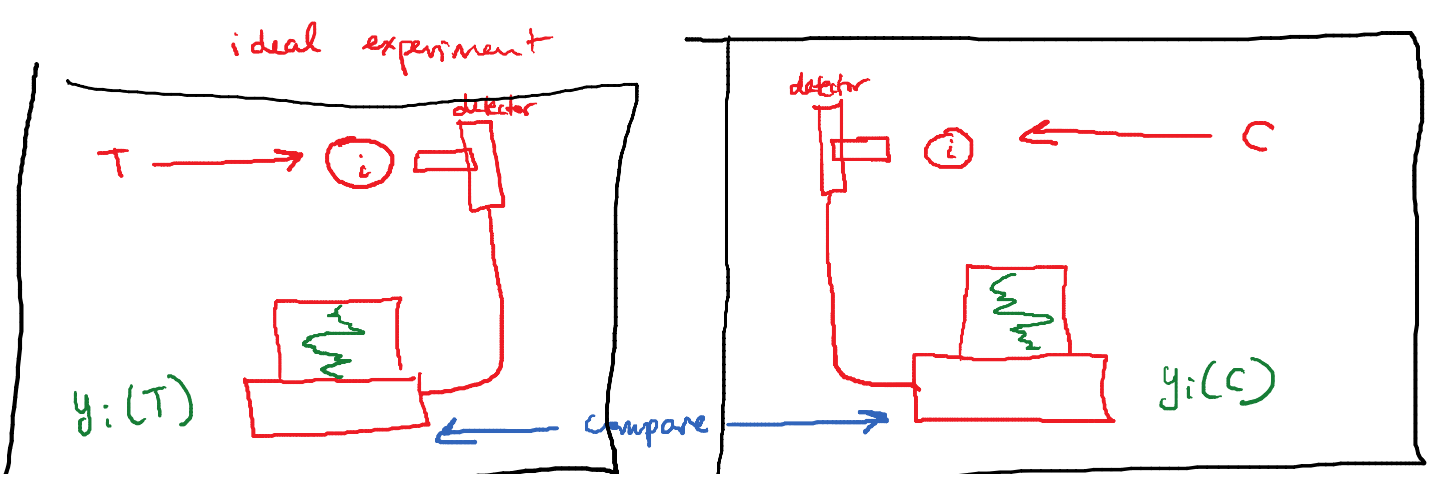
- Experimentally, T is often (by no means always) conceived as binary

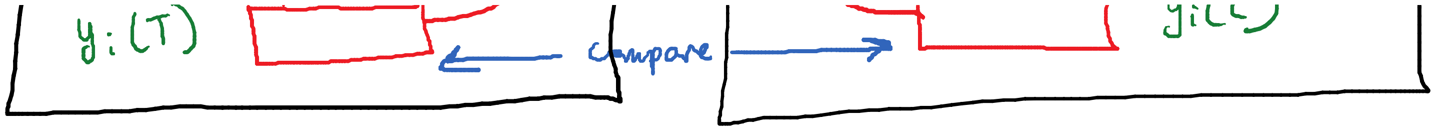
treatment   control

δ_i = "treatment effect" for person i

$$= y_i(T) - y_i(C)$$

- The fundamental problem of causal inference





We can't expose i to 2 treatments (at least not at the same time.)

sequential? order effect
temporal effect
background effects

- we expose $i \leftarrow T$ and compare i & j.
  $j \leftarrow C$

$$y_i(T) - y_j(C)$$

How do we know that y is attributable to treatment differences and not other differences between i & j?

$y = f(S)$ $S \equiv$ state of the world

$y = f(T, \phi)$ T: treatment condition
ϕ: everything else

$$\frac{y(T + dT, \phi_0) - y(T, \phi_1)}{dT} = \frac{dy}{dT} \Big| \phi$$

ϕ_0 : person 0's state

ϕ_1 : person 1's state

$\phi_0 \neq \phi_1$

$$\text{ATE} : E[\delta_i] = E_\phi\left[y(T, \phi) - y(C, \phi) \right]$$

$$= \int \left(f(T, \phi) - f(C, \phi) \right) g(\phi)\, d\phi$$

$$= \int f(T, \phi) g(\phi) d\phi - \int f(C, \phi) g(\phi) d\phi$$

$$\frac{1}{N} \sum_{i=1}^{N} f(T, \phi_i) - \frac{1}{N} \sum_{i=1}^{N} f(C, \phi_i)$$

Treatment — Control

$N \rightarrow \infty$

$$\frac{\#(\phi = \phi_0)}{N} \rightarrow g(\phi)$$

Assumptions that make causal inference plausible

Tuesday, November 13, 2012
9:30 PM

- What is a simple set of assumptions we can make in order to identify causality?

$$g(\phi | T) = g(\phi | C) = g(\phi)$$

$$\mathrm{cov}(\phi, T/C)$$

$$E[(\phi | T) - (\phi | C)] = 0$$

experiments: random assignment of T/C.

Assume / measure away

① SUTVA: $f(C, \phi \mid A = C) = f(C, \phi \mid A = T)$

reaction to the treatment/control is not a function of the assignment process.

Matching.

* ② $g(\phi \mid A = C) = g(\phi \mid A = T) = g(\phi)$

- Even this simple assumption can be problematic...

omitted variable bias as a threat to SUTVA.

- Calculate average treatment effect, where the average is taken over ϕ

When does ATE = Causal relationship?

Tuesday, November 13, 2012
9:46 PM

- Simply calculating and comparing treatment averages requires two conditions:

1) Balance - we need ϕ to be distributed similarly between the treatment conditions

2) SUTVA - the stable unit treatment value assumption:

   $f(C, \phi | D = 1) = f(C, \phi | D = 0)$

   The treatment assignment process does not change the causal relationship.

Calculating Treatment Effects using Matching

Tuesday, November 13, 2012
10:02 PM

- Matching methods are an attempt to construct a data set where ϕ is equal between the treatments to allow the calculation of treatment effects

$$E[SAT \mid private] - E[SAT \mid public]$$

use matching

$$\rightarrow E[SAT \mid private, rich] - E[SAT \mid public, rich] = \delta_{rich}$$

$$\rightarrow E[SAT \mid private, poor] - E[SAT \mid public, poor] = \delta_{poor}$$

$$E[SAT \mid private, w] - E[SAT \mid public, w] = \delta_w$$

1) Average Treatment Effect (ATE)

$$\delta_{rich} \cdot pr(rich) + \delta_{poor} \cdot pr(poor)$$

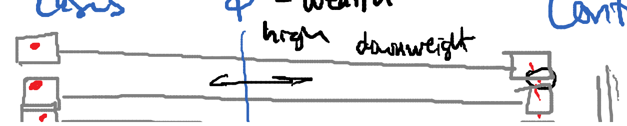
$$\int \delta_w \cdot g(w)\, dw = ATE$$

2) Average Treatment Effect on the Treated (ATT)

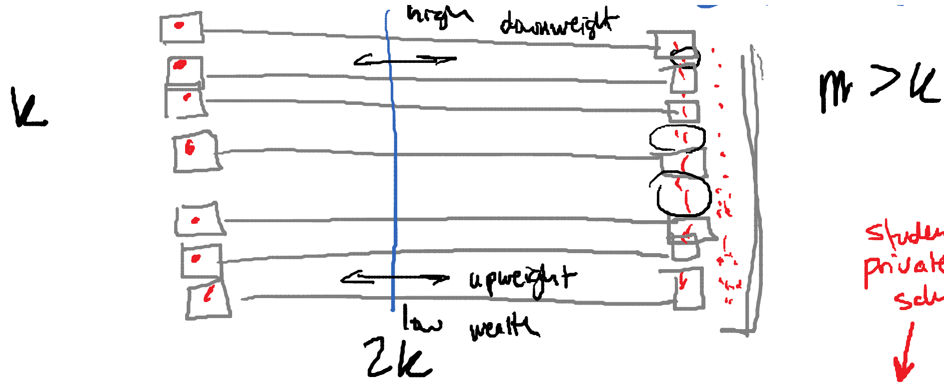
$$E[SAT \mid private, \text{went to private school}] - E[SAT \mid public, \text{went to private school}]$$

- This can typically be calculated with a mean comparison test or (weighted) regression on the matched data

Treatment cases ϕ = wealth Control Cases



$m > k$



students private school → SAT ~ wealth, if public

- We can also calculate this with post-matching regression analysis

   1) Regress on control data (or on all data, if data set is small)
   2) Use fitted model to generate fitted predictions on the treatment data
   3) Compare fitted predictions to observed values

* SAT ~ treat + wealth (treat = 0; private school students)

all treatment data is matched already.

N = # private school sample

$$\frac{1}{N} \sum_i SAT^i_{private} - \widehat{SAT}^i_{private|public} = ATT.$$

on matched data only

   1) Regress on treatment data (or on all data, if data set is small)
   2) Use fitted model to generate fitted predictions on the control data
   3)
   4) Compare fitted predictions to observed values

include data not matched.

M = # public school sample

$$\frac{1}{M} \sum \widehat{SAT}^i_{public|private} - SAT_{public} = ATC$$

$$ATE. = \frac{1}{N+M} \left[\sum_i SAT^i_{public|private} - SAT_{public} \right] + \left[\sum_i SAT^i_{private} - \widehat{SAT}^i_{priv|pub} \right]$$

Propensity scores

Tuesday, November 13, 2012
11:25 PM

- How do we match when there are multiple potential confounders (i.e., almost always)?
- One technique: propensity score matching

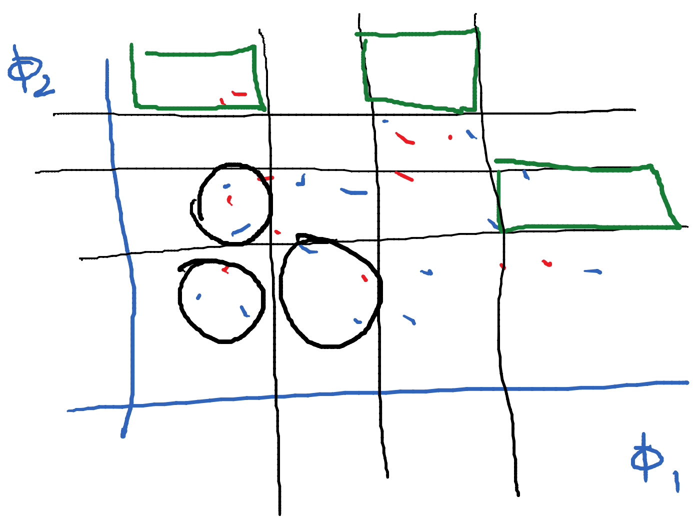
ϕ : background confounders

$$pr(\text{treatment}) = \Lambda(\phi\beta)$$

match on pr(treatment) instead of directly on ϕ.

- Another idea: Coarsened Exact Matching

— Stratifies ϕ on every subdimension of ϕ



Matching techniques

Tuesday, November 13, 2012
11:07 PM

- For multidimensional matching problems (i.e., all the important ones), there are different ways of achieving matches

- Each matching technique is designed to achieve better balance
- Many techniques (and others besides!) are implemented in software designed to implement matching

- Discuss some techniques implemented in MatchIt

1) Nearest-neighbor matching via propensity score
2) Matching via genetic optimization of balancing metrics
3) Coarsened exact matching

assess balance manually

rgenoud automatically optimizes balance.