Introduction to observational studies

Thomas Leavitt and Ben Hansen

June 26, 2024

Review: Randomized Experiments

Randomized experiments

- **Treatment**: z_i is indicator of treatment for unit i, where $i = 1, \dots, N$
  - The subscript i is a placeholder referring to arbitrary unit
  - E.g., z_5 is the treatment indicator for the 5th unit

$$z_i = \begin{cases} 1 & \text{if unit } i \text{ assigned to treatment} \\ 0 & \text{otherwise} \end{cases}$$

- In randomized controlled trial,

  each subject assigned to treatment and control via known, chance process
  - E.g., if we flip a fair coin for each unit, probability of treatment assignment is 0.5 for all units; $\Pr(Z_i = 1) = 0.5$ for all i

Randomized experiments

- Randomization $\implies$

  procedure for generating estimates of ATE will be correct, on average

- Randomization is foundation for valid hypothesis tests about causal effects

  - The "reasoned basis" for inference (Fisher, 1935, p. 14)

Observational studies

Observational studies

- Observational study:

  Empirical investigation in which "it is not feasible to use controlled experimentation in the sense of being able to ... assign subjects at random to different procedures" (Cochran, 1965)

- In observational study, subjects select into treatment and control conditions via unknown, chance process

  - Think of this as
    "assignment to treatment group on the basis of a covariate" (Rubin, 1977)

Observational studies

- In observational study, subjects select into treatment and control conditions via unknown, chance process

- Propensity score

  Probability of selecting into treatment as a function of baseline covariates

  - Propensity scores are unknown and cannot be directly measured

  - In principle, baseline covariates that determine treatment assignment probabilities can be measured

Observational studies

Model of an observational study (Rosenbaum, 2002)

- Independent but not necessarily identically distributed assignments

$$\Pr(\boldsymbol{Z} = \boldsymbol{z}) = \prod_{i=1}^{N} \pi_i^{z_i} (1 - \pi_i)^{1 - z_i}, \tag{1}$$

  where $\pi_i \in [0, 1]$ is probability of treatment

- Propensity score

- For each unit, π_i is function of covariates, either observed, $\boldsymbol{x}_i$, or unobserved, $\boldsymbol{u}_i$
  - Under no hidden bias, $\pi_i = \lambda(\boldsymbol{x}_i)$, where λ is a function, $\lambda : \mathbb{R}^K \mapsto [0, 1]$, whose form is unknown
  - Under common support, $\pi \in (0, 1)$ for all $i = 1, \ldots, N$

Observational studies

- We want to compare "apples to apples," *not* "apples to oranges" (Rubin and Waterman, 2006)

  - I.e., make strata of treated and control units homogeneous in propensity scores

- Although propensity scores unknown, if any two units are same on covariates that determine treatment assignment probability

  $\implies$ identical propensity scores

Observational studies

- If we can justify we are comparing treated and control units with same propensity scores

  $\rightarrow$ Then act as if we have mini randomized experiment within each block

- We do **not** need to know units' true propensity scores

- We need to justify only that propensity scores are the same within blocks

  - E.g., in block A with one treated and one control unit, it does not matter if propensity scores are 0.7 and 0.7; 0.2 and 0.2; or any other values

Observational studies

- More formally, let $s = 1, \ldots, S$ tun over S strata

- If $\pi_{s,i} = \pi_{s,j}$ for all $i, j = 1, \ldots, n_s$ in stratum s, then

$$\Pr\left(\boldsymbol{Z}_z = \boldsymbol{z}_s\right) = \frac{1}{|\Omega_s|} \text{ for all } \boldsymbol{z}_s \in \Omega_s, \text{ where } \Omega_s = \left\{ \boldsymbol{z}_s : \sum_{i=1}^{n_s} z_{s,i} = n_{1,s} \right\}$$

⋆ As-if randomization holds conditional on observed number of treated units

Observational studies

- E.g., stratum with 3 units, 1 treated and 2 control, all with propensity score of 0.35
- Without conditioning on 1 treated unit, model in (1) yields

$$\Pr\left(\boldsymbol{Z}_s = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^\top\right) = (1 - 0.35)(1 - 0.35)(1 - 0.35) = 0.274625$$

$$\Pr\left(\boldsymbol{Z}_s = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^\top\right) = (1 - 0.35)(1 - 0.35)(0.35) = 0.147875$$

$$\Pr\left(\boldsymbol{Z}_s = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^\top\right) = (1 - 0.35)(0.35)(1 - 0.35) = 0.147875$$

$$\Pr\left(\boldsymbol{Z}_s = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^\top\right) = (0.35)(1 - 0.35)(1 - 0.35) = 0.147875$$

$$\Pr\left(\boldsymbol{Z}_s = \begin{bmatrix} 0 & 1 & 1 \end{bmatrix}^\top\right) = (1 - 0.35)(0.35)(0.35) = 0.079625$$

$$\Pr\left(\boldsymbol{Z}_s = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}^\top\right) = (0.35)(1 - 0.35)(0.35) = 0.079625$$

$$\Pr\left(\boldsymbol{Z}_s = \begin{bmatrix} 1 & 1 & 0 \end{bmatrix}^\top\right) = (0.35)(0.35)(1 - 0.35) = 0.079625$$

$$\Pr\left(\boldsymbol{Z}_s = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^\top\right) = (0.35)(0.35)(0.35) = 0.042875$$

Observational studies

- With conditioning on 1 treated unit, model in (1) and def. of conditional probability yield

$$\Pr\left(\boldsymbol{Z}_s = \begin{bmatrix}0 & 0 & 1\end{bmatrix}^\top\right) = 0.147875 \,/\, [(0.147875)(3)] = 1/3$$
$$\Pr\left(\boldsymbol{Z}_s = \begin{bmatrix}0 & 1 & 0\end{bmatrix}^\top\right) = 0.147875 \,/\, [(0.147875)(3)] = 1/3$$
$$\Pr\left(\boldsymbol{Z}_s = \begin{bmatrix}1 & 0 & 0\end{bmatrix}^\top\right) = 0.147875 \,/\, [(0.147875)(3)] = 1/3$$

- We now have uniform distribution on Ω_s

  I.e., completely randomized experiment within stratum s

Two concerns

- In practice, even under no hidden bias, exact stratification on covariates determining assignment probabilities is difficult or impossible
  - ⋆ Exact stratification on covariates **sufficient**, but not **necessary**, for homogeneous assignment probabilities within strata
  - Common instead to stratify on **estimated** propensity score
- Propensity scores may differ within strata due to unobserved covariates, $\boldsymbol{u}$

We will address both concerns in days on matching and sensitivity analysis

Estimation and testing

Estimation and testing

- Create strata of similar units (in their unobservable propensity scores)
- Then estimate ATE or test hypotheses about effects within blocks
  - Act as-if we have a mini randomized experiment within blocks
- Overall estimate or test then averages over results in all mini-experiments

Example: Estimation and testing

- Rosenbaum (2017, p. 66-70)
  - Treatment vs. control comparison
  - Outcome: mortality
  - Two covariates
    - age (young or old)
    - sex (male or female)
  - Probability of treatment for young men and young women $= 0.2$
  - Probability of treatment for old men and old women $= 0.8$

Table 5.1. A small simulated example, with randomized treatment assignment inside each of four strata, and with no treatment effect

Stratum 1: Older men				
Group	*Dead*	*Alive*	*Total*	*Morality rate (%)*
Treated	31,868	47,960	79,828	39.9
Control	8,132	12,040	20,172	40.3
Stratum 2: Older women				
Group	*Dead*	*Alive*	*Total*	*Morality rate (%)*
Treated	23,983	55,796	79,779	30.1
Control	6,017	14,204	20,221	29.8
Stratum 3: Younger men				
Group	*Dead*	*Alive*	*Total*	*Morality rate (%)*
Treated	3,993	16,028	20,021	19.9
Control	16,007	63,972	79,979	20.0
Stratum 3: Younger women				
Group	*Dead*	*Alive*	*Total*	*Morality rate (%)*
Treated	2,021	17,777	19,798	10.2
Control	7,979	72,223	80,202	9.9

Figure 1: (Rosenbaum, 2017, p. 67)

Example: Estimation and testing

Table 5.2. The four strata from Table 5.1 collapsed, leading to the false impression of a treatment effect

Merged table				
Group	*Dead*	*Alive*	*Total*	*Morality rate (%)*
Treated	61,865	137,561	199,426	31.0
Control	38,135	162,439	200,574	19.0

Figure 2: (Rosenbaum, 2017, p. 69)

- Older people more likely to be treated (prob $= 0.8$)
- Older people also have greater baseline mortality rate

$\Longrightarrow$ treated group composed mainly of older people
and control group composed mainly of younger people

$\Longrightarrow$ appears to be an effect when there is none

References

Cochran, W. G. (1965). The planning of observational studies of human populations. *Journal of the Royal Statistical Society. Series A (General) 128*(2), 234–266.

Fisher, R. A. (1935). *The Design of Experiments*. Edinburgh, SCT: Oliver and Boyd.

Rosenbaum, P. R. (2002). *Observational Studies* (2nd ed.). New York, NY: Springer.

Rosenbaum, P. R. (2017). *Observation and Experiment: An Introduction to Causal Inference*. Cambridge, MA: Harvard University Press.

Rubin, D. B. (1977). Assignment to treatment group on the basis of a covariate. *Journal of Educational Statistics 2*(1), 1–26.

Rubin, D. B. and R. P. Waterman (2006). Estimating the causal effects of marketing interventions using propensity score methodology. *Statistical Science 21*(2), 206–222.