

CompSci 590.6

# Understanding Data: Theory and Applications

## Lecture 17

# Causality in Statistics

Instructor: Sudeepa Roy

Email: [sudeepa@cs.duke.edu](mailto:sudeepa@cs.duke.edu)

Fall 2015

# Today's Reading

## **Rubin**

Journal of the American Statistical Association, 2005

*Causal Inference Using Potential Outcomes: Design, Modeling, Decisions*

## **Rosenbaum-Rubin**

Biometrika, 1983

*The Central Role of the Propensity Score in Observational Studies for Causal Effects*

# Potential Outcome Model

- Referred to as Neyman-Rubin model or Rubin's model
  - First proposed in Neyman's Ph.D. thesis (1923)
  - A model for "Randomized Experiments" by Fisher (1920s-30s)
  - Further developed by Rubin (1978) and others
- Establish a causal relationship between a potential cause (treatment) and its effect (outcome)

# Potential Outcome Model

Widely used in

- **Medicine**
  - Christakis and Iwashyna 2003; Rubin 1997
- **Economics**
  - Abadie and Imbens 2006; Galiani, Gertler, and Schargrotsky 2005; Dehejia and Wahba 2002, 1999
- **Political science**
  - Bowers and Hansen 2005; Imai 2005; Sekhon 2004b
- **Sociology**
  - Morgan and Harding 2006; Diprete and Engelhardt 2004; Winship and Morgan 1999; Smith 1997
- **Law**
  - Rubin 2001

References in [Sekhon 2007]

# Units

- N “units”
  - physical objects at particular points in time
  - e.g. individual people, one person at different points of time, plots of lands

Units	Covariates	Treatment assignment	Potential Outcome: Treatment	Potential Outcome: Control	Unit-level causal effects	Summary of causal effects
1	$X_1$	$T_1$	$Y_{11}$	$Y_{01}$	$Y_{11} - Y_{01}$	$E[Y_1 - Y_0]$
2	$X_2$	$T_2$	$Y_{12}$	$Y_{02}$	$Y_{12} - Y_{02}$	
...						
N	$X_n$	$T_N$	$Y_{1N}$	$Y_{0N}$	$Y_{1N} - Y_{0N}$	

# Treatment and Control

- Each unit  $i$  can be exposed or not to a treatment  $T_i$ 
  - e.g. individuals taking an Aspirin vs. placebo ,
- “Active Treatment” or “Treatment” ( $T_i = 1$ )
  - if exposed
- “Control Treatment” or “Control” ( $T_i = 0$ )
  - if not exposed

Units	Covariates	Treatment assignment	Potential Outcome: Treatment	Potential Outcome: Control	Unit-level causal effects	Summary of causal effects
1	$X_1$	$T_1$	$Y_{11}$	$Y_{01}$	$Y_{11} - Y_{01}$	$E[Y_1 - Y_0]$
2	$X_2$	$T_2$	$Y_{12}$	$Y_{02}$	$Y_{12} - Y_{02}$	
...						
N	$X_n$	$T_N$	$Y_{1N}$	$Y_{0N}$	$Y_{1N} - Y_{0N}$	

# Covariates

- Variables that take their values before the treatment assignment
- Cannot be affected by the treatment
  - e.g. pre-aspirin headache pain, gender, blood-pressure

Units	Covariates	Treatment assignment	Potential Outcome: Treatment	Potential Outcome: Control	Unit-level causal effects	Summary of causal effects
1	$X_1$	$T_1$	$Y_{11}$	$Y_{01}$	$Y_{11} - Y_{01}$	$E[Y_1 - Y_0]$
2	$X_2$	$T_2$	$Y_{12}$	$Y_{02}$	$Y_{12} - Y_{02}$	
...						
N	$X_n$	$T_N$	$Y_{1N}$	$Y_{0N}$	$Y_{1N} - Y_{0N}$	

# Potential Outcome

- $Y_1$  (for treatment,  $T_i = 1$ )
- $Y_0$  (for control,  $T_i = 0$ )
- for  $i$ -th unit :  $Y_{1i}$  and  $Y_{0i}$
- **Observed outcome**  $Y = T_i Y_{1i} + (1 - T_i) Y_{0i}$

Units	Covariates	Treatment assignment	Potential Outcome: Treatment	Potential Outcome: Control	Unit-level causal effects	Summary of causal effects
1	$X_1$	$T_1$	$Y_{11}$	$Y_{01}$	$Y_{11} - Y_{01}$	$E[Y_1 - Y_0]$
2	$X_2$	$T_2$	$Y_{12}$	$Y_{02}$	$Y_{12} - Y_{02}$	
...						
N	$X_n$	$T_N$	$Y_{1N}$	$Y_{0N}$	$Y_{1N} - Y_{0N}$	



# Unit-level causal effect

- The comparisons of  $Y_{1i}$  and  $Y_{0i}$ 
  - difference or ratio
  - Typically  $Y_{1i} - Y_{0i}$
- For any unit  $i$ , only one of them can be observed
  - we cannot go back in time and expose it to the other treatment
- **Fundamental problem of causal inference**

Units	Covariates	Treatment assignment	Potential Outcome: Treatment	Potential Outcome: Control	Unit-level causal effects	Summary of causal effects
1	$X_1$	$T_1$	$Y_{11}$	$Y_{01}$	$Y_{11} - Y_{01}$	$E[Y_1 - Y_0]$
2	$X_2$	$T_2$	$Y_{12}$	$Y_{02}$	$Y_{12} - Y_{02}$	
...						
N	$X_n$	$T_N$	$Y_{1N}$	$Y_{0N}$	$Y_{1N} - Y_{0N}$	

# Summary of causal effect

- Defined for a collection of units
- e.g.
  - the mean (or expected) unit-level causal effect -- standard
  - the median unit-level causal effect for all males
  - the difference between the median  $Y_{1i}$  and  $Y_{0i}$  for all females

Units	Covariates	Treatment assignment	Potential Outcome: Treatment	Potential Outcome: Control	Unit-level causal effects	Summary of causal effects
1	$X_1$	$T_1$	$Y_{11}$	$Y_{01}$	$Y_{11} - Y_{01}$	$E[Y_1 - Y_0]$
2	$X_2$	$T_2$	$Y_{12}$	$Y_{02}$	$Y_{12} - Y_{02}$	
...						
N	$X_n$	$T_N$	$Y_{1N}$	$Y_{0N}$	$Y_{1N} - Y_{0N}$	

# Remark..

- To be a causal effect, the comparisons of  $Y_1$  and  $Y_0$  should be for a common set of units
  - e.g. females
  - we cannot apply control to males and treatment to females

Units	Covariates	Treatment assignment	Potential Outcome: Treatment	Potential Outcome: Control	Unit-level causal effects	Summary of causal effects
1	$X_1$	$T_1$	$Y_{11}$	$Y_{01}$	$Y_{11} - Y_{01}$	$E[Y_1 - Y_0]$
2	$X_2$	$T_2$	$Y_{12}$	$Y_{02}$	$Y_{12} - Y_{02}$	
...						
N	$X_n$	$T_N$	$Y_{1N}$	$Y_{0N}$	$Y_{1N} - Y_{0N}$	

# Average Treatment Effect (ATE)

- $ATE = E[Y_1 - Y_0]$
- Recall observed outcome  $Y = T Y_1 + (1-T) Y_0$
- Suppose Treatment Assignment ( $T$ ) is independent of  $Y_1, Y_0$
- Then

$$\begin{aligned} & E[Y_1 - Y_0] \\ = & E[Y_1] - E[Y_0] \\ = & E[Y_1 \mid T = 1] - E[Y_0 \mid T = 0] \\ = & E[Y \mid T = 1] - E[Y \mid T = 0] \end{aligned}$$

- e.g. in a Randomized Experiment (Fisher 1920-30), when each unit is randomly assigned to a Treatment or Control Group
- Still need additional assumptions

# SUTVA

## Stable Unit Treatment Value Assumptions

- Cox 1958, Rubin 1978

### 1. No “interference” or “spill-over effect” among units

- For unit  $i$ ,  $Y_{1i}$  and  $Y_{10}$  are NOT affected by what action any other unit  $j$  received

### 2. Unique Treatment Level or “Dose”

- There are no hidden versions of treatments
- No matter how (mechanism) unit  $i$  received treatment 1, the outcome that would be observed would be  $Y_{1i}$  -- similarly for treatment 0

# Violations of SUTVA

## 1. No interference

- (wiki) Two units Joe and Mary for effect of a drug for high blood pressure
- They share the same household
- Mary cooks
- Mary got drug (treatment) – her pressure reduces – cooks salty food
  - In practice, Mary may not know if she got the drug or placebo
- Joe's pressure increases

## 2. Unique Treatment Level or “Dose”

- Different doses of the medicine for drug pressure

# More assumptions

- Compliance issue
  - People assigned to treatment may refuse it
  - People assigned to control may try to get treatment
    - Barnard, Frangakis, Hill, and Rubin 2003
  - People started taking a medicine, then stopped in the middle because it made them too sick to work

# Notes on Neyman-Rubin Model

- At least half of the potential outcomes are missing
  - Still it is important to explicitly represent both potential outcomes
  - Considered to be a significant contribution by Neyman (Rubin 2005)
- Assumptions are critical
  - without them the causal inferences are meaningless



# The Power of Randomized Experiments

## Recall

- Covariates ( $X$ ) represent the set of variables that take their values before the assignment of the units into treatment or control groups
  - e.g., the gender of a human subject
  - cannot be affected by treatments
- What do we get by randomly assigning units to treatment/control groups?

# The Power of Randomized Experiments

- The assigned treatment is statistically independent of any (measured or unmeasured) covariate in the population before the experiment has been started
  - The distribution of any covariate is the same in the treatment and control groups
- Any difference in outcomes is due to the treatment and not any other pre-existing differences
- The average of control/treatment group outcomes is an unbiased estimate of average outcome under control/treatment for whole population
  - $ATE = E[Y_1 - Y_0] = E[Y | T = 1] - E[Y | T = 0]$

# But, Randomized Experiments are not always feasible

## 1. Infeasibility or high cost

- e.g., how allocation of government funding in different research areas will affect the number of academic jobs in these areas

## 2. Ethical reasons

- e.g., effect of availability to better resources during childhood on higher education in the future

## 3. Prohibitive delay

- e.g., effect of childhood cholesterol on teen obesity)

## 4. In some scenarios randomization may not estimate effects for the groups we are interested in

## 5. Experiments can be on a small population, may have a large variance

# Observational Study

- Alternative to true randomized experiments
  - Tries to simulate the ideal situation
- Create treatment and control groups that appear to be random
  - at least on **observed/measured** variables by choosing individuals with similar covariate values
  - do not use the outcome while selecting the groups

# Balancing Scores

- A balancing score  $b(X)$  is a function of the observed covariates  $X$  such that
  - the conditional distributions of  $X$  given  $b(X)$  are the same on the treatment ( $T = 1$ ) and the control groups ( $T = 0$ ), i.e.,
    - $X \perp T \mid b(X)$
- Example:  $b(X) = X$ 
  - The finest balancing score
- Propensity score  $e(X)$ 
  - The coarsest balancing score

# Propensity-Score Methods

- Make coarse (bigger) groups
  - May not match on all measured covariates
  - But the distributions of covariates are the same for treatment and control
- Cannot say anything about unmeasured/unobserved covariates

# Propensity Score

- The conditional probability of assignment to treatment given the covariates
  - $e(X) = \Pr(T = 1 \mid X)$
- Known for Randomized Experiments
- Not known for Observational Study

# Strongly Ignorable Treatment Assignment

- Treatment assignment is “strongly ignorable given a vector of covariates  $V$ ” if for all  $V$ 
  1.  $(Y_1, Y_0) \perp T \mid V$
  2.  $0 < \Pr[T = 1 \mid V] < 1$
- Simply “strongly ignorable” when  $V = X$

[Rosenbaum-Rubin 1983]

1. If treatment assignment is strongly ignorable given  $X$ , then it is strongly ignorable given any balancing score  $b(X)$
2. For any function  $b(X)$  of  $X$ ,  $b(X)$  is a balancing score if and only if  $e(X) = f(b(X))$  for some function  $f$ 
  - In particular,  $X \perp T \mid e(X)$



# ATE in Observational Study

- Recall,  $ATE = E[Y_1 - Y_0]$
- Consider a two-phase sampling approach
  1. Suppose a specific value of the vector of covariates  $X = x$  is randomly sampled from the entire population (both treated and control groups)
  2. Then a treated and a control units are sampled with this value  $X = x$
- The expected difference in response is

$$E_X[ E[Y_1 | X, T = 1] - E[Y_0 | X, T = 0] ]$$

- If the treatment assignment is strongly ignorable, then

$$\begin{aligned} & E_X[ E[Y_1 | X, T = 1] - E[Y_0 | X, T = 0] ] \\ = & E_X[ E[Y_1 | X] - E[Y_0 | X] ] \\ = & E[Y_1 - Y_0] \quad (\text{why?}) \end{aligned}$$

- **Challenge:** Too many (measured) covariates, individual groups will be too sparse

# Three methods for using balancing score on observational data

1. Pair matching on balancing scores
2. Sub-classification on balancing scores
3. Covariance adjustment on balancing scores

# Pair-matching on balancing score

- Sample  $b(X)$  at random
- Then sample one treated and one control units with this value of  $b(X)$ .
- The expected difference in response equals the ATE at this  $b(X)$ 
  - the mean of matches pair differences in this two-step process is an unbiased estimator of the ATE

# Sub-classification on balancing scores

- Sample a group of units using  $b(X)$  such that
  - $b(X)$  is constant for all units in this group
  - at least one unit in the group received each treatment ( $T = 1, 0$ ).
- The expected difference in treatment means equals the ATE at this  $b(X)$ 
  - the weighted average of such differences (weight = fraction of population at  $b(X)$ ) is an unbiased estimator of the ATE.

# Covariance adjustment on balancing scores

- Assumes that the conditional expectation of  $Y_t$  given  $b(X)$  is linear
  - $E[Y_t \mid b(X), S = t] = \alpha_t + \beta_t b(X)$  for  $t = 0, 1$
- Gives an unbiased estimator of the treatment effect at  $b(X) = E[Y_1 - Y_0 \mid b(X)]$  in terms of unbiased estimators of  $\alpha_1, \beta_1, \alpha_0, \beta_0$

# Neyman-Rubin vs. Pearl's Model

- Potential Outcome (Neyman-Rubin) = Counterfactuals (Pearl)
- Treatment (Neyman-Rubin)  $\approx$  intervention (Pearl)
- Structural causal graph on variables assumed by Pearl
  - Causal inference is on (variable-value) pairs
- No causal structure assumed in Neyman-Rubin's model
  - Infers causal relationships by experiments or from evidence

*“Some authors (e.g., Greenland, Pearl, and Robins 1999; Dawid 2000) call the potential outcomes “counterfactuals,” borrowing the term from philosophy (e.g., Lewis 1973). 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”*

**-- Rubin' 2005**

## Other References

1. [Holland 1986]: Holland, Paul W. 1986. “Statistics and Causal Inference.” *Journal of the American Statistical Association* 81(396): 945-960
2. [Sekhon 2007]: Sekhon, Jasjeet S. 2007. “The Neyman-Rubin Model of Causal Inference and Estimation via Matching Methods”

## Next Topic:

- Exploring Data with Humans in the Loop