Biost 536 / Epi 536
Categorical Data Analysis in Epidemiology

Scott S. Emerson, M.D., Ph.D.
Professor of Biostatistics
University of Washington

Lecture 16:
Causal Inference

December 3, 2013

1

Lecture Outline

- Causal Estimands
- Analysis Methods
- (Next: Propensity Scores)

2

Causal Estimands

3

Common Questions

- Clustering
  - Individuals
  - Variables
- Associations between “risk factors” and “events”
  - “Markers” of incident events
  - “Causes” of incident events
- Prediction of events
  - Estimation of tendencies (means, medians, …)
  - Estimation of ranges (necessary tolerance, …)
  - Classification (diagnosis, prognosis)

4

My Claim

- Ultimate interest is almost always understanding cause and effect relationships
- Clustering is something we do at the very beginning of investigations
- "Associative" analysis for markers is merely recognition that we cannot infer cause and effect from observational data
  - Often a logical step on the way to investigating cause and effect
  - Sometimes all that is feasible
- "Predictive" models are most often stop-gap solutions that will work in the short term
  - If we understood the cause and effect, we would use that information

5

Causation

- A term that is at best imprecisely defined
- When considering a single "risk factor" we might consider whether it is
  - "necessary" for the event to occur
  - "sufficient" for the event to occur
- Most often, it is neither
  - The incidence of the event will depend on other variables as well
    - At least with respect to our current level of understanding
- Nonetheless, we try to refine what we mean within a context that allows us to define three broad levels of estimands

6

Three Levels of Estimands

- Effect of exposure on the population
- Effect of exposure on some subpopulation
- Effect of exposure within an individual

7

Causal Effect: Definition

- We consider a counterfactual setting in which we could know outcomes for each individual when exposed and unexposed

u Index (covariate vector) for unit of treatment

$Y_E(u)$ Outcome for unit u in presence of exposure

$Y_{\bar{E}}(u)$ Outcome for unit u in absence of exposure

Causal effect of exposure for unit u

$$CE(u) = Y_E(u) - Y_{\bar{E}}(u)$$

8

Causal Effect: Comments

- The effect within an individual
- If we consider the outcome to be deterministic, we presume that *u* captures all pertinent variables that identify the sampling unit and its determinants of outcome
  - Essentially a random effect measuring all pertinent covariates at the time of the "treatment" (exposure)
- Of course, it is not possible to observe the outcome for both the exposed and unexposed setting for any unit
  - Time (at the very least) marches on
- And it is probably not really of interest to condition on every single covariate
  - We most often find it useful to average over covariates that we do not yet understand enough to measure

9

Indiv Average Causal Effect: Definition

- We consider a counterfactual setting in which we could know a distribution of outcomes for each individual when exposed and unexposed that may depend on time-varying covariates

u Index (time invariant covariate vector) for unit of treatment

$\vec{V}$ Time - varying covariate vector for unit of treatment

$Y_E\left(u \mid \vec{V}\right)$ Outcome for unit *u* in presence of exposure

$Y_{\bar{E}}\left(u \mid \vec{V}\right)$ Outcome for unit *u* in absence of exposure

Average causal effect of exposure for unit *u*

$$ACE(u) = E_{\vec{V}|u}\left[Y_E\left(u \mid \vec{V}\right) - Y_{\bar{E}}\left(u \mid \vec{V}\right)\right]$$

10

Indiv Average Causal Effect: Comments

- The average effect within an individual
  - Averaged over any effect modification by *V*
- We now model the possibility that there might be a distribution of settings in which the "treatment" (exposure) would be administered to an individual
- For any arbitrary auxiliary covariate vector *V*, it is impossible to measure the outcome in both the exposed and unexposed settings
- However, depending on how *V* is defined, we may through repeated experiments on *u* estimate *ACE(u)*
  - We must include time itself as a covariate that we do not want to hold constant
  - We must also presume no time trend that would confound or modify our estimate of the desired estimand

11

Average Causal Effect: Definition

- We average the individual average causal effect across the population

u Index (time invariant covariate vector) for unit of treatment

$\vec{V}$ Time - varying covariate vector for unit of treatment

$Y_E\left(u \mid \vec{V}\right)$ Outcome for unit *u* in presence of exposure

$Y_{\bar{E}}\left(u \mid \vec{V}\right)$ Outcome for unit *u* in absence of exposure

Average causal effect of exposure

$$ACE = E_U\left\{E_{\vec{V}|u}\left[Y_E\left(u \mid \vec{V}\right) - Y_{\bar{E}}\left(u \mid \vec{V}\right)\right]\right\}$$

12

Average Causal Effect: Comments

- The average individual effect within some population
  - May or may not consider variable effect within an individual
- Again, depending on how *V* is defined, we may through repeated experiments on *u* estimate *ACE(u)*
  - Because we are using expectation, we can consider varying exposure randomly across individuals

13

Population Average Causal Effect: Defn

- We consider a counterfactual setting in which we could know outcomes for a population of individuals when exposed and unexposed

u Index for unit of treatment having some defined distribution within a population of interest

$Y_E(u \mid \vec{V})$ Outcome for unit u in presence of exposure

$Y_{\bar{E}}(u \mid \vec{V})$ Outcome for unit u in absence of exposure

Population average causal effect of exposure

$$PACE = E_U[Y_E(u)] - E_U[Y_{\bar{E}}(u)]$$

14

PACE: Comments

- This measures the average effect in the population
- Note that because we are using expectation, we expect the *ACE* and *PACE* to be equal

Average causal effect of exposure

$$ACE = E_U[Y_E(u) - Y_{\bar{E}}(u)]$$

Population average causal effect of exposure

$$PACE = E_U[Y_E(u)] - E_U[Y_{\bar{E}}(u)]$$

15

Subpopulations

- We can of course define an estimand similar to *PACE* for each of several subpopulations
  - We can consider causal effect in each subgroup separately, or
  - We can consider some average of the separate subgroup causal effects

Population average causal effect of exposure in subgroup

$$PACE(\vec{x}) = E_U[Y_E(u)] - E_U[Y_{\bar{E}}(u)]$$

Average subpopulation causal effect of exposure in subgroup

$$ASCE(\vec{x}) = \sum_{\vec{x}} w_{\vec{x}} PACE(\vec{x})$$

16

Modifications to Causal Estimands

- In the previous measures, we used means and differences of means to characterize the causal effect
- With binary outcomes, we have also considered
  - Ratios of means, which are equivalent to differences of log means
  - Odds ratios, which are equivalent to differences of log odds

17

Modified CE, ACE, PACE

- We consider alternatives to the expectation
  - Note that transformations of binary exposures may not make sense
  - And not that *mACE* and *mPACE* may not be equal

$f(\ )$ is some alternative transformation (e.g., log or logit)

Causal effect of exposure for unit u

$$m\text{CE}(u) = f(Y_E(u)) - f(Y_{\bar{E}}(u))$$

Average causal effect of exposure for unit u

$$mACE = E_U\left[f(Y_E(u)) - f(Y_{\bar{E}}(u))\right]$$

Population average causal effect of exposure for unit u

$$mPACE = f(E_U[Y_E(u)]) - f(E_U[Y_{\bar{E}}(u)])$$

18

Analysis Methods

19

Impact on Analysis Models

- When considering the analysis models used for each of the estimands, we must consider the role of:
- Matching in design
- Variable adjustment
  - Effect modification, confounding, precision
- Measure of association
  - Risk difference, risk ratio, odds ratio

20

Available Tools

- Regression models
  - Fixed effects
  - Random effects
- Matched analyses
  - Fixed effects
  - Random effects
- Stratified analyses using weighted averages

21

Population Estimands: RR and RD

- In absence of confounding we would treat as a two sample problem
- In presence of confounding we can
  - Estimate effects within strata
  - Combine strata according to some “standard” population
- RD and RR are collapsible
  - Hence, adjusting for covariates should give an estimate of a common RD or RR across all covariate defined groups
  - In the presences of effect modification, however, we still need to use weighted combinations to recreate the population

22

Population Estimands: OR

- In absence of confounding we would treat as a two sample problem
- In presence of confounding we can
  - Estimate effects within strata
  - Combine strata according to some “standard” population
    - Even with a common OR in all subgroups, we will obtain different standardized OR depending upon the population used.
- OR is not collapsible
  - We must recognize that adjusted / matched analyses are not estimating the same OR as desired for mPACE
    - Hence either logistic or conditional logistic regression would have to use weighted averages
    - In case-control studies, we may not have the information about the distribution of subjects in the entire population

23

Example

- Hypothetical case-control study: Original data

	Cases			Controls			OR
Age	**Exp**	**Unexp**	**P(stratum)**	**Exp**	**Unexp**	**P(stratum)**	
50-55	77	9	0.04	395	33	0.21	**1.40**
55-60	102	28	0.06	418	76	0.25	**1.51**
60-65	121	73	0.10	277	113	0.19	**1.48**
65-70	119	161	0.14	147	133	0.14	**1.50**
70-75	98	292	0.19	64	130	0.10	**1.47**
75-80	64	430	0.25	24	106	0.06	**1.52**
80-85	27	401	0.21	8	78	0.04	**1.52**
Total	608	1394		1333	669		**4.57**

24

Example

- Hypothetical case-control study: Standardization 1

	Cases			Controls			OR
Age	**Exp**	**Unexp**	**P(stratum)**	**Exp**	**Unexp**	**P(stratum)**	
50-55	18	2	0.01	18	2	0.01	**1.40**
55-60	31	9	0.02	34	6	0.02	**1.51**
60-65	37	23	0.03	43	17	0.03	**1.48**
65-70	51	69	0.06	63	57	0.06	**1.50**
70-75	60	180	0.12	79	161	0.12	**1.47**
75-80	65	436	0.25	92	408	0.25	**1.52**
80-85	63	938	0.50	93	908	0.50	**1.52**
Total	326	1656		423	1559		**1.38**

25

Example

- Hypothetical case-control study: Standardization 2

	Cases			Controls			OR
Age	**Exp**	**Unexp**	**P(stratum)**	**Exp**	**Unexp**	**P(stratum)**	
50-55	256	30	0.14	264	22	0.14	**1.40**
55-60	224	62	0.14	242	44	0.14	**1.51**
60-65	178	108	0.14	203	83	0.14	**1.48**
65-70	122	164	0.14	150	136	0.14	**1.50**
70-75	72	214	0.14	94	192	0.14	**1.47**
75-80	37	249	0.14	53	233	0.14	**1.52**
80-85	18	268	0.14	27	259	0.14	**1.52**
Total	907	1095		1033	969		**1.29**

26

Final Comments

- It should be apparent that we can never truly estimate a “within individual” effect
- Sometimes we can get closer than others
  - Matching, adjusting for random effects
- It should also be apparent that we rarely truly estimate a “population” effect
  - We almost always have restrictions placed on our sampling scheme
- The major point that should be made is:
  - When using OR, recognize that variations in the adjustment variables and the way they are modeled can mean that precision variables will change our estimand

27

Propensity Scores

28

Confounding: Definition

- The association between a predictor of interest and the response variable is confounded by a third variable if
  - The third variable is associated with the response
    - causally (in truth)
    - in groups that are homogeneous with respect to the predictor of interest, and
    - not in the causal pathway of interest,

    AND

    - The third variable is associated with the predictor of interest in the sample.

29

Interpretation of Adjusted Parameters

- Intercept
  - Corresponds to a population with all modeled covariates equal to zero
  - Most often outside range of data; quite often impossible; very rarely of interest by itself
- Slope
  - A comparison between groups differing by 1 unit in corresponding covariate, but agreeing on all other modeled covariates
    - Sometimes impossible to use this definition when modeling interactions or complex curves
- Adjustment thus handles confounding by modeling the association between the confounder and the response

30

Propensity Scores

- Motivation: Handle confounding by modeling the association between the confounders and the POI
  - An analogy to RCT: Randomization ensures independence
  - We model the propensity for individuals to receive the "treatment" (or have the exposure)
- In RCT all subjects have same probability of assignment to arms
- In an observational study, subjects with different covariate values might have different "randomization ratios" (propensity scores)
  - But we assume that all subjects with similar covariates have the same propensity to exposure

31

General Methods

- Fit a model using
  - Exposure as response variable
  - All possible covariates as potential predictors
- Analysis models
  - Logistic regression
  - Other predictive models for a probability
- Model building for propensity score can be highly adaptive
- Take estimated probability of "assignment" to exposure as an adjustment predictor in model of outcome vs POI
  - Usually use very coarse grouping of propensity scores

32

Impact

- We can show that performing analyses adjusted for similar propensity scores guarantees that within each stratum the distribution of covariates is equivalent for exposed and unexposed
  - Hence, after adjustment for propensity scores all covariates are precision variables
- HOWEVER: Above is true providing we have correctly modeled all covariates that affect the propensity to treatment
  - All covariates
  - Highly flexible model
    - Overfitting is not an issue from bias standpoint
    - But we are just estimating the propensity, so need to worry somewhat about variability

33

Advantages / Disadvantages

- Relative to adjustment for covariates
  - We end up just adjusting for one propensity score
  - May better handle sparse data
    - We just have to consider about matched propensity, the matched distribution of covariates is then addressed
  - May better address population level inference
    - But there will be some adjustment for the most dominant variables in the propensity score
  - Exposure must be binary
  - OR coverage probabilities may be poor

34

Evaluation

- Cepeda et al. (2003) studied performance of the propensity score incohort studies with moderately rare events. Comparing covariate adjustment for all confounders in logistic regression to dummy variable adjustment for quintiles of the propensity score they found:
- In the settings they considered, when there were seven or fewer events per confounder, propensity score adjustment yielded OR estimates with less bias and more precision than ordinary confounder adjustment.
- With more events per confounder, ORs based on propensity score adjustment were more biased.

35

Evaluation (cont)

- Estimates based on covariate adjustment were biased by mis-specification of the model for association between a confounder and disease, but not biased by mis-specificationof the model for the association between confounder and exposure in propensity score adjustment.
- Tests based on propensity score adjustment had greater power than tests based on covariate adjustment. (Test level wasn't evaluated.)
- Estimates based on logistic regression covariate adjustment had much worse precision than estimates based on propensity score adjustment when there were fewer than eight outcome events per confounder.

36

Evaluation (cont)

- Confidence intervals based on propensity score adjustment had much better coverage probability when there were fewer than eight observations per confounder; logistic regression covariate adjustment had much better coverage probability when there were eight or more outcome events per confounder.
- With enough events (8 per confounder) covariate adjustment was better.

37

My View

- Other attempts at evaluation also appear in the literature, with similarly mixed reviews
- Propensity scores appear attractive, but do not really seem to work well enough, if you ask me

38