500 Class 6 Slides

github.com/THOMASELOVE/2020-500

2020-03-05

Sensitivity Analysis for Matched Samples

The Role of Assumptions

Scenario	Analytic Goal	Role of Assumptions
Randomized Experiments	Testing H_0_: No effect	None
Randomized Experiments	Estimating effects, CIs	Minor
Observational Studies	Anything at all	**Massive**

- **All** observational studies are potentially affected by hidden bias.
  - Sensitivity analyses are needed in any such study.

Stability Analyses

- Are our conclusions mere artifacts of analytic decisions, or are they stable over analyses that differ in (apparently) innocuous ways?
- Examine a series of discrete decisions, hoping conclusions mostly don't change if the decision is changed.

Sensitivity Analyses

- How much do plausible changes in assumptions change conclusions?
- How much hidden bias would have to be present to alter the study's conclusions?

We want to assess the potential for **unmeasured** covariates to change our conclusions.

- To maximize trouble, the unmeasured covariate must be independent of the variables in our propensity model.
  - We missed a dimension of the problem, or our measure is terribly weak.
  - This is part of the motivation to be inclusive in building the PS model.

Idealized Standards for a Sensitivity Evaluation

- Logic, Theory and Empirical Evidence
- "It is unlikely that a non-huge hidden bias would substantially change our conclusions"
  - Measured and incorporated every major known factor that we could identify.
  - Effects on health outcomes were generally large, consistent with earlier work and clinically plausible.

An omitted variable is most likely to change our conclusions about the exposure if it is

- closely related to the outcome,
- seriously imbalanced by exposure,
- uncorrelated with the propensity score.

Does PS Matching Balance "Omitted" Covariates?

No.

- We fit[1] a published PS model to data from the RHC study, using 82 covariates.
- Then we obtained data on 17 other covariates, not included in the PS model.

Corr. with PS	Covariates	Balance Improved	Median Bias Reduction
Sig. ($p < .05$)	10	9 (90%)	45%
Not Sig.	7	2 (29%)	-36%

[1] Love et al 2003 abstract

Sensitivity Analysis Approach

- When describing possible hidden bias, we refer to characteristics we did not observe, and therefore did not control for in PS.
- If our study was randomized, or somehow free of hidden bias, we would have strong evidence of a treatment effect.
- To explain away the observed effect, an unobserved covariate would need to increase the odds of exposure to treatment and the odds of outcome by at least a factor of Γ.

The Role of Sensitivity Analysis

- Cameron and Pauling (1976, 1978) concluded Vitamin C increased survival time of colon cancer patients.

The Role of Sensitivity Analysis

- Cameron and Pauling (1976, 1978) concluded Vitamin C increased survival time of colon cancer patients.
- The result was not sensitive to an unmeasured binary covariate which led to a 10-fold increase in odds of exposure to vitamin C and was a perfect predictor of the outcome[2].

[2] See Rosenbaum 2002, 2017

The Role of Sensitivity Analysis

- Cameron and Pauling (1976, 1978) concluded Vitamin C increased survival time of colon cancer patients.
- The result was not sensitive to an unmeasured binary covariate which led to a 10-fold increase in odds of exposure to vitamin C and was a perfect predictor of the outcome[2].
- Sensitivity analysis looks great, yet the findings were contradicted in a Mayo Clinic RCT.

[2] See Rosenbaum 2002, 2017

The Role of Sensitivity Analysis

- Cameron and Pauling (1976, 1978) concluded Vitamin C increased survival time of colon cancer patients.
- The result was not sensitive to an unmeasured binary covariate which led to a 10-fold increase in odds of exposure to vitamin C and was a perfect predictor of the outcome[2].
- Sensitivity analysis looks great, yet the findings were contradicted in a Mayo Clinic RCT.
- Conclusion: Sensitivity analyses cannot indicate what biases are present, it can only indicate the magnitude needed to alter the conclusion.

[2] See Rosenbaum 2002, 2017

Surgery vs. Medicine for Coronary Artery Disease

CAD: Surgery or Medicine

Coronary bypass surgery or medical/drug therapy for CAD[3]?

- 1,515 subjects
  - 590 surgical patients (39%), the rest medical
- PS included 74 observed covariates
  - Hemodynamic, Angiographic, Lab and Exercise Test Results
  - Patient histories, and demographics
- Stratified into PS quintiles, then combined estimates of Pr(sustained improvement at 6 mo).

[3] See Rosenbaum and Rubin 1983, 1984

Surgery vs. Medicine Study: Results + Sensitivity

Substantial Improvement	Prob (SE)
Medical	0.359 (0.042)
Surgical	0.669 (0.059)

Conclusion: Pr(improved | Surgery) > Pr(improved | medicine)

- A hypothetical unobserved binary covariate would have to **more than triple** the odds of surgery and **more than triple** the odds of improvement, before altering the conclusion[4].

[4] See Rosenbaum and Rubin 1983, 1984

Goal of a Formal Sensitivity Analysis

To replace a general qualitative statement that applies in all observational studies. . .

> *"The association we observe between treatment and outcome does not imply causation."*
>
> *"Hidden biases can explain observed associations."*

with a quantitative statement that is specific to what is observed in a particular study

> *"To explain the association seen in this particular study, one would need a hidden bias of this particular magnitude."*

Hidden Bias

- Two units (patients, subjects, whatever) with the same observed covariates have different chances of receiving the exposure.
- If the association is strong, the hidden bias needed to explain it would be large.
- If a study is free of hidden bias (one example: RCT), this means that any two units that appear similar in terms of their observed covariates actually have the same chance of assignment to exposure.

How would inferences about treatment effects be altered by hidden biases of various magnitudes? - How large would these differences have to be to alter the qualitative conclusions of the study?

The Design Sensitivity Parameter Γ

Γ measures degree of departure from a study that is free of hidden bias. A sensitivity analysis will consider possible values of Γ and show how the inference might change.

- Γ describes the odds ratio comparing the odds of being selected for treatment for two units who are similar on all observed covariates.
  - If $\Gamma = 1$, this means that the study is free of hidden bias
    - Subjects with the same observed covariates have the same odds (hence same probability) of exposure.
  - If $\Gamma = 2$, then two units who appear similar, who have the same set of observed covariates $\mathbf{X}$, could differ in their odds of receiving the treatment by as much as a factor of 2, so that one could be twice as likely as the other to receive the exposure.

Relating Γ to Sensitivity Statements

A study is **sensitive** if values of Γ close to 1 could lead to inferences that are very different from those obtained assuming the study is free of hidden bias.

- A study is *insensitive* if extreme values of Γ are required to alter the inference.

  "To attribute the (observed significant) outcome to an unobserved covariate rather than to the treatment, that unobserved covariate has to increase the odds of treatment by a factor of Γ*, and also predict our outcome quite well."*

Using a Spreadsheet

Getting the Message Across

Straightforward, Spreadsheet-Based Formal Sensitivity Analysis for Matched Samples

- Separate tabs for Outcomes: Binary, Continuous, Survival (w/censoring)
- All calculations based on base case formulas using sign-score tests as described in Rosenbaum 2002 [some nuances ignored (dealing with ties, etc.)]
- Available documents on our Data and Code page describe three examples

Demonstration: A Binary Outcome

- Exposure: Heavy Smoker vs. Non-Smoker
- Outcome: death due to lung cancer (no censoring)

Suppose we paired 1042 heavy smokers to 1042 nonsmokers on the basis of a series of baseline characteristics. Resulting data table on **the 1042 PAIRS** is (LCD = dies from lung cancer)...

–	Heavy Smoker LCD	Heavy Smoker No LCD	Total
Nonsmoker LCD	175	54	**229**
Nonsmoker No LCD	110	703	**813**
Total	**285**	**757**	**1042**

McNemar Test (if this was randomized)

```
test1 <- matrix(c(175, 54, 110, 703), nrow = 2)
mcnemar.test(test1, correct = F)
```

```
    McNemar's Chi-squared test

data:  test1
McNemar's chi-squared = 19.122, df = 1, p-value =
1.226e-05
```

This is an appropriate result if $\Gamma = 1$, but how much hidden bias would need to be present to change this conclusion?

How much hidden bias is needed?

- Pairs = 1042 matched pairs overall
- D = 164 discordant pairs (exactly one member had LCD)
- T = 110 discordant pairs in which Heavy Smoker had LCD

We find the binomial probability of obtaining a value of T = 110 or higher assuming a binomial distribution with D = 164 trials and common probability = p^+ and p^- for the upper and lower bounds. PDF specifies the details.

Spreadsheet Demonstration

	A	B	C	D	E	F	G
1	**Sensitivity Analysis for McNemar's Test: Simplied Formula**						
2	**Section 4.3.2. of Rosenbaum PR (2002) Observational Studies, 2nd Edition.**						
3	**INSERT VALUES (IN RED) IN CELLS HIGHLIGHTED IN YELLOW.**						
4	**Two-By-Two Table**	**Treated, outcome = Yes**	**Treated, outcome = No**				
5	**Control, outcome = Yes**	**175**	**54**	**229**			
6	**Control, outcome = No**	**110**	**703**	**813**			
7		**285**	**757**	**1042**			
8							
9	**Computed Summaries**						
10	# of Pairs	**1042**	# of matched pairs (overall)				
11	# of Discordant Pairs	**164**	# of matched pairs in which exactly one has the outcome				
12	Test Statistic	**110**	# of discordant pairs where Treated has outcome				
13							
14	**Sensitivity Analysis**						
15	**Gamma Values**	2-tail P value (lower bound)	?-tail P value (upper bound)	P-	P+		
16	**1.0**	**0.0000**	**0.0000**	0.500	0.500		
17	**1.5**	**0.0000**	**0.0516**	0.400	0.600		
18	**2.0**	**0.0000**	**0.8538**	0.333	0.667		
19	**2.5**	**0.0000**	**1.7477**	0.286	0.714		
20	**3.0**	**0.0000**	**1.9723**	0.250	0.750		
21	**3.5**	**0.0000**	**1.9979**	0.222	0.778		
22	**4.0**	**0.0000**	**1.9999**	0.200	0.800		
23	**4.5**	**0.0000**	**2.0000**	0.182	0.818		
24	**5.0**	**0.0000**	**2.0000**	0.167	0.833		
25	**5.5**	**0.0000**	**2.0000**	0.154	0.846		
26	**6.0**	**0.0000**	**2.0000**	0.143	0.857		
27							
28	**Insert Gamma Value Below**	2-tail P value (lower bound)	?-tail P value (upper bound)	P-	P+		
29	**1.43**	**0.0000**	**0.0249**	0.412	0.588		

Interpretation of $\Gamma = 1.43$

From Rosenbaum, Chapter 9, we can identify the range of probabilities of treatment assignment:

$$\frac{1}{1+\Gamma} \leq \theta_p \leq \frac{\Gamma}{1+\Gamma}. \text{ Here, } \frac{1}{2.43} = 0.412, \frac{1.43}{2.43} = 0.588$$

So the possible values of treatment assignment are (0.412, 0.588) for $\Gamma = 1.43$.

Demonstration: A Continuous Outcome

- Exposure: Parent working (or not) in a factory where lead was used to make batteries
- Outcome: level of lead found in the child's blood (in ??g/dl of whole blood)

Morton et al. (1982) studied lead in the blood of 33 kids (from different families) whose parents worked in a factory where lead was used in making batteries. They matched each exposed child to one control child of the same age and neighborhood whose parents were employed in other industries not using lead.

The Wilcoxon signed rank statistic for S matched pairs is computed by ranking the absolute value of the differences within each pair from 1 to S, and then summing the ranks of the pairs where the exposed unit had a higher response than the matched control. Value in Morton is 527.

Spreadsheet Demonstration

	A	B	C	D	E	F	G	H	I	J
1	**Sensitivity Analysis for Wilcoxon Signed Rank Test: Simplied Formula**									
2	**Section 4.3.3. of Rosenbaum PR (2002) Observational Studies, 2nd Edition.**									
3	**INSERT VALUES (IN RED) IN CELLS HIGHLIGHTED IN YELLOW.**									
4										
5	**Data**									
6	**Total # of Pairs (S)**	**33**								
7	**Wilcoxon Test Statistic, T**	**527**	sum of ranks of absolute differences for pairs where exposed had higher value than control							
8										
9	**Sensitivity Analysis**									
10	**Gamma Values**	2-tail P value (lower bound)	2-tail P value (upper bound)	P-	P+	E(T-)	E(T+)	SD(T+)		
11	**1.0**	**0.0000**	**0.0000**	0.500	0.500	280.50	280.50	55.97		
12	**1.5**	**0.0000**	**0.0005**	0.400	0.600	224.40	336.60	54.84		
13	**2.0**	**0.0000**	**0.0037**	0.333	0.667	187.00	374.00	52.77		
14	**2.5**	**0.0000**	**0.0125**	0.286	0.714	160.29	400.71	50.57		
15	**3.0**	**0.0000**	**0.0284**	0.250	0.750	140.25	420.75	48.47		
16	**3.5**	**0.0000**	**0.0514**	0.222	0.778	124.67	436.33	46.53		
17	**4.0**	**0.0000**	**0.0807**	0.200	0.800	112.20	448.80	44.77		
18	**4.5**	**0.0000**	**0.1152**	0.182	0.818	102.00	459.00	43.17		
19	**5.0**	**0.0000**	**0.1538**	0.167	0.833	93.50	467.50	41.71		
20	**5.5**	**0.0000**	**0.1952**	0.154	0.846	86.31	474.69	40.39		
21	**6.0**	**0.0000**	**0.2388**	0.143	0.857	80.14	480.86	39.17		
22										
23	**Insert Gamma Value Below**	2-tail P value (lower bound)	2-tail P value (upper bound)	P-	P+	E(T-)	E(T+)	SD(T+)		
24	**2.91**	**0.0000**	**0.0250**	0.256	0.744	143.48	417.52	48.83		
25	**Stop when value for the upper bound of the P value (cell C24) is just below desired two-tailed significance level**									
26										
27										
28										
29										

Demonstration: A Survival Outcome

- Secondary analysis of DIG trial (Ali Ahmed, PI)
- Exposure: Either Normal or Low Serum Potassium Level in the DIG trial (Heart Failure Pts) 1187 matched pairs of "normal potassium" and "low potassium" HF patients with similar baseline characteristics.
- Outcome: All cause mortality during the follow-up period (i.e. there is LOTS of censoring)
  - There are 440 pairs with clear winners
  - In 335 of these 440, winner is normal potassium.

Spreadsheet Demonstration

	A	B	C	D	E	F	G	H
1	Sensitivity Analysis for A Simple Comparison for Censored Survival							
2	Section 4.4.8. of Rosenbaum PR (2002) Observational Studies, 2nd Edition.							
3	INSERT VALUES (IN RED) IN CELLS HIGHLIGHTED IN YELLOW.							
4								
5	Data							
6	Total # of Pairs With A Clear Winner	440						
7	T = # of Pairs Where Exposed Outlives Contr	335						
8								
9	Sensitivity Analysis							
10	Gamma Values	2-tail P value (lower bound)	2-tail P value (upper bound)	P-	P+	E(T-)	E(T+)	SD(T+)
11	1.0	0.0000	0.0000	0.500	0.500	220.00	220.00	10.49
12	1.5	0.0000	0.0000	0.400	0.600	176.00	264.00	10.28
13	2.0	0.0000	0.0000	0.333	0.667	146.67	293.33	9.89
14	2.5	0.0000	0.0288	0.286	0.714	125.71	314.29	9.48
15	3.0	0.0000	0.5820	0.250	0.750	110.00	330.00	9.08
16	3.5	0.0000	1.0000	0.222	0.778	97.78	342.22	8.72
17	4.0	0.0000	1.0000	0.200	0.800	88.00	352.00	8.39
18	4.5	0.0000	1.0000	0.182	0.818	80.00	360.00	8.09
19	5.0	0.0000	1.0000	0.167	0.833	73.33	366.67	7.82
20	5.5	0.0000	1.0000	0.154	0.846	67.69	372.31	7.57
21	6.0	0.0000	1.0000	0.143	0.857	62.86	377.14	7.34
22								
23	Insert Gamma Value Below	2-tail P value (lower bound)	2-tail P value (upper bound)	P-	P+	E(T-)	E(T+)	SD(T+)
24	2.484	0.0000	0.0248	0.287	0.713	126.29	313.71	9.49
25	Stop when value for the upper bound of the P value (cell C24) is just below desired two-tailed significance level							
26								

Communicating the Results

In the absence of hidden bias, a sign-score test for matched data with censoring provides strong evidence ($p < .001$) that low (vs. normal) potassium decreases survival time, even after adjustment via PS matching. To attribute the lower survival time to an unobserved binary covariate unrelated to our propensity model rather than the effect of low potassium, that covariate would need to both

- increase the odds of low potassium more than 2.5-fold and
- be an excellent predictor of mortality.

Estimating Γ in R with `rbounds`

Using `rbounds`: Two Examples

See the `rbounds_demos` document on our web site.

Using rbounds: the toy example

```
psens(match1, Gamma = 5, GammaInc = 0.25)
```

```
 Rosenbaum Sensitivity Test for Wilcoxon Signed Rank P-Value

Unconfounded estimate ....  0

 Gamma Lower bound Upper bound
  1.00           0      0.0000
  1.25           0      0.0000
  1.50           0      0.0003
  1.75           0      0.0030
  2.00           0      0.0160
  2.25           0      0.0524
  2.50           0      0.1232
  2.75           0      0.2286
  3.00           0      0.3574
  3.25           0      0.4927
  3.50           0      0.6190
  3.75           0      0.7265
  4.00           0      0.8114
  4.25           0      0.8745
  4.50           0      0.9190
  4.75           0      0.9491
  5.00           0      0.9688

 Note: Gamma is Odds of Differential Assignment To
```

Summary: Sensitivity Analysis

Hidden bias is the great problem with observational studies, and with PS models.

- Sensitivity analysis after matching can be applied in many scenarios.
- We hope to find that an unobserved covariate would have to be very powerful to alter our conclusions.
- That doesn't mean that such a covariate (or set of them) doesn't exist.

Our PS "Formula" for the Heart Failure papers

1. Identify available data related to selection for the exposure, and to risk for the outcome.
2. Demonstrate need for PS modeling (imbalance in key characteristics), and evaluate PS balance after matching, usually through standardized difference plots (usually significance, too, unfortunately.)
3. Model exposure effect (Cox models stratified by matched pair identifiers, typically.)
4. Formal sensitivity analysis if effect is significant.