Sensitivity Analysis for Observational Studies

November 10, 2010

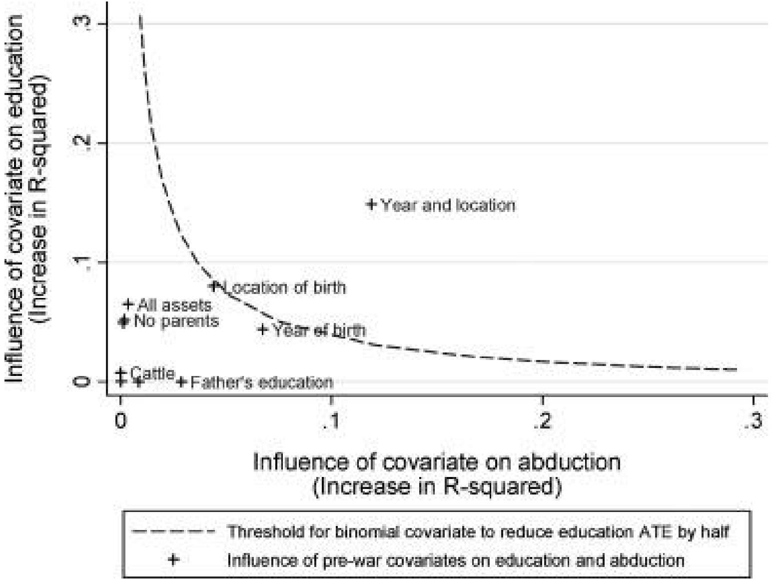
Paul Rosenbaum

An Example

TABLE 4.1. Sensitivity Analysis for Hammond's Study of Smoking and Lung Cancer: Range of Significance Levels for Hidden Biases of Various Magnitudes.

Γ	Minimum	Maximum
1	< 0.0001	< 0.0001
2	< 0.0001	< 0.0001
3	< 0.0001	< 0.0001
4	< 0.0001	0.0036
5	< 0.0001	0.03
6	< 0.0001	0.1

Another Example



An Observational Study

R_1	R_2	Z_1	Z_2	X_1	X_2	π_1	π_2	$\frac{\pi_1}{\pi_1+\pi_2}$	Γ
6	5	1	0	5	5	.293	.293	.5	1
3	7	1	0	46	46	.83	.83	.5	1
4	7	1	0	3	3	.2	.2	.5	1
7	14	1	0	25	25	.44	.44	.5	1

Table: Under the Naive Model

Model of an Observational Study

- M units, each with an observed covariate vector $\mathbf{x}$. Number the M units $j = 1, \ldots, M$, so $\mathbf{x}_{[j]}$ and $Z_{[j]}$ is the covariate and the treatment assignment for the jth unit.

Model of an Observational Study

- M units, each with an observed covariate vector $\mathbf{x}$. Number the M units $j = 1, \ldots, M$, so $\mathbf{x}_{[j]}$ and $Z_{[j]}$ is the covariate and the treatment assignment for the jth unit.
- Unit j is assigned to treatment with probability $\pi_j = \mathrm{prob}(Z_{[j]} = 1)$ and to control with probability $1 - \pi_j = \mathrm{prob}(Z_{[j]} = 0)$

Model of an Observational Study

- M units, each with an observed covariate vector $\mathbf{x}$. Number the M units $j = 1, \ldots, M$, so $\mathbf{x}_{[j]}$ and $Z_{[j]}$ is the covariate and the treatment assignment for the jth unit.
- Unit j is assigned to treatment with probability $\pi_j = \mathrm{prob}(Z_{[j]} = 1)$ and to control with probability $1 - \pi_j = \mathrm{prob}(Z_{[j]} = 0)$
- Treatments are assigned by flipping biased coins (each unit might have a different biased coin):

$$\mathrm{prob}(Z_{[1]} = z_1, \ldots, Z_{[M]} = z_M) = \prod_{j=1}^{M} \pi_{[j]}^{z} \{1 - \pi_{[j]}\}^{1-z_j}$$

Overt Bias

- An observational study is free of *hidden* bias if the π's, though unknown, are known to only depend on the observed covariates, so two units with the same value of $\mathbf{x}$ have the same chance π of receiving the treatment.

- We are assuming **"randomization on the basis of a covariate"**.

Overt Bias

- An observational study is free of *hidden* bias if the π's, though unknown, are known to only depend on the observed covariates, so two units with the same value of $\mathbf{x}$ have the same chance π of receiving the treatment.
- The probability that j will be in treatment is some unknown function of $\mathbf{x}$: $\lambda(\mathbf{x}_{[j]})$, so the probability of treatment assignment becomes:

$$\mathrm{prob}(Z_{[1]} = z_1, \ldots, Z_{[M]} = z_M) = \prod_{j=1}^{M} \lambda(\mathbf{x}_{[j]})^{z} \{1 - \lambda(\mathbf{x}_{[j]})\}^{1-z_j}$$

- We are assuming **"randomization on the basis of a covariate"**.

Probability Distribution of $\mathbf{Z}|\mathbf{m}$

- For inference, Rosenbaum proposes we use the conditional distribution of $\mathbf{Z}$ given $\mathbf{m}$.

Probability Distribution of $\mathbf{Z}|\mathbf{m}$

- For inference, Rosenbaum proposes we use the conditional distribution of $\mathbf{Z}$ given $\mathbf{m}$.
- For randomization inference, we want the full set of possibile treatment assignments (Ω) and their associated probabilities. There are $K = \prod_{s=1}^{S} \binom{n_s}{m_s}$ possible assignments.

Probability Distribution of $\mathbf{Z}|\mathbf{m}$

- For inference, Rosenbaum proposes we use the conditional distribution of $\mathbf{Z}$ given $\mathbf{m}$.
- For randomization inference, we want the full set of possibile treatment assignments (Ω) and their associated probabilities. There are $K = \prod_{s=1}^{S} \binom{n_s}{m_s}$ possible assignments.
- Every treatment assignment $z \in \Omega$ has the same conditional probability: $\frac{1}{K}$, which means we can analyze the data as a uniform randomized experiment.

A Model for Sensitivity Analysis

A sensitivity analysis asks: How would inferences about treatment effects be altered by hidden biases of various magnitudes?

- There is *hidden* bias if two units with the same observed covariates x have differing chances of receiving the treatment, i.e. if $\mathbf{x}_{[j]} = \mathbf{x}_{[k]}$, but $\pi_{[j]} \neq \pi_{[k]}$ for some j and k.

A Model for Sensitivity Analysis

A sensitivity analysis asks: How would inferences about treatment effects be altered by hidden biases of various magnitudes?

- There is *hidden* bias if two units with the same observed covariates x have differing chances of receiving the treatment, i.e. if $\mathbf{x}_{[j]} = \mathbf{x}_{[k]}$, but $\pi_{[j]} \neq \pi_{[k]}$ for some j and k.
- If units j and k are matched into pairs, the odds that units j and k receive the treatment are, respectively, $\pi_{[j]}/(1 - \pi_{[j]})$ and $\pi_{[k]}/(1 - \pi_{[k]})$, and the odds ratio is the ratio of these odds.

A Model for Sensitivity Analysis

A sensitivity analysis asks: How would inferences about treatment effects be altered by hidden biases of various magnitudes?

- There is *hidden* bias if two units with the same observed covariates x have differing chances of receiving the treatment, i.e. if $\mathbf{x}_{[j]} = \mathbf{x}_{[k]}$, but $\pi_{[j]} \neq \pi_{[k]}$ for some j and k.
- If units j and k are matched into pairs, the odds that units j and k receive the treatment are, respectively, $\pi_{[j]}/(1-\pi_{[j]})$ and $\pi_{[k]}/(1-\pi_{[k]})$, and the odds ratio is the ratio of these odds.
- Conditional on the matching procedure, the probability of assignment to treatment:

$$\mathrm{P}(Z_1 = 1 | Z_{s1} + Z_{s2}) = \frac{\pi_{s1}(1-\pi_{s2})}{\pi_{s1}(1-\pi_{s2}) + \pi_{s2}(1-\pi_{s1})}$$

Γ

- The key parameter in a Rosenbaum-style sensitivity analysis is the treatment odds ratio, known as Γ.

Γ

- The key parameter in a Rosenbaum-style sensitivity analysis is the treatment odds ratio, known as Γ.
- In a sensitivity analysis, we ask the effect of varying Γ on our inferences, where Γ is bounded as follows:

$$\frac{1}{\Gamma} \leq \frac{\pi_{[j]}(1 - \pi_{[k]})}{\pi_{[k]}(1 - \pi_{[j]})} \leq \Gamma$$

Γ

- The key parameter in a Rosenbaum-style sensitivity analysis is the treatment odds ratio, known as Γ.
- In a sensitivity analysis, we ask the effect of varying Γ on our inferences, where Γ is bounded as follows:

$$\frac{1}{\Gamma} \leq \frac{\pi_{[j]}(1-\pi_{[k]})}{\pi_{[k]}(1-\pi_{[j]})} \leq \Gamma$$

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

An Alternative Expression: Bias Due to an Unobserved Covariate

- Unit j has an observed covariate $\mathbf{x}_{[j]}$ and an unobserved covariate $u_{[j]}$. The model links the probability of assignment to treatment as follows:

$$\log\left(\frac{\pi_{[j]}}{1-\pi_{[j]}}\right) = k(\mathbf{x}_{[j]}) + \gamma u_{[j]}$$

  with $0 \leq u_{[j]} \leq 1$ and where $k(\cdot)$ is an unknown function and γ is an unknown parameter.

An Alternative Expression: Bias Due to an Unobserved Covariate

- Unit j has an observed covariate $\mathbf{x}_{[j]}$ and an unobserved covariate $u_{[j]}$. The model links the probability of assignment to treatment as follows:

$$\log\left(\frac{\pi_{[j]}}{1-\pi_{[j]}}\right) = k(\mathbf{x}_{[j]}) + \gamma u_{[j]}$$

  with $0 \leq u_{[j]} \leq 1$ and where $k(\cdot)$ is an unknown function and γ is an unknown parameter.
- After adjusting for $\mathbf{x}$, the odds ratio for two units in the same matched pair can be written as:

$$\frac{\pi_{[j]}(1-\pi_{[k]})}{\pi_{[k]}(1-\pi_{[j]})} = \exp\{\gamma(u_{[j]} - u_{[k]})\}$$

Sensitivity of Significance Levels

$$\mathrm{prob}(\mathbf{Z} = \mathbf{z}|\mathbf{m}) = \prod_{j=1}^{M} \left[\frac{e^{\gamma u_{s1}}}{e^{\gamma u_{s1}} + e^{\gamma u_{s2}}} \right]^{z_{s1}} \left[\frac{e^{\gamma u_{s2}}}{e^{\gamma u_{s1}} + e^{\gamma u_{s2}}} \right]^{1-z_{s1}}$$

An Observational Study

R_1	R_2	Z_1	Z_2	X_1	X_2	U_1	U_2	$\frac{\pi_1}{\pi_1+\pi_2}$	Γ
6	5	1	0	5	5	1	0	.6667	2
3	7	1	0	46	46	0	1	.333	2
4	7	1	0	3	3	0	0	.5	2
7	14	1	0	25	25	1	1	.5	2

Table: With an Omitted Variable

Sign-Score Statistics

- General form of a Sign-Score test statistic:

$$T = t(\mathbf{Z}, \mathbf{r}) = \sum_{s=1}^{S} d_s \sum_{i=1}^{2} c_{si} Z_{si}$$

Sign-Score Statistics

- General form of a Sign-Score test statistic:

$$T = t(\mathbf{Z}, \mathbf{r}) = \sum_{s=1}^{S} d_s \sum_{i=1}^{2} c_{si} Z_{si}$$

- Wilcoxon signed rank statistic fo S matched pairs is computed by ranking the absolute differences $|r_{s1} - r_{s2}|$ from 1 to S and summing the ranks of the pairs in which the treated unit had a higher response than the matched control.

Sign-Score Statistics

- General form of a Sign-Score test statistic:

$$T = t(\mathbf{Z}, \mathbf{r}) = \sum_{s=1}^{S} d_s \sum_{i=1}^{2} c_{si} Z_{si}$$

- Wilcoxon signed rank statistic fo S matched pairs is computed by ranking the absolute differences $|r_{s1} - r_{s2}|$ from 1 to S and summing the ranks of the pairs in which the treated unit had a higher response than the matched control.
- d_s is the rank of $|r_{s1} - r_{s2}|$ with average ranks used for ties, and $c_{s1} = 1$, $c_{s2} = 0$ if $|r_{s1} > r_{r2}|$ with average ranks used for ties, and $c_{s1} = 1$, $c_{s2} = 0$ if $r_{s1} > r_{s2}$ or $c_{s1} = 0$, $c_{s2} = 1$ if $r_{s1} < r_{s2}$, and $c_{s1} = 0$, $c_{s2} = 0$ if $r_{s1} = r_{s2}$ (pairs are tied).

Sign Test

- A particularly simple test for matched data is the sign test, which is simply the number of positive (or negative) within match differences
- d_s is 1 for all matched pairs, and $c_{s1} = 1$, $c_{s2} = 0$ if $r_{s1} > r_{r2}$. Similarly, $c_{s1} = 0$, $c_{s2} = 1$ if $r_{s1} < r_{r2}$.

Sign Test

- A particularly simple test for matched data is the sign test, which is simply the number of positive (or negative) within match differences
- d_s is 1 for all matched pairs, and $c_{s1} = 1$, $c_{s2} = 0$ if $r_{s1} > r_{r2}$. Similarly, $c_{s1} = 0$, $c_{s2} = 1$ if $r_{s1} < r_{r2}$.
- Exact p-values can be obtained using the binomial distribution.

The Sign Test

R_1	R_2	D	C_1	C_2	U_1	U_2	$\frac{\pi_1}{\pi_1+\pi_2}$	Γ
6	5	1	1	0	1	0	.6667	2
3	7	1	0	1	0	1	.333	2
4	7	1	0	1	0	0	.5	2
7	14	1	0	1	1	1	.5	2

Table: With an Omitted Variable

Unknowns

R_1	R_2	D	C_1	C_2	U_1	U_2	$\frac{\pi_1}{\pi_1+\pi_2}$
6	5	1	1	0	?	?	?
3	7	1	0	1	?	?	?
4	7	1	0	1	?	?	?
7	14	1	0	1	?	?	?

Table: With an Omitted Variable

Inference

- In a randomized experiment, $t(\mathbf{Z}, \mathbf{r})$ is compared to the randomization distribution under the null hypothesis. In effect, $t(\mathbf{Z}, \mathbf{r})$ is the sum of S independent random variables where the sth variable equals d_s with probability $1/2$.
- If there is hidden bias, we don't know what the randomization distribution is under the null hypothesis! But we can still **bound** the possible distributions under a given amount of possible hidden bias.

Inference with an Unknown Confounder

- For each possible $(\gamma, \mathbf{u})$, the statistic $t(\mathbf{Z}, \mathbf{r})$ is the sum of S independent random variables, where the sth variable equals d_s with probability

$$p_s = \frac{c_{s1}\exp(\gamma u_{s1}) + c_{s2}\exp(\gamma u_{s2})}{\exp(\gamma u_{s1}) + \exp(\gamma u_{s2})}$$

Inference with an Unknown Confounder

- For each possible $(\gamma, \mathbf{u})$, the statistic $t(\mathbf{Z}, \mathbf{r})$ is the sum of S independent random variables, where the sth variable equals d_s with probability

$$p_s = \frac{c_{s1}\exp(\gamma u_{s1}) + c_{s2}\exp(\gamma u_{s2})}{\exp(\gamma u_{s1}) + \exp(\gamma u_{s2})}$$

- With $\Gamma = exp(\gamma)$ define p_s^+ and p_s^- in the following way:

$$p_s^+ = \frac{\Gamma}{1+\Gamma} \tag{1}$$

$$p_s^- = \frac{1}{1+\Gamma} \tag{2}$$

Assume the worse case scenario

R_1	R_2	D	C_1	C_2	U_1	U_2	$\frac{\pi_1}{\pi_1+\pi_2}$
6	5	1	1	0	1	0	?
3	7	1	0	1	0	1	?
4	7	1	0	1	0	1	?
7	14	1	0	1	0	1	?

Table: With an Omitted Variable

Choose a Γ

R_1	R_2	D	C_1	C_2	U_1	U_2	$\frac{\pi_1}{\pi_1+\pi_2}$	Γ
6	5	1	1	0	1	0	.33	2
3	7	1	0	1	0	1	.66	2
4	7	1	0	1	0	1	.66	2
7	14	1	0	1	0	1	.66	2

Table: With an Omitted Variable

Choose another Γ

R_1	R_2	D	C_1	C_2	U_1	U_2	$\frac{\pi_1}{\pi_1+\pi_2}$	Γ
6	5	1	1	0	1	0	.2	4
3	7	1	0	1	0	1	.8	4
4	7	1	0	1	0	1	.8	4
7	14	1	0	1	0	1	.8	4

Table: With an Omitted Variable

Bounds

- Define T^+ to be the sum of S independent random variables, where the sth variable takes the value of d_s with probability p_s^+ and takes the value of 0 with probability $1 - p_s^+$. Define T^- similarly with p_s^-
- If the treatment has no effect, then for each fixed $\gamma \geq 0$,

$$\mathrm{prob}(T^+ \geq a) \geq \mathrm{prob}\{T > a|\mathbf{m}\} \geq \mathrm{prob}(T^- \geq a)$$

for all a and $\mathbf{u} \in U$.

More on Bounds

What do these bounds actually mean?

- The upper bound $\text{prob}(T^+ \geq a)$ is the distribution of $t(\mathbf{Z}, \mathbf{r})$ when $u_{si} = c_{si}$ and the lower bound $\text{prob}(T^- \geq a)$ is the distribution when $u_{si} = 1 - c_{si}$

More on Bounds

What do these bounds actually mean?

- The upper bound $\text{prob}(T^+ \geq a)$ is the distribution of $t(\mathbf{Z}, \mathbf{r})$ when $u_{si} = c_{si}$ and the lower bound $\text{prob}(T^- \geq a)$ is the distribution when $u_{si} = 1 - c_{si}$
- This means that the bounds are attained values of $\mathbf{u}$ that exhibit a strong, near perfect, relationship with $\mathbf{r}$, as c_{si} is a function of r_{si}.

Calculating P-Values

- The exact P-values can be calculated, but not practical in moderate or larger datasets.

Calculating P-Values

- The exact P-values can be calculated, but not practical in moderate or larger datasets.
- Large sample approximations using:

$$E(T^+) = \sum_{s=1}^{S} d_s p_s^+$$

$$\mathrm{var}(T^+) = \sum_{s=1}^{S} d_s^2 p_s^+ (1 - p_s^+)$$

Amplification

- What if we believe that a confounder exists but it's not perfectly correlated with the outcome?

Amplification

- What if we believe that a confounder exists but it's not perfectly correlated with the outcome?
- An amplification allows the researcher to ask, how would my inferences change under different correlations between u and r, as well as u and r?

Amplification

- What if we believe that a confounder exists but it's not perfectly correlated with the outcome?
- An amplification allows the researcher to ask, how would my inferences change under different correlations between u and r, as well as u and r?
- 3 unknowns:

Amplification

- What if we believe that a confounder exists but it's not perfectly correlated with the outcome?
- An amplification allows the researcher to ask, how would my inferences change under different correlations between u and r, as well as u and r?
- 3 unknowns:
  1. Δ is degree of association between u and r

Amplification

- What if we believe that a confounder exists but it's not perfectly correlated with the outcome?
- An amplification allows the researcher to ask, how would my inferences change under different correlations between u and r, as well as u and r?
- 3 unknowns:
  1. Δ is degree of association between u and r
  2. Λ is degree of association between u and Z

Amplification

- What if we believe that a confounder exists but it's not perfectly correlated with the outcome?
- An amplification allows the researcher to ask, how would my inferences change under different correlations between u and r, as well as u and r?
- 3 unknowns:
  1. Δ is degree of association between u and r
  2. Λ is degree of association between u and Z
  3. Γ is degree of association between u and Z when u is perfectly correlated with Z.

Amplification

- What if we believe that a confounder exists but it's not perfectly correlated with the outcome?
- An amplification allows the researcher to ask, how would my inferences change under different correlations between u and r, as well as u and r?
- 3 unknowns:
  1. Δ is degree of association between u and r
  2. Λ is degree of association between u and Z
  3. Γ is degree of association between u and Z when u is perfectly correlated with Z.
- Γ can be decomposed as follows:

$$\Gamma = \frac{\Delta\Lambda + 1}{\Delta + \Lambda}$$