Sensitivity Analysis for Observational Studies

November 19, 2012

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

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

- For M units, with observed covariates $\mathbf{x}$, 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

- For M units, with observed covariates $\mathbf{x}$, 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

$$\pi_j = \Pr(Z_{[j]} = 1)$$

  and to control with probability

$$1 - \pi_j = \Pr(Z_{[j]} = 0)$$

Model of an Observational Study

- For M units, with observed covariates $\mathbf{x}$, 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

$$\pi_j = \Pr(Z_{[j]} = 1)$$

  and to control with probability

$$1 - \pi_j = \Pr(Z_{[j]} = 0)$$

- Treatments are assigned by flipping biased coins (each unit might have a different biased coin):

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

Overt Bias

- An observational study is free of *hidden* bias if every $\pi_{[j]}$ (though unknown), only depend on the observed covariates:

$$\pi_{[j]} = \pi_{[k]} \text{ iff } \mathbf{x}_{[j]} = \mathbf{x}_{[k]}$$

Overt Bias

- An observational study is free of *hidden* bias if every $\pi_{[j]}$ (though unknown), only depend on the observed covariates:

$$\pi_{[j]} = \pi_{[k]} \text{ iff } \mathbf{x}_{[j]} = \mathbf{x}_{[k]}$$

- Define the probability that j will be in treatment as some unknown function of $\mathbf{x}$:

$$\pi_{[j]} = \lambda(\mathbf{x}_{[j]})$$

Overt Bias

- An observational study is free of *hidden* bias if every $\pi_{[j]}$ (though unknown), only depend on the observed covariates:

$$\pi_{[j]} = \pi_{[k]} \text{ iff } \mathbf{x}_{[j]} = \mathbf{x}_{[k]}$$

- Define the probability that j will be in treatment as some unknown function of $\mathbf{x}$:

$$\pi_{[j]} = \lambda(\mathbf{x}_{[j]})$$

- The probability of treatment assignment becomes:

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

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

- For inference, Rosenbaum proposes we use the conditional distribution of $\mathbf{Z}$ given $\mathbf{m}$, (e.g., conditional on fixing $\mathbf{m}$ via matching).

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

- For inference, Rosenbaum proposes we use the conditional distribution of $\mathbf{Z}$ given $\mathbf{m}$, (e.g., conditional on fixing $\mathbf{m}$ via matching).
- For randomization inference, we want the full set of possible treatment assignments (Ω) and their associated probabilities.

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

- For inference, Rosenbaum proposes we use the conditional distribution of $\mathbf{Z}$ given $\mathbf{m}$, (e.g., conditional on fixing $\mathbf{m}$ via matching).
- For randomization inference, we want the full set of possible treatment assignments (Ω) and their associated probabilities.
- With S matched-strata, 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 $\mathbf{m}$, (e.g., conditional on fixing $\mathbf{m}$ via matching).
- For randomization inference, we want the full set of possible treatment assignments (Ω) and their associated probabilities.
- With S matched-strata, there are $K = \prod_{s=1}^{S} \binom{n_s}{m_s}$ possible assignments.
- Conditional on $\mathbf{m}$, 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

- How would inferences about treatment effects be altered by hidden biases of various magnitudes?

A Model for Sensitivity Analysis

- 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:

$$\text{if } \mathbf{x}_{[j]} = \mathbf{x}_{[k]}, \text{ but } \pi_{[j]} \neq \pi_{[k]} \text{ for some } j \text{ and } k.$$

A Model for Sensitivity Analysis

- For units j and k pair-matched into strata s, the odds that units j and k receive the treatment are:

$$O_{[j]} = \pi_{[j]}/(1 - \pi_{[j]}) \text{ and } O_{[k]} = \pi_{[k]}/(1 - \pi_{[k]})$$

A Model for Sensitivity Analysis

- For units j and k pair-matched into strata s, the odds that units j and k receive the treatment are:

$$O_{[j]} = \pi_{[j]}/(1 - \pi_{[j]}) \text{ and } O_{[k]} = \pi_{[k]}/(1 - \pi_{[k]})$$

- The ratio of these odds, $\Gamma = O_{[j]}/O_{[k]}$, measures bias after matching

A Model for Sensitivity Analysis

- For units j and k pair-matched into strata s, the odds that units j and k receive the treatment are:

$$O_{[j]} = \pi_{[j]}/(1-\pi_{[j]}) \text{ and } O_{[k]} = \pi_{[k]}/(1-\pi_{[k]})$$

- The ratio of these odds, $\Gamma = O_{[j]}/O_{[k]}$, measures bias after matching
- Conditional on the matching procedure, the probability of assignment to treatment within s:

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

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

- After adjusting for **x**, the odds ratio for two units in the same matched pair can be written as:

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

An Alternative Expression: Bias Due to an Unobserved Covariate

- After adjusting for $\mathbf{x}$, the odds ratio for two units in the same matched pair can be written as:

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

- Bounding this term, for $Z_{[j]} = 1$:

$$\begin{array}{llll} u_{[j]} - u_{[k]} = 1 & \rightarrow & \exp\{\gamma\} & = \Gamma \\ u_{[j]} - u_{[k]} = -1 & \rightarrow & \exp\{-\gamma\} & = 1/\Gamma \\ u_{[j]} - u_{[k]} = 0 & \rightarrow & \exp\{0\} & = 1 \end{array}$$

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

Wilcoxon Sign-score test: $t(\mathbf{Z}, \mathbf{r}) = 1$

- Permute all $z \in \Omega$, recording $I(T \geq 1) \times \Pr(\mathbf{Z} = \mathbf{z}|\mathbf{m})$

Recall: $\Pr(\mathbf{Z} = \mathbf{z}|\mathbf{m})$ for matched strata

Sensitivity of Significance Levels

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

An Observational Study with Unknowns

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	?	?	?	?
3	7	1	0	46	46	?	?	?	?
4	7	1	0	3	3	?	?	?	?
7	14	1	0	25	25	?	?	?	?

Table: With an Omitted Variable

Wilcoxon Sign-score test: $t(\mathbf{Z}, \mathbf{r}) = 1$

- Permute all $z \in \Omega$, recording $I(T \geq 1) \times \Pr(\mathbf{Z} = \mathbf{z}|\mathbf{m})$

$\Pr(\mathbf{Z} = \mathbf{z}|\mathbf{m})$ is unknown

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 $(\gamma, \mathbf{u})$, statistic $t(\mathbf{Z}, \mathbf{r})$ is the sum of S independent RV, 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 $(\gamma, \mathbf{u})$, statistic $t(\mathbf{Z}, \mathbf{r})$ is the sum of S independent RV, where the sth variable equals d_s with probability

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

- Recall c_{sj} is a function of the outcome R, such that
  $c_{s1} = 1, c_{s2} = 0$ if $r_{s1} > r_{s2}$;
  $c_{s1} = 0, c_{s2} = 1$ if $r_{s1} < r_{s2}$;
  $c_{s1} = c_{s2} = 0$ if $r_{s1} = r_{s2}$

Inference with an Unknown Confounder

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

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

- Recall c_{sj} is a function of the outcome R, such that
  $c_{s1} = 1, c_{s2} = 0$ if $r_{s1} > r_{s2}$;
  $c_{s1} = 0, c_{s2} = 1$ if $r_{s1} < r_{s2}$;
  $c_{s1} = c_{s2} = 0$ if $r_{s1} = r_{s2}$
- With $\Gamma = exp(\gamma)$ define p_s^+ and p_s^- in the following way:

$$p_s^+ = \frac{\Gamma}{1+\Gamma} \quad \text{and} \quad p_s^- = \frac{1}{1+\Gamma}$$

Recall the Example

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	.6667
3	7	1	0	1	0	1	.333
4	7	1	0	1	0	0	.5
7	14	1	0	1	1	1	.5

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 Strategy: Assume the Worst 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

Note: We assume $u_{sj} = f(r_{sj})$, or 'near-perfect' correspondence between U and the outcome R

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

Note: We choose Γ to vary the magnitude of the correspondence between U and Z

Choose a Larger 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 $\gamma \geq 0$,

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

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

More on Bounds

What do these bounds actually mean?

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

More on Bounds

What do these bounds actually mean?

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

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

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

Calculating P-Values

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

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

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

- Resulting test: $\frac{t(\mathbf{Z},\mathbf{r}) - E(T^{+})}{\sqrt{\mathrm{var}(T^{+})}} \sim N(0,1)$

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:

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

Calculating Two-Parameter P-Values

- Large sample approximations using:

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

$$\mathrm{var}(T^+) = \sum_{s=1}^{S} d_s^2 \frac{\frac{\Delta\Lambda+1}{\Delta+\Lambda}}{1 + \frac{\Delta\Lambda+1}{\Delta+\Lambda}} \left(1 - \frac{\frac{\Delta\Lambda+1}{\Delta+\Lambda}}{1 + \frac{\Delta\Lambda+1}{\Delta+\Lambda}}\right)$$

Calculating Two-Parameter P-Values

- Large sample approximations using:

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

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

- Resulting test: $\frac{t(\mathbf{Z},\mathbf{r})-E(T^+)}{\sqrt{\mathrm{var}(T^+)}} \sim N(0,1)$