On Sensitivity Value of Pair-Matched Observational Studies

Qingyuan Zhao

Department of Statistics, University of Pennsylvania

August 2nd, JSM 2017

Manuscript and slides are available at
http://www-stat.wharton.upenn.edu/~qyzhao/.

Sensitivity Analysis

- Observational studies = Treatment is not randomized.
- The core but unverifiable assumption: treatment ignorability, aka no unmeasured confounding. (Fisher's criticism of "smoking causes lung cancer".)
- Sensitivity analysis: what if this assumption is violated (in a controlled way captured by one or a few sensitivity parameters)?
- There is a long list of approaches of sensitivity analysis. I will consider Rosenbaum's sensitivity analysis for a pair-matched study. (Cornfield's response to Fisher.)

What does a sensitivity analysis look like?

- Sensitivity parameter $\Gamma \geq 1$ in Rosenbaum's model: within each matched pair,

$$1/\Gamma \leq \text{odds ratio}(\text{1st unit treated}, \text{2nd unit treated}) \leq \Gamma.$$

- $\Gamma = 1$ corresponds to ignorable treatment. Can test the sharp null hypothesis by e.g. Wilcoxon's signed rank test.
- When $\Gamma > 1$: Rosenbaum obtained lower and upper bounds of the *p*-value of any signed score test.
- An example:

probe set	sensitivity analysis							sensitivity value	
37583_at	Γ	1	2	3	5	7	10	**1.84**	**2.44**
	$\overline{p}_\Gamma$	0.00	0.02	0.13	0.60	1.00	1.00	0.01	0.05

Sensitivity value

Definition ("Truncated" sensitivity value)

$$\Gamma_{\alpha}^{**} = \inf \left\{ \Gamma \geq 1 \,|\, \overline{p}_{\Gamma} > \alpha \right\}.$$

Interpretation: If the unmasured confounder changes the within-pair odds ratio of treatment by more than Γ_{α}^{**}, then the sharp null hypothesis could be not significant.

- $\Gamma_{\alpha}^{**} > 1$ iff $\overline{p}_1 = p_1 \leq \alpha$. What if the null hypothesis is not significant even when $\Gamma = 1$?
- Mathematically, $\overline{p}_{\Gamma}$ can be defined for $0 < \Gamma < 1$ as well. It is more convenient to work with Γ_{α}^{*} that takes the infimum over $\Gamma > 0$.

Some motivations I

- Sensitivity value is a concise summary of the study's "sensitivity to measured confounding".
- A value vs. A table.
- Analogy: p-value for a randomized experiment vs. sensitivity value for an observational study.

Some motivations II

- Asymptotics of sensitivity value $\leftrightarrow$ Power of sensitivity analysis:
$$\mathrm{P}(\Gamma_\alpha^* \geq \Gamma) = \mathrm{P}(\overline{p}_\Gamma \leq \alpha).$$
- The "favorable situation": no unmeasured confounding and nonzero causal effect.
- Fixed Γ asymptotics: Rosenbaum [2015] considered the Bahadur efficiency of a sensitivity analysis by studying how fast $\overline{p}_\Gamma \to 0$.
- Fixed α asymptotics: examine the distribution of Γ_α^*.

Background

- A general and common strategy is to use the signed score test (Y_i is the within-pair difference)

$$T = \frac{\sum_{i=1}^{I} \mathrm{sgn}(Y_i) q_i}{\sum_{i=1}^{I} q_i}, \; q_i = \psi\left(\frac{\mathrm{rank}(|Y_i|)}{I+1}\right).$$

- Rosenbaum found bounding variable $\bar{T}_\Gamma$ in the sense that

$$\mathrm{P}(T \geq t|\mathcal{F}) \leq \mathrm{P}(\overline{T}_\Gamma \geq t|\mathcal{F}) = \overline{p}_\Gamma,$$

- CLT for $\bar{T}_\Gamma$:

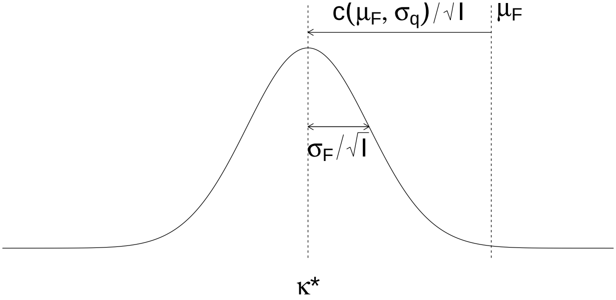
$$\sqrt{I} \cdot \frac{\overline{T}_\Gamma - \Gamma/(1+\Gamma)}{\sqrt{\Gamma/(1+\Gamma)^2 \sigma_{q,I}^2}} \xrightarrow{d} \mathrm{N}(0,1), \; \sigma_{q,I}^2 = \frac{I^{-1}\sum_{i=1}^{I} q_i^2}{\left(I^{-1}\sum_{i=1}^{I} q_i\right)^2}.$$

CLT for sensitivity value

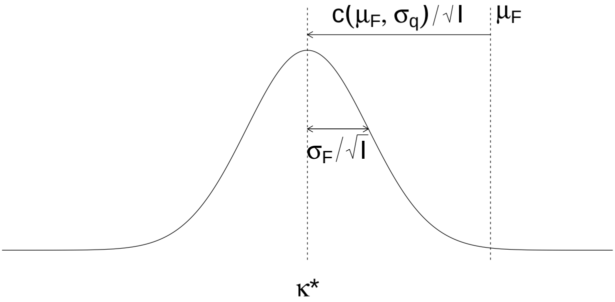
Theorem (Z, 2017)

Suppose $Y_i \overset{i.i.d.}{\sim} F$ *and* $\sqrt{I} \cdot (T - \mu_F)/\sigma_F \overset{d}{\to} \mathrm{N}(0,1)$, *the transformed sensitivity value* $\kappa^*_\alpha = \Gamma^*_\alpha/(1+\Gamma^*_\alpha)$ *for fixed* $0 < \alpha < 1$ *satisfies*

$$\sqrt{I} \cdot \left[\kappa^*_\alpha - \mu_F\right] \overset{d}{\to} \mathrm{N}\left(-\sigma_q \bar{\Phi}^{-1}(\alpha)\sqrt{\mu_F(1-\mu_F)},\ \sigma_F^2 \right).$$

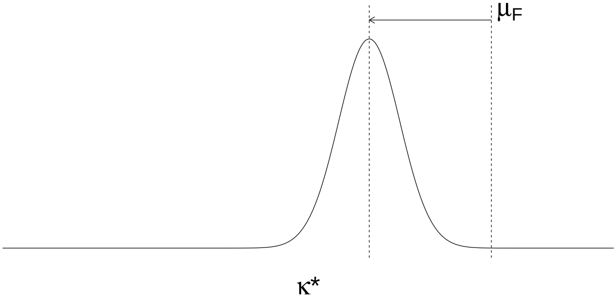


Design sensitivity



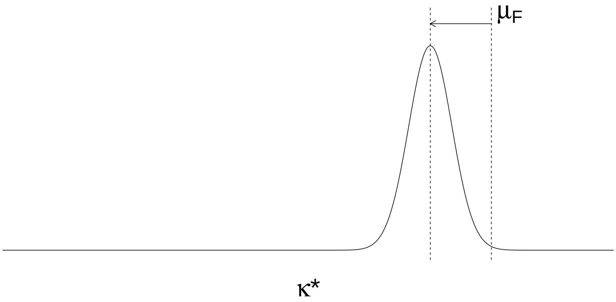
- Rosenbaum [2004] noticed a phase transition at μ_F:
  - If $\kappa > \mu_F$, sensitivity analysis has no asymptotic power.
  - If $\kappa < \mu_F$, power $\to 1$ as $I \to \infty$.
  - He calls the value $\mu_F/(1-\mu_F)$ "design sensitivity".
- The CLT for sensitivity value separates the contribution by design and by obtaining more sample.

Design sensitivity



- Rosenbaum [2004] noticed a phase transition at μ_F:
  - If $\kappa > \mu_F$, sensitivity analysis has no asymptotic power.
  - If $\kappa < \mu_F$, power $\to 1$ as $I \to \infty$.
  - He calls the value $\mu_F/(1-\mu_F)$ "design sensitivity".
- The CLT for sensitivity value separates the contribution by design and by obtaining more sample.

Design sensitivity



- Rosenbaum [2004] noticed a phase transition at μ_F:
  - If $\kappa > \mu_F$, sensitivity analysis has no asymptotic power.
  - If $\kappa < \mu_F$, power $\to 1$ as $I \to \infty$.
  - He calls the value $\mu_F/(1-\mu_F)$ "design sensitivity".
- The CLT for sensitivity value separates the contribution by design and by obtaining more sample.

Implication: select statistics

- The goal: maximize Γ^*_α stochastically.

$$\sqrt{I} \cdot \left[\kappa^*_\alpha - \mu_F\right] \xrightarrow{d} \mathrm{N}\left(-\sigma_q \bar{\Phi}^{-1}(\alpha)\sqrt{\mu_F(1-\mu_F)},\ \sigma_F^2\right).$$

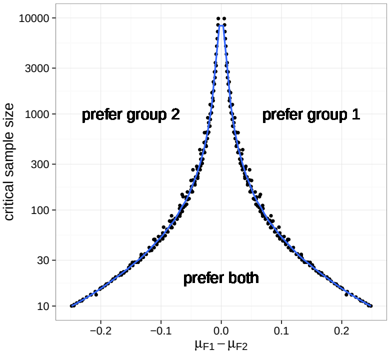
- The expecation of Γ^*_α is determined by two quantities:
  1. $\mu_F =< \psi, g_F >$;
  2. $\sigma_q = \|\psi\|_2^2/\|\psi\|_1^2$.
- Tradeoff: want to maximize μ_F without making σ_q too large.
- Manuscript has detailed comparison of different choices of ψ under different F.
- Practically, can estimate μ_F by sample splitting and then decide.

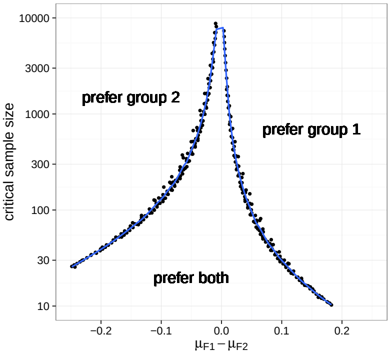
Implication: select subgroups

- Hsu, Small, and Rosenbaum [2013] noticed a dilemma. Suppose there are two subgroups with unequal μ_F. Heuristically,
  - If $I \to \infty$, should use the subgroup with larger μ_F.
  - If I is small, should combine the samples.

Implication: select subgroups

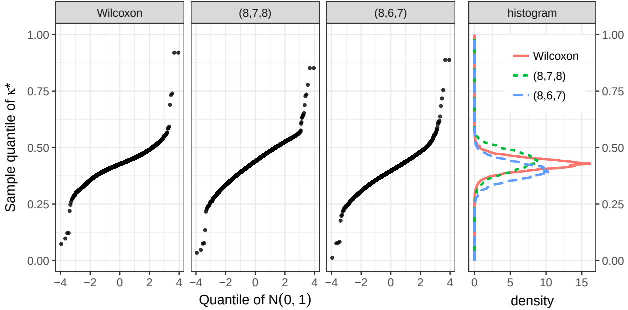
- Hsu, Small, and Rosenbaum [2013] noticed a dilemma. Suppose there are two subgroups with unequal μ_F. Heuristically,
  - If $I \to \infty$, should use the subgroup with larger μ_F.
  - If I is small, should combine the samples.
- Can use the asymptotics to compute the critical sample size.
- For example, for Noether's statistics $\psi(u) \equiv 1$ (ratio of sample sizes below: 1:1 and 3:1).





Implication: select hypotheses

- One treatment and hundreds of outcomes that are susceptible to unmeasured confounding.
- Can use sensitivity value to screen causal hypotheses.
- The manuscript has an application to genomics screening.



- Zhao, Small, and Rosenbaum [2017] proposed a related method called "Cross Screening".

References

- Manuscript:
  - Q. Zhao. On sensitivity value of pair-matched observational studies. arXiv:1702.03442.
- Additional references:

J. Y. Hsu, D. S. Small, and P. R. Rosenbaum. Effect modification and design sensitivity in observational studies. *Journal of the American Statistical Association*, 108(501):135–148, 2013.

P. R. Rosenbaum. Design sensitivity in observational studies. *Biometrika*, 91(1): 153–164, 2004.

P. R. Rosenbaum. Bahadur efficiency of sensitivity analyses in observational studies. *Journal of the American Statistical Association*, 110(509):205–217, 2015.

Q. Zhao, D. S. Small, and P. P. Rosenbaum. Cross-screening in observational studies that test many hypotheses. *arXiv:1703.02078*, 2017.