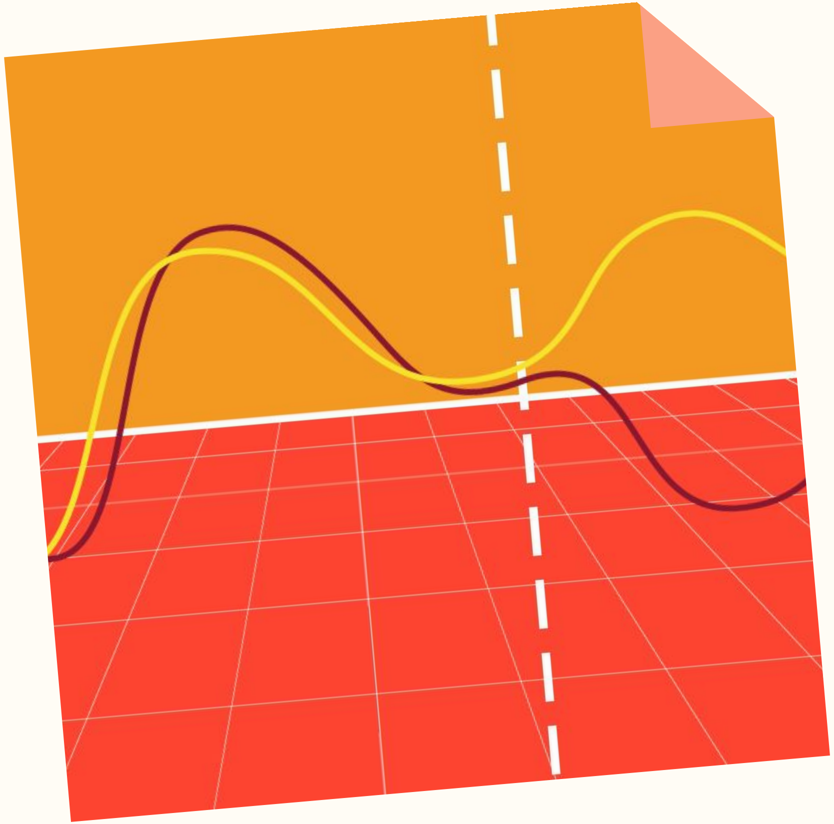
Sensitivity Analysis in Causal Inference

- Shuozishan(Doris) Wang
- Mentor: Kenny Zhang

Table of Contents



What is Causal Inference & Sensitivity Analysis?

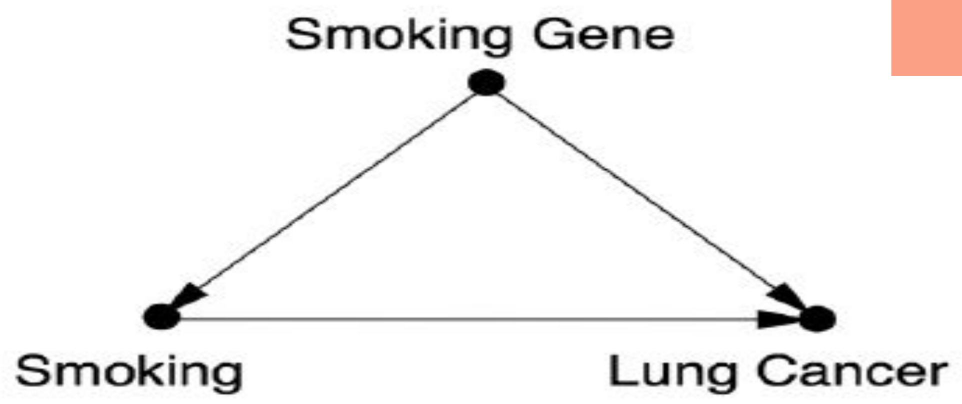
📌**Causal Inference:** Understanding **cause-and-effect** relationships.

📌 **Challenge: Unmeasured confounders** in observational studies can bias conclusions.

📌 **Why Sensitivity Analysis?** Helps assess the **robustness** of causal claims.

Correlation vs. Causation – The Smoking & Cancer Example

"Does smoking really cause lung cancer, or could an unmeasured confounder—like genetics—be responsible?"



The DAG

What is the Potential Outcomes Framework?

- $Y(1)$ = The outcome were an individual to take treatment
- $Y(0)$= The outcome were an individual to take control

Equation for the Causal Effect:

$$\tau = Y(1) - Y(0)$$

Individual	Smokes?	$Y(1)$	$Y(0)$	Observed Outcome?
A	Yes	1	?	1 (Developed Cancer)
B	No	?	0	0 (No Cancer)

OVB & Sensitivity Analysis In Linear Regression

📌 What is Omitted Variable Bias (OVB)?

- **Occurs when an important confounder is missing from the regression model.**
- Leads to **biased estimates** of causal effects.
- Example: If **genetics influences both smoking and lung cancer**, but we do not include it in our model, the effect of smoking may be **overestimated or underestimated**.

📌 How Does Sensitivity Analysis Help?

- Sensitivity analysis **quantifies how much an omitted confounder** could impact our conclusions.
- Helps us determine **if our results are robust** or **sensitive to hidden bias**.

Robustness Value & Partial R^2

✔ A Simple Formula Representation of Robustness Value (RV):

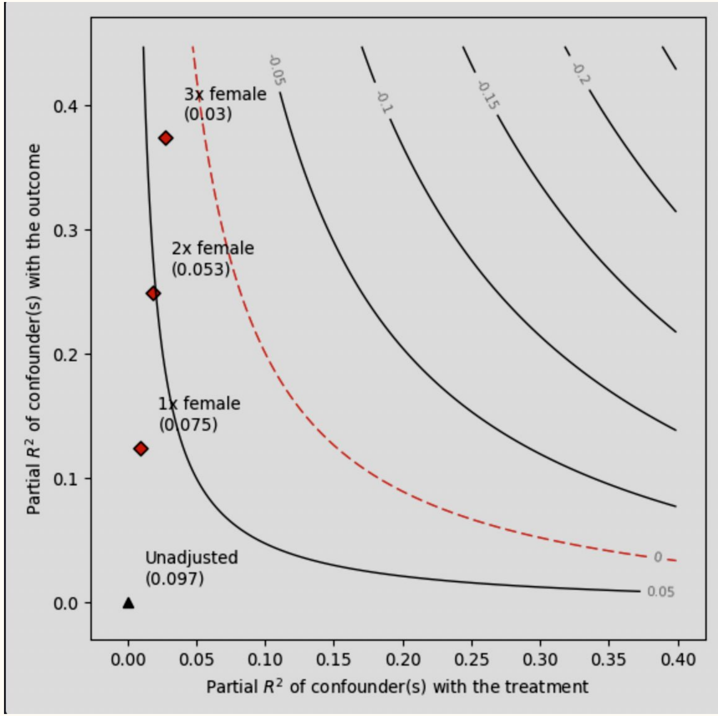
RV=Minimum strength of confounder needed to erase causal effect

📌 What is Robustness Value (RV)?

- Measures **how strong an unmeasured confounder must be** to completely eliminate the observed effect.
- If **RV is high**, the results are **robust** to hidden bias.
- If **RV is low**, the results are **sensitive** to omitted variables.

📌 What is Partial R^2?

- Measures **how much variation in the treatment (smoking) and outcome (lung cancer)** is explained by an unmeasured confounder (e.g., genes).



We are now ready to express the bias in terms of partial R^2. First, by the FWL theorem,

$$\begin{aligned}\widehat{\text{bias}} &= \hat{\delta}\hat{\gamma}\\ &= \left(\frac{\text{cov}(D^{\perp \boldsymbol{X}},\ Z^{\perp \boldsymbol{X}})}{\text{var}(D^{\perp \boldsymbol{X}})}\right)\left(\frac{\text{cov}(Y^{\perp \boldsymbol{X},D},\ Z^{\perp \boldsymbol{X},D})}{\text{var}(Z^{\perp \boldsymbol{X},D})}\right)\\ &= \left(\frac{\text{cor}(D^{\perp \boldsymbol{X}},\ Z^{\perp \boldsymbol{X}})\text{sd}(Z^{\perp \boldsymbol{X}})}{\text{sd}(D^{\perp \boldsymbol{X}})}\right)\left(\frac{\text{cor}(Y^{\perp \boldsymbol{X},D},\ Z^{\perp \boldsymbol{X},D})\text{sd}(Y^{\perp \boldsymbol{X},D})}{\text{sd}(Z^{\perp \boldsymbol{X},D})}\right)\\ &= \left(\frac{\text{cor}(Y^{\perp \boldsymbol{X},D},\ Z^{\perp \boldsymbol{X},D})\text{cor}(D^{\perp \boldsymbol{X}},\ Z^{\perp \boldsymbol{X}})}{\frac{\text{sd}(Z^{\perp \boldsymbol{X},D})}{\text{sd}(Z^{\perp \boldsymbol{X}})}}\right)\left(\frac{\text{sd}(Y^{\perp \boldsymbol{X},D})}{\text{sd}(D^{\perp \boldsymbol{X}})}\right)\end{aligned} \tag{7}$$

Noting that $\text{cor}(Y^{\perp \boldsymbol{X},D}, Z^{\perp \boldsymbol{X},D})^2 = R^2_{Y\sim Z|\boldsymbol{X},D}$, that $\text{cor}(Z^{\perp \boldsymbol{X}},\ D^{\perp \boldsymbol{X}})^2 = R^2_{D\sim Z|\boldsymbol{X}}$, and that $\frac{\text{var}(Z^{\perp \boldsymbol{X},D})}{\text{var}(Z^{\perp \boldsymbol{X}})} = 1 - R^2_{Z\sim D|\boldsymbol{X}} = 1 - R^2_{D\sim Z|\boldsymbol{X}}$, we can write 7 as

$$|\widehat{\text{bias}}| = \sqrt{\frac{R^2_{Y\sim Z|D,\boldsymbol{X}}\ R^2_{D\sim Z|\boldsymbol{X}}}{1 - R^2_{D\sim Z|\boldsymbol{X}}}}\left(\frac{\text{sd}(Y^{\perp \boldsymbol{X},D})}{\text{sd}(D^{\perp \boldsymbol{X}})}\right). \tag{8}$$

Reference and Further Reading

1. Meghanath, Ganga. Causal Analysis Overview: Causal Inference versus Experimentation versus Causal Discovery. Data Science at Microsoft, 5 Nov. 2024, https://medium.com/data-science-at-microsoft/causal-analysis-overview-causal-inference-versus-experimentation-versus-causal-discovery-d7c4ca99e3e4.
2. Cinelli, Carlos, and Chad Hazlett. Making Sense of Sensitivity: Extending Omitted Variable Bias. Journal of the Royal Statistical Society, Series B (Statistical Methodology), 2020, https://doi.org/10.1111/rssb.12348.
3. Ding, Peng. A First Course in Causal Inference. arXiv:2305.18793v2, 3 Oct. 2023,https://arxiv.org/abs/2305.18793.

Questions?