Extrapolation Parameterizations for Assessing Sensitivity to Unmeasured Confounding

UC SANTA BARBARA

Alexander Franks[1], **Alexander D'Amour**[2], Avi Feller[2]

UC Santa Barbara[1], UC Berkeley[2]

Berkeley UNIVERSITY OF CALIFORNIA

Sensitivity Analysis

If **unconfoundedness** is violated

$$[Y(0), Y(1)] \not\perp\!\!\!\perp T \mid X$$

Estimated observed data distribution $\hat{f}(Y^{obs}, T \mid X)$ consistent with many causal conclusions.

Sensitivity Analysis: explore **ignorance region**, or space of causal conclusions consistent with fixed $\hat{f}(Y^{obs}, T \mid X)$.

Latent Confounder Sensitivity Analysis

Posit a model with latent confounder U satisfying

$$[Y(0), Y(1)] \perp\!\!\!\perp T \mid X, U$$

Sensitivity parameters γ link U to Y^{obs} and T.

Method 1: γ-Conditional Analysis (Classical sensitivity analysis, cf, Rosenbaum & Rubin 1983)

Obtain **conditional complete data posterior predictive distribution** (fixing γ at grid of values, integrating out U):

$$\tilde{f}_\gamma([Y(0), Y(1)], T \mid X)$$

Report **range** of implied treatment effects.

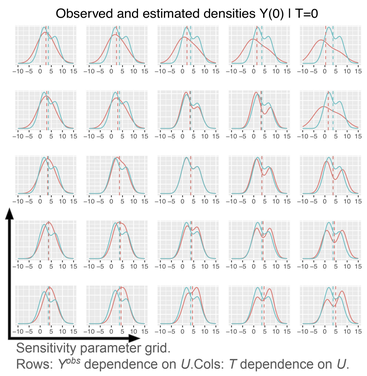
Method 2: γ-Marginal Analysis ("Bayesian Sensitivity Analysis", cf, McCandless et al. 2006)

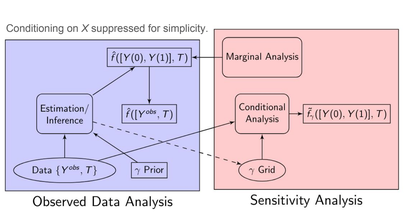
Obtain **complete data posterior predictive distribution** (integrating out both U and γ):

$$\hat{f}([Y(0), Y(1)], T \mid X)$$

Report **posterior uncertainty** in treatment effect.

Problem: Sensitivity parameters often (partially) identified.





(Above): Schematic of standard Bayesian sensitivity analysis. Data analysis and sensitivity analysis operations are intertwined in complex ways.

(Left): Illustration of complications from identified sensitivity parameter in γ-conditional analysis. Each setting of γ implies a different observed data posterior predictive distribution. The observed data fit changes at each setting of the sensitivity parameters.

✘**Conflates posterior uncertainty with sensitivity/ignorance.**

✘**Observed data fit depends on sensitivity parameter setting.**

✘**Computationally inefficient: New estimation for each γ setting.**

Extrapolation Factorization

Key idea: Factorize complete data density into **fully identified** and **fully unidentified** factors. Original missing data application, cf, Franks et al. 2016, Linero and Daniels 2017.

All densities conditional on X. Conditioning on X suppressed for simplicity.

$$f([Y(0), Y(1)], T) = \underbrace{f(Y(0) \mid T = 0)}_{\text{Observed control outcome density}} f(T = 0) \cdot \frac{f(T \mid Y(0))}{f(T = 0 \mid Y(0))} \cdot$$

Complete data density

$$\underbrace{f(Y(1) \mid T = 1)}_{\text{Observed treated outcome density}} f(T = 1) \cdot \frac{f(T \mid Y(1))}{f(T = 1 \mid Y(1))} \cdot \frac{1}{f(T)} \cdot$$

$$c(F(Y(0) \mid T), F(Y(1) \mid T) \mid T)$$

Marginal Selection Models
Dependence between treatment assignment and potential outcomes.

Conditional Copula
Dependence between potential outcomes. Irrelevant for standard estimands.

Marginal selection models sufficient to **extrapolate** from observed to **missing potential outcome** distributions, e.g.,

$$\underbrace{f(Y(0) \mid T = 1)}_{\text{Missing control outcome density}} \propto \underbrace{f(Y(0) \mid T = 0)}_{\text{Observed control outcome density}} \cdot \frac{f(T = 1 \mid Y(0))}{f(T = 0 \mid Y(0))}$$

Extrapolation-Based Sensitivity Analysis

Extrapolation Analysis: Specify marginal selection models; extrapolate from **observed data posterior predictive distribution**

$$\hat{f}(Y^{obs}, T \mid X) \overset{\gamma}{\mapsto} \{\tilde{f}_\gamma(Y(0), T \mid X), \tilde{f}_\gamma(Y(1), T \mid X)\}$$

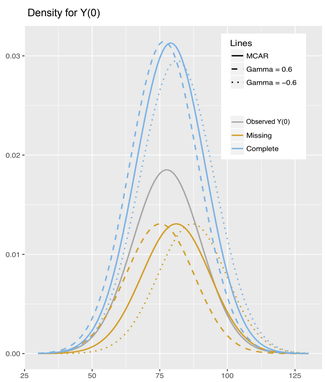
Example selection specification: For a grid of values for (γ_0, γ_1):

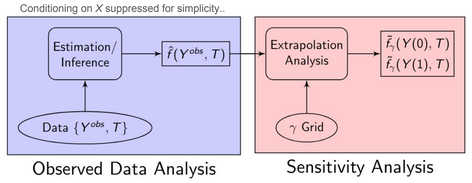
Extrapolation is simple when $f(Y^{obs} \mid T, X)$ is in the **exponential family** (e.g., Normal, Bernoulli, Poisson) and selection model is **logistic**.

$$f(T \mid X, Y(1)) \propto \text{Bern}[\text{Logit}^{-1}(\mu_1(X) + \gamma_1 Y(1))]$$

$$f(T \mid X, Y(0)) \propto \text{Bern}[\text{Logit}^{-1}(\mu_0(X) + \gamma_0 Y(0))]$$

Calibrate γ grid w/ largest coefficient magnitude in propensity score model.





(Above): Schematic of extrapolation-based sensitivity analysis. Data analysis and sensitivity analysis are cleanly separated. Data is only touched once; all extrapolation analysis is post-hoc.

(Left): Example extrapolation analysis assuming a normal model for observed outcomes and logistic missingness. Observed outcome density is gray, extrapolated missing outcome density is orange and combined complete outcome density is blue. The logistic missingness model implies that the missing data density is simply a shifted version of the observed data density, because the observed data density is in the exponential family.

✔**Guarantees sensitivity parameters are unidentified.**

✔**Clean separation of estimation/inference and sensitivity analysis.**

✔**Computationally efficient: Estimation happens only once.**

NHANES III Analysis

Modeled on Dorie et al 2016

Goal: Estimate effect of **taking two or more anti-hypertensives** on average **diastolic blood pressure** (BP). Normal ~ 80, High ~ 90.

Data: Third National Health and Nutrition Examination Survey (NHANES III).

Observed data model: BART on treated and control outcome surfaces

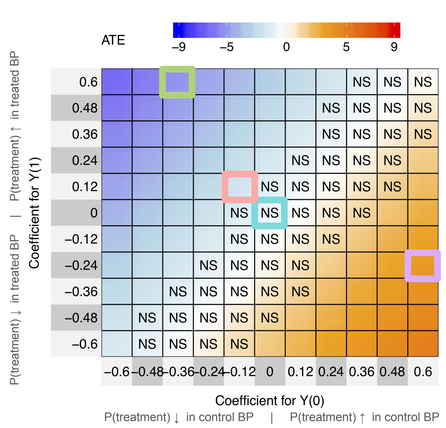
$$f(Y(t) \mid T = t, X) \sim N(\alpha_t(X), \sigma_t^2)$$

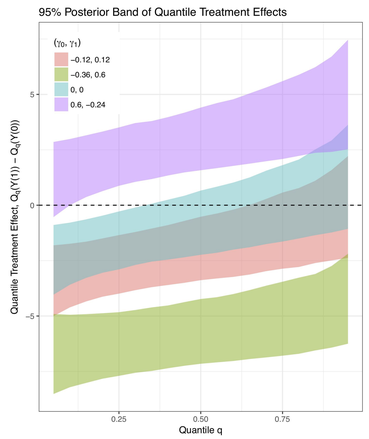
$$\alpha_t(X), \sigma_t^2 \sim BART(X \mid T = t)$$

Marginal selection model: Logistic in outcome

$$f(T \mid Y(t), X) \sim \text{Bern}(\text{Logit}^{-1}(\phi + \beta X + \gamma_t Y(t)))$$

Calibration: Largest coefficient in propensity score (given all X) is 0.33 (BMI). Set (γ_0, γ_1) grid with extreme values ~2x this magnitude.





Sensitivity analysis for **Average Treatment Effect** (left) and **Quantile Treatment Effects** (right).

Boxes marked **NS** (not significant) imply ATE posterior credible intervals containing zero. Posterior bands for QTEs from parameter settings in **highlighted boxes** are plotted on the right.

- **Teal:** unconfounded.
- **Lavender** and **light green:** extreme selection on the control and treated outcomes.
- **Pink:** smallest outcome-based selection that yields a credible, favorable ATE estimate.

Computation: Sample from observed data model posterior **once**; analysis at each (γ_0, γ_1) is *post hoc*, parallel.

References

Vincent Dorie, Masataka Harada, Nicole Bohme Carnegie, and Jennifer Hill. A flexible, interpretable framework for assessing sensitivity to unmeasured confounding. *Statistics in medicine*, 35(20):3453–3470, 2016.

Alexander M Franks, Edoardo M Airoldi, and Donald B Rubin. Non-standard conditionally specified models for non-ignorable missing data. *arXiv preprint arXiv:1603.06045*, 2016.

Antonio R Linero and Michael J Daniels. Bayesian approaches for missing not at random outcome data: The role of identifying restrictions. 2017.

Lawrence C. McCandless, Paul Gustafson, and Adrian Levy. Bayesian sensitivity analysis for unmeasured confounding in observational studies. *Statistics in Medicine*, 26(11):2331–2347, 2006.

Paul R Rosenbaum and Donald B Rubin. Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 212–218, 1983.