Fraunhofer
IKS

Fraunhofer-Institut für Kognitive Systeme IKS

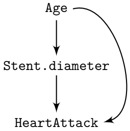
Leopold Mareis (Narges Ahmidi [IKS], Mathias Drton [TUM])

Nonparametric Effect Estimation and Guideline Search for Continuous Treatments

Motivating Example

- We want to measure the effect of the stent diameter on the heart attack rate. Is `lm(HeartAttack ~ Stent.diameter)` an appropriate model?

- How can we estimate the effect non-parametrically assuming `Stent.diameter` is continuous?

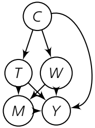


- And what is a good guideline telling physicians how to select the stent diameter based on the patient's age?

Do-Operator and Interventions [Pearl, 2009]

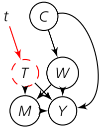
We are assuming a continuous SEM $X_i = h_i(X_{\text{pa}(i)}, \varepsilon_i)$, ε_i independent, with a binary response $Y = h_Y(X_{\text{pa}(Y)}, \varepsilon_Y)$. No unobserved confounding.

P_X



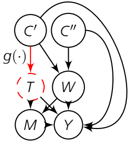
$$\begin{aligned} C &= \varepsilon_1 \\ T &= h_1(C, \varepsilon_2) \\ W &= h_2(C, \varepsilon_3) \\ M &= h_3(T, W, \varepsilon_4) \\ Y &= h_4(C, M, W, \varepsilon_5) \end{aligned}$$

$P_{X;do(T=t)}$



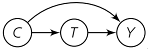
$$\begin{aligned} C &= \varepsilon_1 \\ T &= \color{red}{t} \\ W &= h_2(C, \varepsilon_3) \\ M &= h_3(T, W, \varepsilon_4) \\ Y &= h_4(C, M, W, \varepsilon_5) \end{aligned}$$

$P_{X;do(T=g(A))}$



$$\begin{aligned} (C', C'') &= \varepsilon_1 \\ T &= \color{red}{g(C')} \\ W &= h_2(C, \varepsilon_3) \\ M &= h_3(T, W, \varepsilon_4) \\ Y &= h_4(C, M, W, \varepsilon_5) \end{aligned}$$

Marginal Integration (MI, Backdoor Adjustment)

Under 

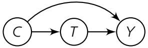
, the average causal effect can be written as an observational quantity:

$$\mathbb{E}[Y; do(T = g(C))] = \int \mathbb{E}[Y|c, T = g(C)] f(c)\, \mathrm{d}c \approx \frac{1}{n} \sum_{i=1}^{n} \hat{\mathbb{E}}[Y|C^i, T = g(C^i)]$$

In practice, this requires estimating the conditional expectation (linear model, logistic regression, local linear estimator) and approximating the integral.

[Ernest and Bühlmann, 2015]

Inverse Probability Weighting

Again, 

. For this slide, assume T binary:

$$\mathbb{E}[Y; do(T = g(C))] = \int yf(y|a, T = g(c))f(c)\,\mathrm{d}(y, c) = \mathbb{E}\left[Y \frac{\mathbb{1}_{T=g(C)}}{f(T|C)}\right]$$

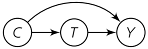
Example

	T	f(T|C)
Alice	1	0.2
Bob	0	0.4
Charlie	1	0.5

Goal: $\mathbb{E}[Y; do(T = 1)]$
Virtual dataset: $\tilde{D} := \{A, A, A, A, A, C, C\}$
Estimate interventional expectation as average of Y in $\tilde{D}$.

[Robins et al., 1994]

Continuous IPW

Assume 

. Let T be continuous and K^n be kernels.

$$\mathbb{E}[Y; do(T = g(C))] \overset{n\to\infty}{\longleftarrow} \mathbb{E}\left[Y \frac{K^n(T - g(C))}{f(T|C)}\right] \approx \frac{1}{n}\sum_{i=1}^{n} Y^i \frac{K(T^i - C^i)}{\hat{f}(T^i|C^i)}$$

- $K(T - g(C))$ quantifies the distance of the natural treatment to the intervention guideline instead of $\mathbb{1}_{T=g(C)}$.
- In CIPW, the propensity score $f(T|C)$ has to be estimated opposed to the conditional expectation in MI.

[Kallus and Zhou, 2018]

Heuristic for (explainable) guidelines

Goal: Find guideline $\hat{g}(\cdot)$ with small $\mathbb{E}[Y; do(T = \hat{g}(C))]$

Idea: We have for any guideline g the natural lower bound

$$\mathbb{E}[Y; do(T = g(C))] \overset{MI}{\approx} \frac{1}{n} \sum \hat{\mathbb{E}}[Y|T = g(C^i), C^i] \geq \frac{1}{n} \sum \min_{\tau} \hat{\mathbb{E}}[Y|T = \tau, C^i].$$

Steps:

1. Find $\hat{\tau}^i = \text{argmin}_{\tau \in supp(T)} \hat{\mathbb{E}}[Y|T = \tau, C^i]$, the risk minimizing treatment for sample i.
2. Find an (explainable) guideline $\hat{g}$ such that $\hat{g}(C^i) \approx \hat{\tau}^i$ for all samples.
3. Estimate $\mathbb{E}[Y; do(T = \hat{g}(C))]$ with CIPW, which is based on $\hat{f}(C|T)$.

Advantage: Data-driven heuristic and not biased by expert assessment.

Dataset: Stent Implantation

Sets of Variables[1]**:**

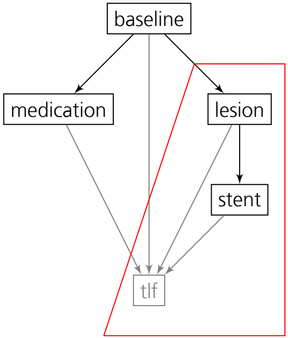
- baseline information of the patient (age, smoking status, nr. of prior heart attacks),
- medication information (anti-platelet, heparin),
- lesion information of the target vessel (length, diameter, residual blood flow, eccentricity),
- stent information (length, diameter, applied pressure, inflation time) and
- target lesion failure [response, binary] (heart attack or revascularization) information for up to 5 years.

Goal: Estimate intervention expectations $\mathbb{E}\left[\texttt{tlf}; do\left(\texttt{stent.attribute} = \xi\right)\right]$ and $\mathbb{E}\left[\texttt{tlf}; do\left(\texttt{stent.attribute} = \xi \cdot \texttt{lesion.attribute}\right)\right]$ and apply heuristic.

[1] In this talk, we neglect pre- and post-dilatation balloon information

Expert Knowledge Causal DAG

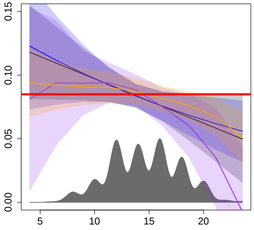
- We have baseline, medication $\perp\!\!\!\perp$ stent | lesion. Thus, we can restrict our analysis on the three red groups of random variables on the right.
- Every node may contain a random vector. To allow for slimmer regression models, we trim superfluous edges with based on conditional independence tests (assuming additive noise).



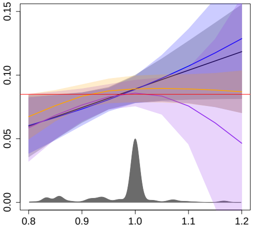
[Peters et al., 2014], similar to their phase 2

Results: Effect estimation

$\mathbb{E}[tlf; do(T = t)]$ against t



$\mathbb{E}[tlf; do(T = \xi \cdot C)]$ against ξ

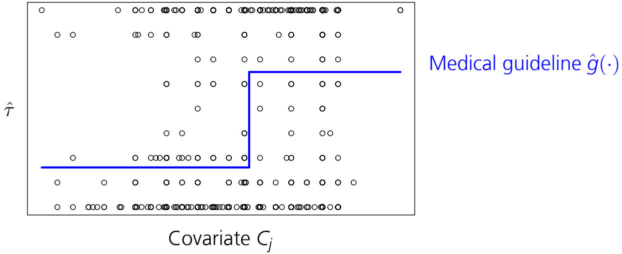


MI with linear model, MI with logistic regression, MI with local linear estimator, CIPW with kernel density estimator
dataset tlf risk, gray area: underlying density, shaded areas: 90% bootstrap CI

$\Rightarrow$ Recommend higher values of t | $\Rightarrow$ Recommend lower values of ξ

- Not shown here: similar results when computed with the original and trimmed DAG.

Results: Heuristic guideline



- Here, $\hat{g}(\cdot)$ is just a function of C_j. Simultaneously construct guidelines for all covariates and check intervention expectation.
- Few guidelines only seem to decrease the heart attack rate as often $\hat{g}(C_j) - \hat{\tau}$ is large.
- Explainable guidelines that can be directly implemented in hospitals and observed be further in future medical studies.

References

Ernest, J. and Bühlmann, P. (2015).
Marginal integration for nonparametric causal inference.
Electronic Journal of Statistics, 9(2):3155–3194.

Kallus, N. and Zhou, A. (2018).
Policy evaluation and optimization with continuous treatments.
In *International conference on artificial intelligence and statistics*, pages 1243–1251. PMLR.

Pearl, J. (2009).
Causality.
Cambridge university press.

Peters, J., Mooij, J. M., Janzing, D., and Schölkopf, B. (2014).
Causal discovery with continuous additive noise models.
Journal of Machine Learning Research, 15(58):2009–2053.

Robins, J. M., Rotnitzky, A., and Zhao, L. P. (1994).
Estimation of regression coefficients when some regressors are not always observed.
Journal of the American statistical Association, 89(427):846–866.

Fraunhofer IKS

Fraunhofer-Institut für Kognitive Systeme IKS

Contact

Leopold Mareis
Division RAD
leopold.mareis@fraunhofer.de
leopold.mareis@tum.de

Funded by IKS Evolution Project