Testing the Causal Effects of Continuous Treatments Learned by Deep Neural Networks

Yan-Yu Chen [1] Qixian Zhong [2] Jane-Ling Wang [1]

[1]University of California, Davis [2]Xiamen University

Motivation

A fundamental objective of causal inference is to comprehend the causal effect of a treatment on an outcome of interest. In the case of continuous treatment, researchers are interested in the average dose-response function (ADRF). Neural networks have been used to model the unknown functions in causal studies. For instance, varying coefficient neural network (VCNet) [4] produces a doubly robust estimator of ADRF. However, previous works predominantly focus on estimation.

To evaluate the significance of continuous treatments, we treat the ADRF as a constant function when the continuous treatment has no causal effect. So far, there are two methods exploring this problem. However, those methods critically rely on the quality of predicted individual potential outcome. Specially, the prediction may lead to a significant error when there is no causal effect under the null.

- Test based on empirical regression process: **Westling** L_1, L_2, and L_∞ [5]
- Smooth-based test: **DRDRtest** [1]

Our goal in this project is to develop a nonparametric testing procedure for causal neural network modeling that ensures a valid type-I error.

Dose-Response Function

- **Potential outcome** $Y(t)$: the outcome Y receiving the treatment $T = t$
- **Dose-response function**: For a continuous treatment, $Y(t)$ is a dose-response function
- **Average dose-response function**:

$$g(t) := E[Y(t)]$$

Assess the causal effects of T on Y by the shape of g

Test Statistic

- **Null hypothesis**: Let $\mathcal{G}_0 = \{g : \mathcal{T} \to \mathbb{R} \mid g(t) = c \text{ for some constant } c\}$

$$H_0 : g \in \mathcal{G}_0$$

This is an extension of binary treatment to test ATE $= g(1) - g(0) = 0$

- **Risk difference**: $\Delta := \min_g R(g) - \min_{g \in \mathcal{G}_0} R(g)$, where

$$R(g) = E\left[\int_{\mathcal{T}} (Y(t) - g(t))^2 dt\right]$$

Then, $g \in \mathcal{G}_0$ iff $\Delta = 0$ and $g \notin \mathcal{G}_0$ iff $\Delta < 0$

One-split test statistic

- Split sample into estimation set $\mathcal{E}$ and inference set $\mathcal{I}$
- Train VCNet on $\mathcal{E}$ and evaluate $\hat{\Delta}_i = \hat{R}_i(\hat{g}) - \hat{R}_i(\hat{g}_0)$ on $\mathcal{I}$
- Let $\ell = |\mathcal{I}|$ and propose

$$\theta_n = \sqrt{\ell} \frac{\frac{1}{\ell}\sum_{i=1}^{\ell}(\hat{\Delta}_i + \rho\epsilon_i)}{\sqrt{\hat{\sigma}^2(\hat{\Delta}_i) + \rho^2}},$$

where $\rho > 0$ and $\epsilon_i \overset{iid}{\sim} N(0,1)$ is a perturbation

K-fold test statistic

- $\theta_n^* = \frac{1}{K}\sum_{k=1}^{K} \theta_n^{(k)}$. If $K = 1$, $\theta_n^* = \theta_n$

Theoretical Analysis

- **Doubly-robust asymptotic normality of** θ_n^*: Under mild assumptions, if $\rho > 0$ and $\ell^{(k)} = o(n^{4\gamma})$, where $\gamma > 0$ satisfies $\|\hat{g}^{(k)} - g\|_2 = O_p(n^{-\gamma})$, then $\theta_n^* \xrightarrow{d} N(0,1)$ under H_0.
- p**-value**: $P_n^* = F_{N(0,1)}(\theta_n^*)$
- **Alternative hypothesis**:

$$H_a : \Delta = -\delta/\sqrt{\ell} < 0$$

for $\delta > 0$, where δ controls the distance from H_a to H_0

- **Consistency**: Under mild assumptions, if $\ell^{(k)} = o(n^{4\gamma})$, then for any $0 < \alpha < 1$,

$$\lim_{n\to\infty} P(\theta_n^* \geq \alpha | H_a) = F_{N(0,1)}\left(z_{1-\alpha} - \frac{\delta}{\rho}\right)$$

and

$$\lim_{\delta\to\infty} \lim_{n\to\infty} P(\theta_n^* \geq \alpha | H_a) = 0$$

Simulation

Irrelevant Covariates

Simu1 is a synthetic dataset containing features that are irrelevant to the causal problem. We generate covariates $X = (X_1, \ldots, X_6) \in \mathbb{R}^6$ with $X_j \overset{\text{i.i.d.}}{\sim} \text{Unif}(0,1)$, treatment

$$T = \text{logit}^{-1}(\tilde{T}),$$

$$\tilde{T} \mid X = \frac{10\sin(\max(X_1, X_2, X_3)) + \max(X_3, X_4, X_5)^3}{1 + (X_1 + X_5)^2} + \sin(0.5X_3)(1 + \exp(X_4 - 0.5X_3)) + X_3^2 + 2\sin(X_4) + 2X_5 - 6.5 + N(0, 0.5^2),$$

and a continuous outcome of interest

$$Y \mid X, T \sim N(\mu(T \mid X), 0.5^2),$$

$$\mu(T \mid X) = \cos(2\pi(T - 0.5))\frac{4\max(X_1, X_6)^3}{1 + 2X_3^2}\sin(X_4 - 0.5) + g(T),$$

where $g(t) = E[\mu(t \mid X)] = \delta(t + 0.2)^2\cos(2\pi(t - 0.3)) + 1$ is the ture ADRF

Irrelevant Covariates with Weak Alternatives

Simu2 is modified from Simu1 to demonstrate a partially effective treatment curve scenario with the presence of irrevalant covariates. The ADRF becomes

$$g(t) = \delta t^2\cos(2\pi(t - 0.5)) + 1,$$

where the first quarter of g is nearly constant. The rest of the setting is identical to Simu1.

	Simu1			Simu2		
Method	$\delta = 0$	$\delta = 0.3$	$\delta = 0.5$	$\delta = 0$	$\delta = 0.3$	$\delta = 0.5$
One-split	**0.03**	1.00	1.00	**0.04**	0.85	0.99
K-fold	**0.06**	1.00	1.00	**0.06**	0.96	1.000
Westling L_1	0.615	1.00	1.00	0.69	0.985	1.00
Westling L_2	0.625	1.00	1.00	0.705	0.99	1.00
Westling L_∞	0.645	1.00	1.00	0.725	0.99	1.00
DRDRtest	0.405	0.895	0.985	0.405	0.855	0.965

Table 1. Empirical rejection rate for synthetic datasets at level $\alpha = 0.05$ for 1000 observation with 200 replications. The parameter δ controls the distance between the true treatment effect model and the null hypothesis

Application

Smoking Intensity on Medical Expenditures

NMES is a real dataset with 7896 observation from the 1987 National Medical Expenditure Survey. We investigate the relationship between smoking intensity and medical expenditures of smokers. The dataset consists of age, age to start smoking, gender, race, marital status, education level, census region, poverty status, and seat belt usage. The literature [2] suggested a positive causal effect using parametric models

Obesity on Risk of Depression

NHANES contains 12215 adults from the 2005–2016 National Health and Nutrition Examination Survey. We study the relationship between obesity and depression with gender, age, race, marital status, education level, family poverty-to-income ratio, smoking, and drinking alcohol. The literature [3] suggested a strong association ($p < 0.001$)

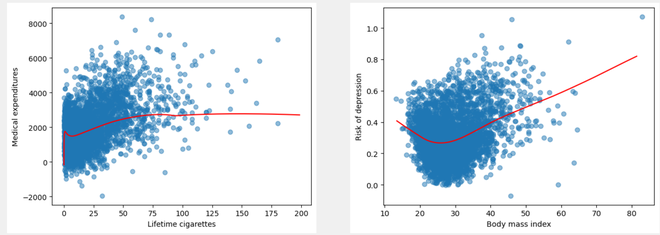

Figure 1. ADRF for NMES (left) and NHANES (right). Red lines represent the ADRF; blue dots represent the predicted potential outcomes

Dataset	One-split	K-fold	Westling L_1	Westling L_2	Westling L_∞	DRDRtest
NMES	0.031	0.021	0.005	0.008	0.04	0.06
NHANES	0.000	0.000	0.000	0.000	0.000	0.000

Table 2. p-value for real datasets

Results

In simulation, only our test successfully reserves a valid type-I error. We observe both existing methods tend to be too aggressive to reject the null when irrelevant covariates exist. In real datasets, we observe that all methods affirm with the findings of previous studies. It indicates the power of our test is comparable to that of the existing methods

References

[1] Charles R. Doss, Guangwei Weng, Lan Wang, Ira Moscovice, and Tongtan Chantarat.
A nonparametric doubly robust test for a continuous treatment effect, 2023.

[2] Kosuke Imai and David A van Dyk.
Causal inference with general treatment regimes.
Journal of the American Statistical Association, 99(467):854–866, 2004.

[3] Wen Ma, Zhengwei Yan, Wentao Wu, Daning Li, Shuai Zheng, and Jun Lyu.
Dose-response association of waist-to-height ratio plus bmi and risk of depression: Evidence from the nhanes 05-16.
Int J Gen Med, 14:1283–1291, April 2021.

[4] Lizhen Nie, Mao Ye, qiang liu, and Dan Nicolae.
Vcnet and functional targeted regularization for learning causal effects of continuous treatments.
In *International Conference on Learning Representations*, 2021.

[5] Ted Westling.
Nonparametric tests of the causal null with nondiscrete exposures.
Journal of the American Statistical Association, 117(539):1551–1562, 2022.