ETH zürich

# Distributional Instrumental Variable Method

**Anastasiia Holovchak**

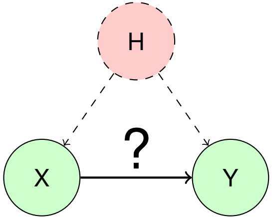
Joint work with Sorawit Saengkyongam, Nicolai Meinshausen and Xinwei Shen

15 April 2026, EuroCIM 2026, Oxford, UK

# Outline

1. Motivation
2. Setting & supporting theory
3. DIV method

# Motivation




What would happen if everyone were given treatment $X := x$?

Estimand: do-interventional distribution $P_Y^{\mathrm{do}(X:=x)}$

**D**istributional modelling + **IV** technique

# Instrumental Variable Setting

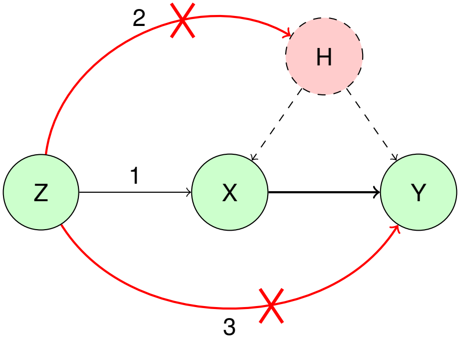



Figure: An example of a DAG compatible with IV model setting.

# Estimand

Traditional IV methods:

$$\mathbb{E}_Y^{\mathrm{do}(X:=x)}$$

Our target:

$$P_Y^{\mathrm{do}(X:=x)}$$

# **Engression**[1]: Distributional learning framework via generative modelling

Target: distribution of $Y|X = x$.

Build a **generative model** to describe the target distribution:

$$Y = g(X, \varepsilon),$$

where $\varepsilon \sim P_\varepsilon$ pre-defined (e.g., standard Gaussian), and map $g\colon (x, \varepsilon) \mapsto y$, parametrized by neural networks.

**Goal**

Find $g$, s.t. $g(x, \varepsilon) \sim P_{Y|X=x}$, allowing us to sample from $P_{Y|X=x}$.

[1] Shen & Meinshausen (2024), JRSS B

# Loss function

We use the **expected negative energy score**[2] as a loss function to train the conditional generative model:

$$\arg\min_{g\in\mathcal{G}} \mathbb{E}\left[\|Y - g(X,\varepsilon)\| - \frac{1}{2}\|g(X,\varepsilon) - g(X,\varepsilon')\|\right],$$

where $\varepsilon, \varepsilon' \sim \mathrm{N}(0,1)$.

- based solely on sampling
- allows for gradient-based optimization

[2]Gneiting & Raftery (2007)

# Outline

1. Motivation

2. Setting & supporting theory

3. DIV method

# Structural causal model

$$\begin{cases} X := g(Z, \eta_X) \\ Y := f(X, \eta_Y), \end{cases} \tag{1}$$

with $Z$ exogenous and independent of $(\eta_X, \eta_Y)$, while $\eta_X$ and $\eta_Y$ correlated due to hidden confounding between $X$ and $Y$.

$$X = g(Z, \eta_X)$$
$$Y = f(X, \eta_Y)$$

**Proposition 1: Identifiability of interventional distribution**

Consider the model in (1) and suppose the following assumptions hold:

1. For all $z \in \mathrm{supp}(Z)$, it holds that $g(z, \cdot)$ is strictly monotone.
2. For all $x \in \mathrm{supp}(X)$, $\mathrm{supp}(\eta_X | X = x) = \mathrm{supp}(\eta_X)$.

Then, for all $x \in \mathrm{supp}(X)$, the interventional distribution $P_Y^{\mathrm{do}(X:=x)}$ is **uniquely determined** from the observed data distribution $P_{(X,Y)|Z}$.

Estimate $P_{(X,Y)|Z}$ $\xrightarrow{\text{sufficient}}$ identify $P_Y^{\mathrm{do}(X:=x)}$

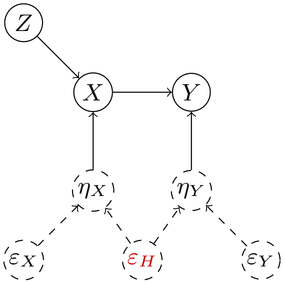
# Outline

# DIV model

$$X = g(Z, \eta_X)$$
$$Y = f(X, \eta_Y)$$

Joint generative model:

$$\begin{aligned} \eta_X &= h_X(\varepsilon_X, \varepsilon_H) \\ \eta_Y &= h_Y(\varepsilon_Y, \varepsilon_H) \\ X &= g(Z, \eta_X) \\ Y &= f(X, \eta_Y) \end{aligned} \tag{2}$$




WLOG assume $\varepsilon_H, \varepsilon_X, \varepsilon_Y \sim \mathrm{N}(0, 1)$.

# DIV solution

$$X = g(Z, h_X(\varepsilon_X, \varepsilon_H))$$
$$Y = f(X, h_Y(\varepsilon_Y, \varepsilon_H))$$

We define the population version of DIV solution as

$$\underset{f,g,h_X,h_Y}{\arg\min}\ \mathbb{E}\left[\|(X,Y) - (\hat{X},\hat{Y})\| - \frac{1}{2}\|(\hat{X},\hat{Y}) - (\hat{X}',\hat{Y}')\|\right], \qquad (3)$$

where

$$\hat{X} = g(Z, h_X(\varepsilon_X, \varepsilon_H)), \quad \hat{Y} = f(\hat{X}, h_Y(\varepsilon_Y, \varepsilon_H))$$
$$\hat{X}' = g(Z, h_X(\varepsilon'_X, \varepsilon'_H)), \quad \hat{Y}' = f(\hat{X}', h_Y(\varepsilon'_Y, \varepsilon'_H))$$

with $\varepsilon_X, \varepsilon_Y, \varepsilon_H, \varepsilon'_X, \varepsilon'_Y, \varepsilon'_H \sim \mathrm{N}(0,1)$ independently.

**Proposition 2 (informal)**

DIV solution induces $P_{(X,Y)|Z}$.

(+ Proposition 1) $\Rightarrow P_Y^{\mathrm{do}(X:=x)}$ can be uniquely identified using DIV approach!

# Estimation of the interventional distribution and its functionals

DIV solution $(f^*, h_Y^*)$ enables sampling from the interventional distribution:

$$f^*(x, h_Y^*(\varepsilon_Y, \varepsilon_H)) \sim P_Y^{do(X:=x)}, \qquad \forall x.$$

- Estimation of the interventional mean function

$$\mu^*(x) := \mathbb{E}_{\varepsilon_Y, \varepsilon_H}[f^*(x, h_Y^*(\varepsilon_Y, \varepsilon_H))]. \tag{4}$$

ATE: $\mu^*(x_1) - \mu^*(x_0)$

- Estimation of the interventional quantile function

$$q_\alpha^*(x) := Q_{\alpha;\varepsilon_Y, \varepsilon_H}[f^*(x, h_Y^*(\varepsilon_Y, \varepsilon_H))]. \tag{5}$$

QTE: $q_\alpha^*(x_1) - q_\alpha^*(x_0)$

# Illustrative example

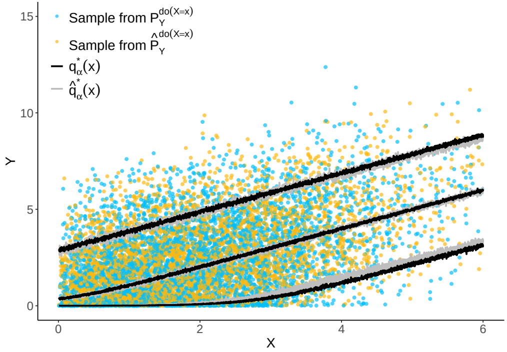



Figure: Samples from $P_Y^{\text{do}(X:=x)}$ (blue) and $\hat{P}_Y^{\text{do}(X:=x)}$ (yellow) along with interventional quantile functions $q^*_\alpha(x)$ and $\hat{q}^*_\alpha(x)$ for $\alpha \in \{0.1, 0.5, 0.9\}$

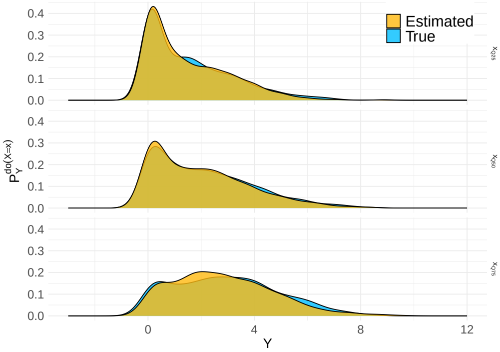



Figure: Kernel density estimates based on samples from $P_Y^{\text{do}(X:=x)}$ (blue) and $\hat{P}_Y^{\text{do}(X:=x)}$ (yellow), $1000$ samples per $x$

# Novelty of DIV

- Interventional distribution $P_Y^{\mathrm{do}(X:=x)}$ can be uniquely identified from $P_{(X,Y)|Z}$ under suitable assumptions
- Estimation is based on sampling
- $P_Y^{\mathrm{do}(X:=x)}$ is a more general estimand than interventional mean/quantiles
- Easy to extend to conditional interventional distribution: target $P_{Y|W=w}^{\mathrm{do}(X:=x)}$
- Can handle multivariate treatments, outcomes and covariates
- `R` package available on CRAN

# References

**Holovchak, A., Saengkyongam, S., Meinshausen, N., and Shen, X.** (2025). *Distributional Instrumental Variable Method*. arXiv:2502.07641

**Shen, X. and Meinshausen, N.** (2024). *Engression: Extrapolation through the Lens of Distributional Regression*. Journal of the Royal Statistical Society Series B. 10.1093/jrsssb/qkae108