



**UNIVERSITÀ
DEGLI STUDI
DI TRIESTE**

From Latent to Deep Latent Variable Models

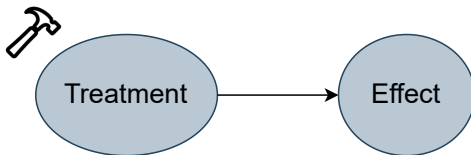
A Study On Causal Effect Inference with CEVAE

Valeria De Stasio, Christian Faccio, Giovanni Lucarelli

June 17, 2025

Objective: estimating causal effects

Estimate how a medical treatment T affects the health Y of a *random* patient.



$$P(Y = y | \text{do}(T = t)) = P(Y = y | T = t)$$

$$\text{ITE} = \mathbb{E}[(Y | \text{do}(T = 1))] - \mathbb{E}[(Y | \text{do}(T = 0))]$$

This is usually the case in a **randomized controlled trial** (RCT), where the treatment is randomly assigned to the patients.

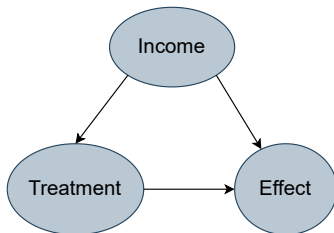
Confounder

In an observational study:

- **no** control over the treatment T assignment
→ there may be a confounder X (e.g. Income) that influences both variables

$$P(Y|\text{do}(t)) \neq P(Y|t)$$

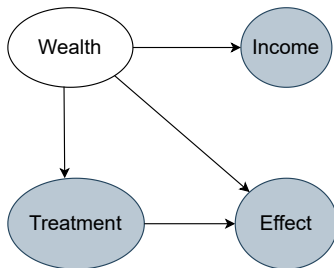
$$P(Y|\text{do}(t), x) = P(Y|t, x)$$



If the confounder is observed even a linear regression can do the job!

Latent confounder

But what if the confounder Z (e.g. Wealth) is **not** observed and the observed X is only a proxy of it?



Idea: estimate the latent variable and condition on it

Vanilla Latent Variable Model

- Assume parametric distributions
- Assume parametric relationships between variables

- Infer parameters using Stochastic Variational Inference (SVI)

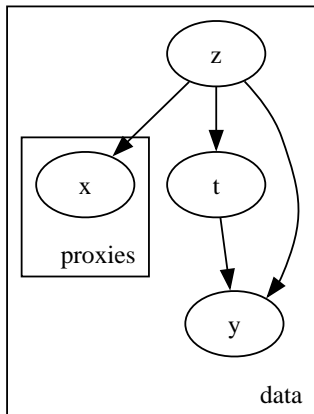
Problem: to infer from new data point x we need to train a new model at test time!

Deep Latent Variable Model: CEVAE

- Assume parametric distributions
- Assume parametric relationships between variables through
Neural Networks

No need for test time training! **amortized inference**

Synthetic linear dataset



$$Z \sim \mathcal{N}(0, 1)$$

$$a_j \sim \mathcal{U}(-10, 10)$$

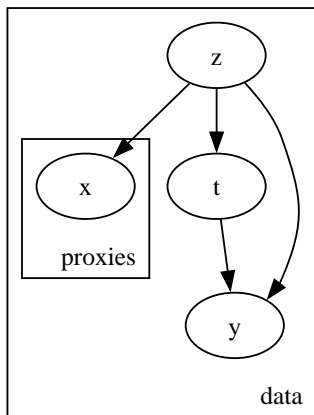
$$X_j \sim \mathcal{N}(a_j Z, \sigma_X^2)$$

$$T|Z \sim \text{Be}(\sigma(\beta Z))$$

$$Y|T, Z \sim \mathcal{N}(z + t, \sigma_Y^2)$$

Remark: ITE is constant for all x !

Synthetic non linear dataset



$$Z \sim \mathcal{N}(0, 1)$$

$$a \sim \mathcal{U}(-10, 10)^d$$

$$\Sigma = \sigma_X^2[(1 - \rho)\mathbb{I} + \rho J]$$

$$X_1, \dots, X_d \sim \mathcal{N}(a \tanh(z), \Sigma)$$

$$T|Z \sim \text{Be}(\sigma(\beta z))$$

$$Y|T, Z \sim \mathcal{N}(\text{NonLin}(z, t), \sigma_Y^2)$$

$$\text{NonLin}(z, t) = \sin(z) + \frac{1}{2}z + t \left(1 + \frac{1}{2}z\right)$$

boxplot1

Results (2/2): Non Linear Dataset

boxplot2

Experiment: latent distribution misspecified

Experiment: changing latent dimension

Experiment: increasing the treatment effect

Conclusions

Thank You!