UNIVERSITÀ DEGLI STUDI DI TRIESTE

Causal Effect Variational Autoencoder

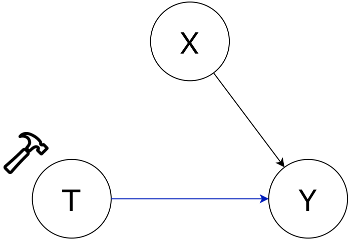
Final Exam Project

Valeria De Stasio, Christian Faccio, Giovanni Lucarelli

June 17, 2025

Objective: estimating causal effects

To estimate how a treatment T affects an outcome Y, we would like to intervene on T for *random* subjects and observe the effect on Y. This corresponds to a Randomized Clinic Trial.

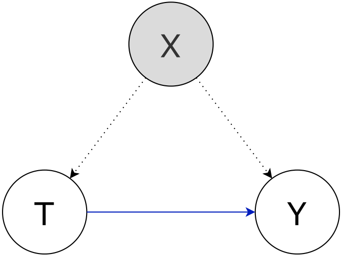


$$P(Y|\text{do}(T=t)) = P(Y|T=t)$$

Confounder

In a observational study we do **not** have control over T: there may be a confounder X that influences both variables!

If we observe all confounders X, then we can block the backdoor path and compute the causal effect of T on Y.

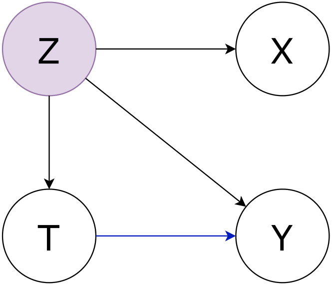


$$P(Y|\mathrm{do}(T=t)) = \sum_x P(Y|T=t, X=x)P(X=x)$$

Latent confounder

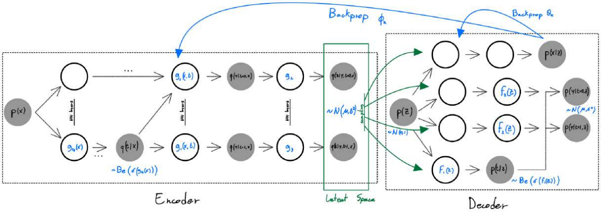
But what if there is a confounder Z that is **not** observed? This is the most likely scenario in which X is only a proxy for Z.

This time the back-door criterion is not satisfied: we cannot compute the causal effect of T on Y.



Causal Effect Variational Autoencoder

We introduce a Variational Autoencoder (VAE) to learn the latent space from the observed variables.



We can now estimate both the causal effect of T on Y and the expected outcome Y if we force a treatment $do(t)$ on a certain subject.

Synthetic dataset

Experiment: noise in the proxies

Experiment: latent distribution misspecified

Experiment: changing latent dimension

Experiment: increasing the treatment effect

Conclusions

Thank You!