

# Semiparametric Generative Invariance

Carlos García Meixide, David Ríos Insua  
Instituto de Ciencias Matemáticas

April 5, 2024

**Abstract:** This work enhances the current statistical comprehension of causality by demonstrating that domain generalization is achievable without heterogeneous data sources and without specific assumptions regarding the strength of perturbations or support overlap. We present a new estimator for predicting outcomes in different distributional settings under hidden confounding without relying on instruments or exogenous variables. The population definition of our estimator identifies causal parameters belonging to certain complexity classes, whose empiricals lead to a generative model capable of replicating the true probability law of the outcome given the covariate distribution at test stage beyond (suboptimal) do-interventional conditionals. We show that the probabilistic alignment between our proposed method and true test distributions is uniformly the best across various interventions.