

# Robust Offline Policy Learning Under Covariate Distribution Shifts

We study the problem of learning a treatment assignment policy based on observable covariates, where there are potential shifts in the distribution of covariates from historical data (training) to deployment (test). We formulate a distributional robust policy optimization problem with the objective of maximizing the worst-case (out-of-sample) expected outcomes, considering all possible distributions of future data within an ambiguity set. We construct the ambiguity set as a variant of the type-1 Wasserstein ball centered at the empirical distribution of the historical data, explicitly requiring that only the covariate distribution can change. Using standard duality techniques, we reformulate the problem as an infinite linear program. For the case of two treatments, we leverage an interpolation technique recently introduced in the newsvendor context to characterize the optimal solution. For settings with more than two treatments, we propose a solution approach inspired by this technique: we construct an in-sample policy and iteratively make assignment decisions as new data becomes available, with the objective of minimizing the optimality gap. We conduct numerical experiments to evaluate the performance of our proposed method.

**Authors:** TANG, Bill (University of Southern California); KOCYIGIT, Cagil (University of Luxembourg); VAYANOS, Phebe (University of Southern California)

**Orateur:** KOCYIGIT, Cagil (University of Luxembourg)

**Classification de Session:** ML

**Classification de thématique:** Machine learning