

## Regression Models for Distributional Data and Time Series

We present two regression models for the rapidly evolving field of distributional data analysis (DDA) for one-dimensional probability distributions and density data, working in the Wasserstein space of continuous distributions. These regression models feature distributional predictors and responses and include the case of i.i.d. predictor/response distributions as well as autoregressive models for distributional time series. The Wasserstein regression model [1] utilizes the Wasserstein manifold and parallel transport, mapping predictor and response distributions to suitable tangent spaces, whereupon any preferred functional regression model is applied in these Hilbert spaces, followed by projecting to a convex invertibility set and mapping back to the Wasserstein space. The optimal transport regression model and its autoregressive version (ATM) [2] is based on a transport algebra that we introduce and relates optimal transports that serve as predictors and responses. This model does not require a projection step and operates on transport geodesics; it is thus intrinsic. Illustrations include age-at-death distributions and financial distributional data. This presentation is based on joint work with Yaqing Chen (Davis), Zhenhua Lin (Singapore) and Changbo Zhu (Davis).

Key words: Autoregressive Transport Model, Distributional Data Analysis, Optimal Transport, Tangent Bundle, Wasserstein Space

### References

- [1] Chen, Y, Lin, Z, Müller, HG (2020). Wasserstein Regression. arXiv:2006.09660
- [2] Zhu, C, Müller, HG (2021). Autoregressive Optimal Transport Models. arXiv:2105.05439.