Studying Earnings Trajectories as Functional Outcomes

Kreske Ecker

Akademisk avhandling

som med vederbörligt tillstånd av Rektor vid Umeå universitet för avläggande av filosofie doktorsexamen framläggs till offentligt försvar i Hörsal HUM.D.220, Humanisthuset, fredagen den 9:e februari, kl. 10:00.

Avhandlingen kommer att försvaras på engelska.

Fakultetsopponent: Dr. Jose Peña, Institutionen för datavetenskap, Enheten för Statistik och maskininlärning, Linköpings universitet, Sverige.
In this thesis, we present methods for studying patterns of income- and earnings accumulation over time using functional data analysis. This is made possible by the availability of large-scale longitudinal register data in Sweden. By modelling individuals' cumulative lifetime earnings trajectories as continuous functions of time, we can explore temporal dynamics as well as divergences in these trajectories based on initial labour market conditions. A major contribution of this thesis consists of extending the potential outcome framework for causal inference to functional data analysis.

In Paper I, we use functional-on-scalar linear regression and an interval-wise testing procedure to study the associations between initial labour market size and income trajectories for one Swedish birth cohort. In Paper II, we present methods to draw causal conclusions in this setting. We introduce the functional average treatment effect (FATE), as well as an outcome-regression based estimator for this parameter. In addition, we show the finite sample distribution of this estimator under certain regularity conditions and demonstrate how simultaneous confidence bands can be used for inferences about the FATE. An application study in this paper estimates the causal effect of initial labour market size on income accumulation trajectories.

In Paper III, these methods are applied to study the effect of initial firm age on earnings accumulation. Paper IV presents an outcome regression based and a double robust estimator for the mean of functional outcomes when some of these outcome functions are missing at random. We derive the asymptotic distributions of these two estimators as well as their covariance structure under more general conditions.

Keywords
functional data analysis, causal inference, earnings trajectories, simultaneous confidence bands, missing data.