Hierarchical Functional Data: Semiparametric and Nonparametric Methods for Modeling Functional Dependence, With Application to Colonic Crypt Signaling

Raymond J. Carroll
Texas A&M University

We consider hierarchical functional data experiments that arise naturally in our colon carcinogenesis experiments. The data are at three levels: individual rats have multiple colonic crypts, and within each colonic crypt, multiple measurements of biomarkers are made depending on the location of cells. In such experiments, it is typical to assume that conditional on the individual, the functions at the crypt level are independent. However, the biology suggests that the functions probably are not independent conditional on the individual, and it is of considerable interest to understand whether this is true and, if so, to quantify the degree of dependence.

This problem is an example of a very general class of semiparametric problems that have a nonparametric component that is evaluated multiple times at the individual level, as in for example longitudinal data, clustered data, and matched studies. We provide the general framework, exhibit the semiparametric efficient solution, and describe computationally convenient variants. We then describe the analysis of p27 marker data (p27 is a predictor of programmed cell death). In our first approach, the multiple functions for each individual are modeled through a low-order penalized spline regression with a separable correlation structure. We exhibit frequentisi and Bayesian analyses of these problems, analyses that suggest surprisingly strong functional correlations. We also provide an alternative, completely nonparametric approach.