Abstract

In this paper we are interested in capturing heterogeneity in clustered or longitudinal data. Traditionally such heterogeneity is modeled by either fixed effects or random effects. In fixed effects models, the number of degree of freedom for the heterogeneity equals the number of clusters/subjects minus 1, which could result in less efficiency. In random effects models, the heterogeneity across different clusters/subjects is described by e.g., a random intercept with 1 parameter (for the variance of random intercept), which could lead to oversimplification and biases (shrinkage estimates). Our "fusion effects" model stands in between these two approaches: we assume that there are unknown number of different levels of heterogeneity, and use the fusion penalty approach for estimation and inference. We evaluate and compare the performance of our method to the fixed and random effects models by simulation studies. We apply our method to the Ocular Hypertension Treatment Study (OHTS) to capture the heterogeneity in the progression rate toward primary open-angle glaucoma of left and right eyes of different subjects.