Summary

by Runze Li

In Ecology Momentary Assessment (EMA) study, data are often highly unbalanced because data were collected at irregular and possibly subject-specific time points. Due to their unbalanced nature, it is difficult to directly apply traditional multivariate regression techniques for analyzing such data. We have extended various modeling tools to analysis of EMA data.

Perhaps the most prominent and theoretically important situational association with smoking is that between smoking and mood. Both negative and positive affect have been proposed as important triggers for smoking, and most theories of smoking emphasize the role of affect in driving smoking. Using ecological momentary assessment, Shiffman, *et al.* (2002) found that after controlling for smoking restrictions, smoking was strongly related to smoking urges and other several activities, but it unrelated to negative or positive affect, or to arousal. Natural questions arising here is whether negative or positive affect and arousal have impacts on the urge to smoke. To address this questions, we investigated the relationship between the urge to smoke and affect & arousal. In the paper entitled "Functional hierarchical linear models with application to analysis of Ecology Momentary Assessment Data", we performed cluster analysis for scores summarized from the EMA data. The collected samples were classified into 4 clusters. The overall trends of different clusters are very different from each other. Cluster 1 consists of 113 subjects, who made good progress during the post-quit period. Cluster 2 consists of 86 subjects, who made slight progress during post-quit period. Cluster 3 consists of 31 subjects, who made no progress during post-quit period. Cluster 4 consists of 23 subjects who made great progress after quitting smoking.

In this paper, we employed functional linear models to identify variables of potential impacts on urge to smoke. We found that the affect, arousal and attention disturbance have

only linear and time-dependent effect on the urge to smoke. All interaction and quadratic effects are statistically insignificant. We further fitted the EMA data using functional mixed effect models. It was found that the negative effects has significant impacts on the urge to smoke, and its effect is time-varying; the effects of arousal is also time-varying, and its pattern has dramatically changes after quitting smoking. The effects of attention disturbance seems to be periodic and vary around at a constant. All of these findings are consistent with those using functional linear models. By using functional mixed effects models, we further discovered that the effects of affect and arousal vary from subject to subject. The variance functions for random effects and random error are also time-varying and significant away from zero.

Analysis of time to event data will be potentially useful for analysis of ecology momentary data as lapse and relapse are important events of interest. Model selection and variable selection are fundamental in statistical modeling. In the paper entitled "Variable- and model-selection for multivariate failure time data", we proposed a class of variable selection procedures via penalized pseudo-partial likelihood. We further demonstrate that the proposed method has some nice asymptotic properties. We show that, for certain penalty functions with proper choices of regularization parameters, the resulting estimate is root n consistent and possesses an oracle property, namely, the resulting estimate can correctly identify the true model as if the true model (the subset of variables with nonvanishing coefficients) were known in advance. Using a simple approximation of the penalty function, the proposed method can be easily carried out with the Newton-Raphson algorithm. We conducted extensive Monte Carlo simulation studies to assess the finite sample performance of the proposed procedures. The proposed method was illustrated by analyzing a data set from the Framingham Heart Study.