Asymptotic Properties of Partial Likelihood Inference for Imprinting and Maternal Effects

Fangyuan Zhang[1*], Abbas Khalili[2*], Shili Lin[1]
[1]Department of Statistics, The Ohio State University, Columbus, OH;
[2]Department of Mathematics and Statistics, McGill University, Montreal, Canada;
[*] Co-First Authors.

ABSTRACT

Both genomic imprinting and maternal effect, as causes of parent-of-origin patterns in complex human diseases, are increasingly explored. Because they are confounded with each other, tests attempting to detect only one of the effects while ignoring the other one may have false positive result. While some methods do model imprinting and maternal effect simultaneously to avoid potential confounding, they rely on strong yet unrealistic assumptions about the population, such as mating symmetry, to avoid over-parameterization with the typical data type, child-parents triads. LIME is a recent *L*ikelihood inference method for detecting *I*mprinting and *M*aternal *E*ffects simultaneously using nuclear families with an arbitrary number of affected and unaffected children. As it uses only part of the full likelihood - partial likelihood - by exploiting the fact the part of the likelihood containing the parameters of interest can be separated from that containing the nuisance parameters, LIME alleviates the need to make typically unrealistic assumptions and leads to a robust procedure with potentially greater power. Although simulation results seem to indicate that the parameters are well estimated, asymptotic behaviors of the procedure have not been studied theoretically or empirically. Here we first use extensive simulations under a variety of settings and different sample sizes to show that the estimates get closer to normal distributions centered around the true parameter values as sample size increases. Inspired by this observation, we then derived the theoretical properties of LIME showing that the estimates are indeed consistent and asymptotically normally distributed. We further show, through simulation, that the difference between theoretical variance-covariance matrix and its empirical counter part diminishes as the sample size increase. Finally based on the theoretical properties, we compare the information contents of 9 data types and make recommendations on designing an experiment if there is a fixed amount of resources.