Large-scale inference for chi-squared data: an empirical Bayes method

LILUN DU

Hong Kong University of Science and Technology

E-mail: dulilun@ust.hk

In a thought-provoking paper, Efron (2011) investigated the merit and limitation of an empirical Bayes method to correct selection bias based on Tweedie's formula first reported by Robbins (1956). The exceptional virtue of Tweedie's formula for the normal distribution lies in its representation of selection bias as a simple function of the derivative of log marginal likelihood. Since the marginal likelihood and its derivative can be estimated from the data directly without invoking prior information, bias correction can be carried out conveniently. We propose a Bayesian hierarchical model for chi-squared data such that the resulting Tweedie's formula has the same virtue as that of the normal distribution. Because the family of noncentral chi-squared distributions, the common alternative distributions for chi-squared tests, does not constitute an exponential family, our results cannot be obtained by extending existing results. Furthermore, the corresponding Tweedie's formula manifests new phenomena quite different from those of the normal distribution and suggests new ways of analyzing chi-squared data. This is a joint work with Inchi HU from Hong Kong University of Science and Technology.