Model Averaging of Count Data Regressions

Kevin Huynh*

Preliminary version. Please do not cite or circulate.

Abstract

Model averaging methods have become an increasingly popular tool for improving predictions and dealing with model uncertainty, especially in Bayesian settings. More recently, frequentist model averaging methods such as information theoretic and least squares model averaging have emerged. In this work, I focus on the issue of covariate uncertainty where managing the computational resources is key: The model space grows exponentially in the number of covariates such that averaged models must often be approximated. Weighted average least squares (WALS), introduced by Magnus et al. (2010, *Journal of Econometrics*) for the Gaussian linear regression model, combines Bayesian and frequentist aspects and additionally employs a semiorthogonal transformation of the regressors to reduce the computational burden. This paper extends WALS for generalized linear models (De Luca et al., 2018, *Journal of Econometrics*) to the negative binomial (NB) regression model for overdispersed count data. The predictive power of WALS for NB regression is compared to traditional estimators in a simulation experiment and in an empirical application using data on doctor visits.

Keywords: WALS, Model Averaging, Negative Binomial Regression, Count Data

*University of Basel, Faculty of Business and Economics, Switzerland, E-Mail: kevin.huynh@unibas.ch