

ASYMPTOTIC PERFORMANCE OF CROSS-VALIDATION METHODS IN LINEAR REGRESSION

A.K. Uludag[†], O. Saracbasi

Hacettepe University, Ankara, Turkey

[†] E-mail: *keremuludag@gmail.com*

Cross-Validation (CV) methods are widely used in order to evaluate the performance of predictive models. It is shown that under fixed-dimensioned model assumption, MCCV method is much more powerful than LOOCV method regarding the selection of the true model (Shao, 1993). The aim of this study is to evaluate and compare the asymptotic performance of these methods under the same assumption with Shao (1993) in the context of linear regression for different data structures; sample sizes ($n=40, 70$), error variances ($\sigma^2 = 0.1, 0.25, 1$). Additionally, the correlation structures of the variables included in the actual model were also considered. In this regard, an algorithm was developed for the simulation study and the methods were investigated for different model sizes. Performances of the methods were evaluated according to the probability of selecting the true model and overall mean errors. In conclusion, it was found that performances of the methods were dramatically decreased in the case of small sample size, high error variance and low correlation between one of the independent variables and the dependent variable. Especially low correlation was found to be the most influential factor affecting the performance of the methods. In addition, despite the findings of Shao (1993), it was observed that MCCV method was not better than LOOCV method in certain situations for selection of the true model.