Bayesian information criteria in misspecified models: validity and consistency

Yoichi MIYATA*

Takasaki City University of Economics

Abstract. Model selection has an important role in modern statistical analysis. Although the Bayesian information criterion (BIC) of Schwarz (1978) is derived by approximating the log-marginal likelihoods of models, for the validity of the approximation, we usually assume that a class of parametric models includes a correctly specified model. Lv and Liu (2014) derive asymptotic expansions for the marginal likelihoods of misspecified generalized linear models (GLMs) under some reasonable conditions, and proposes two kinds of generalized BICs called GBIC and GBIC_p. In this talk, we derive a higher-order asymptotic expansion for the marginal likelihoods under conditions similar to those of Lv and Liu, and present alternative BICs, which modifies the penalty term of GBIC. In addition, we apply the approach of Sin and White (1996) to present some consistency results of the foregoing information criteria in GLMs with non-compact parameter spaces and deterministic design matrices. We also present several numerical examples to illustrate the finite-sample performance of the foregoing BICs in both correctly specified and misspecified linear and logistic models.

References

[1] Schwarz, G. (1978) Estimating the dimension of a model. *Ann. Statist.*, 6, 461-464.

[2] Lv, J. and Liu, J. S. (2014) Model selection principles in misspecified models. *J. R. Statist. Soc.* B, **76**, Part 1, 141-167.

[3] Sin, C.-Y., and H. White. (1996) Information Criteria for Selecting Possibly Misspecified Parametric Models, *Journal of Econometrics*, 71(1-2), 207-225.