Karagrigoriou, Alexandros (University of Cyprus, Cyprus) and Kyriacos Mattheou (University of Cyprus, Cyprus)

A Density Power Divergence Based Robust Model Selection Criterion.

Abstract: The aim of this work is the development of a new robust model selection criterion using a general discrepancy based technique. A model selection criterion can be constructed as an approximately unbiased estimator of an expected 'overall discrepancy' (or divergence), a nonnegative quantity which measures the 'distance' between the true model and a fitted approximating model. A well known discrepancy is the Kullback-Leibler discrepancy that was used by Akaike (1973) to develop AIC. A new discrepancy was recently introduced by Basu (1998). In this work we develop a new model selection criterion which is an approximately unbiased estimator of the expected overall power divergence (discrepancy) between the true and the fitted models that corresponds to Basus's density power divergence.

Note that the family of divergences introduced by Basu is indexed by a single parameter "a" which controls the trade-off between robustness and asymptotic efficiency of the parameter estimators which are the minimizers of this family of divergences. When $a \to 0$, Basu's density power divergence becomes the Kullback-Leibler divergence and the method is the maximum likelihood estimation; when $a = 1$, the divergence is the -distance and a robust but inefficient minimum estimator ensues. The expected overall discrepancy is evaluated and an approximately unbiased estimator is derived which depends on the parameter a.

We are most interested in small values of $a > 0$, say between zero and one, although values greater than one can be considered too. There can be no universal way of selecting the appropriate value of the parameter when applying our estimation methods. The parameter "a" specifies the underlying distance measure and typically dictates to what extent the resulting methods become statistically more robust than the maximum likelihood methods, and should be thought of as an algorithmic parameter. One way of selecting it is to fix the efficiency loss, at the ideal parametric model employed, at some low level, like five or ten percent. Other ways could in some practical applications involve prior motions of the extent of contamination of the model.

The proposed criterion could be used in applications where outliers or contaminated observations are involved. Some applications include medical and epidemiological surveys. Such cases are discussed here. The results show that DIC has a tendency of underestimation by selecting the true model as

well as smaller models in contrast with AIC which overestimates the true model.