On universal principles and model choice

Laurie Davies
Department of Mathematics
University of Duisburg-Essen

February 13, 2004

Abstract

Abstract: A universal principle of model choice is one that does not take into account the specific nature or subject of the data. General principles are Bayes, Minimum Description Length, AIC, BIC etc. Given a data set $\boldsymbol{x}_n$ and a family of stochastic models $\mathcal{P}$ such a principle will select a model $P \in \mathcal{P}$ which is judged to be the best model within the class $\mathcal{P}$ for the data. In general such principles are based on a fidelity term which measures the fit of the given model to the data and either a penalty term which is a measure of the complexity of the data or some prior valuation of the models. The fidelity term consists of single number and is usually based on the likelihood with respect to the given model. It is argued that the fit of a model to a data set is too complex to be reduceable to a single number. Certain aspects of the fit may be of more importance than others and which aspects are to be considered will depend on the nature of the data. There is no reason for supposing that the likelihood will respect all these different aspects. It is further argued that such principles cannot explain the practice of statistics as the decision is based on numbers which given no reason to be satisfied or dissatisfied with the model chosen. In conclusion universal principles of model choice can only be recommended when the statistician has no idea of what to do.