# COUNTEREXAMPLES TO PARSIMONY AND BIC

DAVID F. FINDLEY

*Statistical Research Division, U.S. Bureau of the Census, Washington, D.C. 20233, U.S.A.*
*Institute of Statistical Science, Academia Sinica, Taipei 11529, Taiwan*

(Received February 6, 1990; revised April 1, 1991)

**Abstract.** Suppose that the log-likelihood-ratio sequence of two models with different numbers of estimated parameters is bounded in probability, without necessarily having a chi-square limiting distribution. Then BIC and all other related "consistent" model selection criteria, meaning those which penalize the number of estimated parameters with a weight which becomes infinite with the sample size, will, with asymptotic probability 1, select the model having fewer parameters. This note presents examples of nested and non-nested regression model pairs for which the likelihood-ratio sequence is bounded in probability and which have the property that the model in each pair with *more* estimated parameters has better predictive properties, for an independent replicate of the observed data, than the model with fewer parameters. Our second example also shows how a one-dimensional regressor can overfit the data used for estimation in comparison to the fit of a two-dimensional regressor.

*Key words and phrases*: Model selection, linear regression, misspecified models, AIC, BIC, MDL, Hannan-Quinn criterion, overfitting.