Robust Second-Order Accurate Inference for Generalized Linear Models

E. Ronchetti

[1] Department of Econometrics, University of Geneva, Blv. Pont d'Arve 40, CH-1211 Geneva, Switzerland

Keywords: Empirical likelihood, M-estimators, Robust tests, Saddlepoint techniques

Abstract

Generalized linear models have become the most commonly used class of regression models in the analysis of a large variety of data. In particular, generalized linear model can be used to model the relationship between predictors and a function of the mean of a continuous or discrete response variable.

The estimation of the parameters of the model can be carried out by maximum likelihood or quasi-likelihood methods, which are equivalent if the link is canonical. Standard asymptotic inference based on likelihood ratio, Wald and score test is then readily available for these models. However, two main problems can potentially invalidate p-values and confidence intervals based on standard classical techniques.

First of all, the models are ideal approximations to reality and deviations from the assumed distribution can have important effects on classical estimators and tests for these models (nonrobustness). Secondly, even when the model is exact, standard classical inference is based on (first order) asymptotic theory. This can lead to inaccurate p-values and confidence intervals when the sample size is moderate to small or when probabilities in the extreme tails are required.

The nonrobustness of classical estimators and tests for the parameters is a well known problem and alternative methods have been proposed in the literature. These methods are robust and can cope with deviations from the assumed distribution. However, they are based on first order asymptotic theory and their accuracy in moderate to small samples is still an open question.

In this paper we propose a test statistic which combines robustness and good accuracy for small sample sizes. We combine results from Cantoni and Ronchetti (2001) and Robinson, Ronchetti and Young (2003) to obtain a new test statistic for hypothesis testing and variable selection which is asymptotically χ^2 distributed as the three classical tests but with a *relative error* of order $O(n^{-1})$. Moreover, the accuracy of the new test statistic is stable in a neighborhood of the model distribution and this leads to robust inference even in moderate to small samples.

Joint work with S. N. Lo.

References

E. Cantoni and E. Ronchetti (2001). Robust Inference for Generalized Linear Models. *Journal of the American Statistical Association*, 96, 1022-1030.

J. Robinson, E. Ronchetti and G.A. Young (2003). Saddlepoint Approximations and Tests Based on Multivariate M-estimates. *The Annals of Statistics*, 31, 1154 - 1169.