DIAGNOSTICS IN COUNT DATA MODELS

Sibnarayan Guria; *West Bengal State University, Kolkata, India.*

Abstract

There are primarily two reasons for the paucity of diagnostic studies beyond the linear model. First, most diagnostic techniques exploit the nature of the relationship between the response and the explanatory variables thus making it easier to deal with the linear form of relationship. Second reason is the closed form solution that the normal equations of a linear model provides. Non-linearity requires some form of iteration and hence makes the diagnostic measures difficult to derive algebraically.

Unlike the least-squares estimation method used in the classical linear model, the generalized linear models use the maximum likelihood method of parameter estimation which, as Pregiborn (1981) points out, is extremely sensitive to outlying responses and extreme points in the design space. However, since the likelihood equations are not in a closed form and some iterative techniques are required to obtain the estimates it is difficult to derive diagnostics to identify such observations.

In this paper a log-linear linear model is considered and show that the deletion technique can be used to study the diagnostics of such a model. As is well-known in the literature, the convergence of the likelihood solution is extremely rapid for such models and hence we base our study on a one-step approximation of the likelihood estimates. The model is fitted using the maximum likelihood method and the deletion of observation technique is used to identify outliers. Expressions for the change in the estimates of the parameters after re-fitting are obtained.

In this paper all the computations are done in R, Using simplicity and

beautiful nature of Java programming language, input data files and the output are manipulated and represented in a lucid and meaningful way. Using the Java port to R - "Java2R" library (http://www.sngforge.co.nr/projects/java2r) any R-function can be accessed from Java.

Tools : Java2R

Key Words : DFBETA, DFFIT, Log-Linear model.