A SCORE TEST FOR OVERDISPERSION BASED ON GENERALIZED POISSON MODEL

Z. Yang[†], J.W. Hardin, C.L. Addy

University of South Carolina, Columbia, USA

[†] E-mail: *tonyyangsxz@gmail.com*

Poisson model is the standard tool for count data analysis. A property of the Poisson model is equi-dispersion, that the mean and variance are equal. In practice it is often found that the data exhibit greater variability than is predicted by this implicit mean-variance relationship. This phenomenon is called over-dispersion, and has been widely considered in the literature (Dean and Lawless 1989, Dean 1992). Failure to take account of this over-dispersion leads to serious underestimation of standard errors and misleading inference for the regression parameters. Consequently, a number of models and associated estimation methods have been proposed for handling overdispersed data, including negative binomial and mixed Poisson regression models(Lawless, 1987; Dean and Lawless, 1989), Poisson mixture regression model(Wang, Cockbum and Le, 1996), and quasi-likelihood and random-effect models(Ozmen, 2000, Lee and Nelder, 2000) to deal with the latent heterogeneity.

The presence of overdispersion leads one to consider more complicated count distributions. The generalized Poisson (GP) distribution, introduced by Consul and Jain (1973) can be used to model overdispersed count data. Analysts must justify that the GP model is better than the general Poisson model. A Likelihood ratio test (LRT) or Wald test can be used for the justification, but the Score test has the advantage that we need fit only the Poisson model.

In this paper, we first review some Score test for over-dispersion in Poisson regression, then derived the score test for over-dispersion based on the generalized Poisson (GP) model, and we compare the power of proposed score test with the likelihood ratio test and Wald test. In the simulation study, we illustrate that the score test is more appropriate for testing the difference between the Poisson model and GP model.

Finally, we show an example by using Los Angeles School Dataset (Phil Ender, UCLA), and make some comparison between GP model and negative binomial model. And we conclude the GP model can be used to model some overdispersed data, as well as negative binomial model.