Li, R. and Liang, H. (2008), Variable Selection in Semi-Parametric Regression Modeling, *Annual Statistics*, **36**, 261-286.

Lim, H. K., Li, W. K. and Yu, P. L. H. (2014), Zero-Inflated Poisson Regression Mixture Model, *Computational Statistics and Data Analysis*, **71**, 151-158.

McLachlan, G. J. and Peel, D. (2000), *Finite Mixture Models*, John Wiley, New York.

Ormoz, E. and Eskandari, F. (2016), Variable Selection in Finite Mixture of Semi-Parametric Regression Models, *Communications in Statistics-Theory and Methods*, **45**, 695-711.

Ruppert, R., Wand, M. and Carroll, R. (2003), *Semi-Parametric Regression*, Cambridge University Press, Cambridge.

Tibshirani, R. (1996), Regression Shrinkage and Selection via the LASSO, *Journal of Royal Statistical Society*, **58**, 267-288.

Yatchew, A. (2003), *Semi-Parametric Regression for the Applied Econometrician*, Cambridge University Press, Cambridge.

Zou, Y., Zhang, Y. and Lord, D. (2013), Application of Finite Mixture of Negative Binomial Regression Models with Varying Weight Parameters for Vehicle Crash Data Analysis, *Accident Analysis and Prevention*, **50**, 1042-1051.

# A New Approach of Variable Selection in Finite Mixture of Semi-Parametric Regression Models with Poisson Distribution

**Heidari, M. and Eskandari, F.**

Department of Statistics, Allameh Tabataba'i University, Tehran, Iran.

**Abstract:** In this paper the issue of variable selection with new approach in finite mixture of semi-parametric regression models is studying, although it is supposed that data have Poisson distribution. When we use Poisson distribution, two problems such as overdispersion and excess zeros will happen that can affect on variable selection and parameter estimation. Actually parameter estimation in parametric component of the semi-parametric regression model is done by penalized likelihood approach. However, in nonparametric component after local approximation using Teylor series, the estimation of nonparametric coefficients along with estimated parametric coefficients will be calculated. Using new approach leads to a properly variable selection results. In addition to representing related theories, overdispersion and excess zeros are considered in data simulation section and using EM algorithm in parameter estimation leads to increase the accuracy of end results.

**Keywords:** EM algorithm, Overdispersion, Excess zeros, Semi-parametric regression,

Finite mixture model.

**Mathematics Subject Classification (2010): 62J07، 62J12.**