

**SEMIPARAMETRIC POISSON REGRESSION MODEL
FOR CLUSTERED DATA**

by

Eiffel A. de Vera

A thesis submitted in partial fulfillment
of the requirements for the degree of

Master of Science (Statistics)

School of Statistics
University of the Philippines
Diliman, Quezon City

March 2010

ABSTRACT

A spatially clustered count data is modeled through a semiparametric Poisson regression model. There is a heterogeneous covariate effect across the clusters. In addition, a random clustering effect is added into the model. To avoid overparametrization, the heterogeneous covariate effect is modeled nonparametrically. Two estimation procedures are proposed: (1) the parametric and nonparametric parts are estimated simultaneously through a penalized least squares method; and (2) the parametric and nonparametric parts are estimated iteratively in a backfitting framework. The simulation study confirmed the advantages of these methods over ordinary Poisson and linear models (with transformation) when the aggregate covariate effect is negligible, i.e., sensitivity to the covariate is minimal or the data-generating model is not linear. In general, the two methods are advantageous over the traditional approaches when linear model fit is poor. However, in cases where there is a good linear fit, the proposed methods are at par with the traditional methods, but they can still be advantageous when there are several covariates involved since backfitting can induce computational simplicity in the estimation process.

Keywords: *Poisson Regression; Clustered Data; Nonparametric Regression; Backfitting; Random Effects; Generalized Additive Models*