# SEMIPARAMETRIC PRINCIPAL COMPONENTS POISSON REGRESSION ON CLUSTERED DATA

by:

**Kristina Celene M. Manalaysay**

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 2011

## ABSTRACT

This study aims to address two issues in modeling count data: clustering of observations and interdependency of predictors. We propose to use principal components of predictors instead of the individual predictors to mitigate the multicollinearity problem that is very common in cases where many predictors are available. To abate the information losses due to dimension reduction, a semiparametric link between the count dependent variable and principal components is postulated. Clustering of observations is accounted into the model through a random effect term. The semiparametric poisson regression model is estimated via the backfitting algorithm. The simulation study illustrates the advantages of the proposed model over standard poisson regression in most of the simulation scenarios.



**Keywords**: *Semiparametric Poisson Regression, Clustered Data, Multicollinearity, Principal Components Analysis*