Bivariate regression model for count data based on the generalised Poisson distribution

Vera Hofer[1] **and Johannes Leitner**[1]
[1] University of Graz, Austria

A regression model for bivariate count data is introduced. The most popular approach to construct a bivariate distribution is trivariate reduction. Using three independent Poisson distributed variables X_1, X_2, X_3, a bivariate distribution (Y_1, Y_2) with Poisson marginals is derived as $Y_1 = X_1 + X_3$ and $Y_2 = X_2 + X_3$. The disadvantage of this approach is that the covariance can only be positive.

The bivariate count data model presented here is based on the Sarmanov distributions. The marginals are chosen to be generalised Poisson distributions. Even though the derivation of the distribution is demanding, a closed form of the likelihood function is obtained using the Lambert-W-function. Thus common maximum likelihood parameter estimation without simulation can be carried out. Our model yields less complicated formulas than copula based models and parameter estimation is less time-consuming.

The model is applied to artificial data and a large real dataset on health care demand. Its performance is compared to alternative models presented in the literature. Our model not only turns out to be superior to the other models in many settings. In addition, it gives insights into influences on the variance of the response variables.

Keywords: Sarmanov distribution, bivariate count data, Lambert-W-function