Knudson, Alexander D.; Kozubowski, Tomasz J.; Panorska, Anna K.; Schissler, A. Grant

A flexible multivariate model for high-dimensional correlated count data. (English)

Zbl 1465.62089

Summary: We propose a flexible multivariate stochastic model for over-dispersed count data. Our methodology is built upon mixed Poisson random vectors \( Y_1, \ldots, Y_d \), where the \( Y_i \) are conditionally independent Poisson random variables. The stochastic rates of the \( Y_i \) are multivariate distributions with arbitrary non-negative margins linked by a copula function. We present basic properties of these mixed Poisson multivariate distributions and provide several examples. A particular case with geometric and negative binomial marginal distributions is studied in detail. We illustrate an application of our model by conducting a high-dimensional simulation motivated by RNA-sequencing data.

MSC:
62H05 Characterization and structure theory for multivariate probability distributions; copulas
62H10 Multivariate distribution of statistics
62H30 Classification and discrimination; cluster analysis (statistical aspects)
62P10 Applications of statistics to biology and medical sciences; meta analysis

Keywords:
multivariate count data; copula; distribution theory; big data applications; gamma-Poisson hierarchy; mixed Poisson distribution; negative binomial distribution; high-dimensional multivariate simulation; RNA-sequencing data

Software:
GenOrd

Full Text: DOI

References:


This reference list is based on information provided by the publisher or from digital mathematics libraries. Its items are heuristically matched to zbMATH identifiers and may contain data conversion errors. It attempts to reflect the references listed in the original paper as accurately as possible without claiming the completeness or perfect precision of the matching.