

Project for a Master's Student's Thesis: Modelling Dependence in Discrete Multivariate Extremes

Project Description

In this project, we will explore statistical methods for modelling dependence in multivariate discrete extreme values. Many applications, including environmental sciences, hydrology, insurance, and healthcare, involve count data with heavy tails, excess zeros, and the possibility of extreme events occurring simultaneously across several variables.

The project will focus on copula-based models that combine marginal distributions to construct bivariate models with flexible dependence structures. We will investigate how different copula families capture the dependence between extreme counts, particularly the distinction between asymptotic dependence and asymptotic independence. The project may also consider zero-inflated models and likelihood-based estimation methods, including pairwise likelihood where full likelihood is computationally expensive, for efficient parameter estimation.

The student will study the theoretical foundations of these models, conduct simulation experiments to assess their performance, and, where appropriate, apply the methods to real datasets, such as precipitation measurements recorded by tipping-bucket rain gauges at multiple weather stations. The aim is to understand how dependence structures influence the probability of joint extreme events and to evaluate the suitability of different models for discrete multivariate extremes.

The project is suitable for a Master's student with a background in statistics, probability, and statistical computing and environmental modeling.

Bibliography

1. Ahmad, T., Gaetan, C., and Naveau, P. (2025). An extended generalized Pareto regression model for count data. *Statistical Modelling*, 25(5), 416–431.
2. Aka, S., Kratz, M., and Naveau, P. (2025). Multivariate discrete Generalized Pareto Distributions: Theory, simulation, and applications to dry spells. *arXiv preprint arXiv:2506.19361*.