

COPULA-BASED ANALYSIS FOR COUNT TIME SERIES

Mingze Zhang, Huixia Judy Wang* and James Livsey

*Walmart Global Tech, The George Washington University
and U.S. Census Bureau*

Abstract: We propose a new copula-based Markov model to analyze count time series data. One challenge in using copula to analyze count data is due to the identifiability issue that arises from the discrepancy of using a continuous copula function to characterize the discrete distribution. We find that identifiability can be ensured in the regression setup under one sufficient condition. Resolving the identifiability issue allows us to develop a method to select the appropriate copula to capture different types of temporal dependence, leading to more flexibility in modeling. We propose an estimation procedure and establish the asymptotic properties of the proposed estimators. For capturing temporal dependence, the proposed method is data-adaptive and computationally efficient. It also provides a convenient way to construct both point and interval predictions at a future time. Through a simulation study and the analysis of COVID-19 daily death data, we show that our method produces more stable point and interval predictions than existing methods based on Gaussian copula and autoregression.

Key words and phrases: Conditional quantile, copula, identifiability, Markov, temporal dependence.

1. Introduction

With a surge in applications from finance, environmental science, and social science, there has been increasing interest in developing models and methods for analyzing time series of counts; see Davis and Liu (2016), Jung, Kukuk and Liesenfeld (2006), Brandt et al. (2000), and Song, Li and Yuan (2009). The analysis of count time series is challenging due to the nonuniqueness of the dependence measure caused by the existence of ties. There exist some methods in the literature but many of them either rely on restrictive distributional assumptions, are computationally intensive, or have limitations for capturing complex dependence. In this paper, we aim to develop a flexible and computationally convenient procedure with weaker assumptions that can accommodate various types of dependence.

Some existing methods assume that the data follow a Poisson or negative binomial distribution conditional on past observations or an intensity process; see, e.g., Davis and Wu (2009), Davis, Dunsmuir and Streett (2003),

*Corresponding author. E-mail: judywang@rice.edu