Statistica Sinica **36** (2026), 1-22
doi:https://doi.org/10.5705/ss.202023.0237

JOINT MEAN-ANGLE MODEL FOR SPATIAL BINARY DATA

Cheng Peng[1], Renwen Luo[2], Yang Han[1] and Jianxin Pan*[2]

[1] *The University of Manchester*
and [2] *Beijing Normal-Hong Kong Baptist University*

Abstract: The analysis of spatially correlated binary data has received substantial attention in geo-statistical research but is very challenging due to the intricacy of the distributional form. Two principal objectives include examining the dependence of binary response on covariates of interest and quantifying the covariances or correlations between pairs of outcomes. While the literature has sufficiently addressed the modelling issue of the mean structure of a binary response, the characterization of the covariances between pairs of binary responses in terms of covariates is not clear. In this paper, we propose methods to explain such characterizations by using a latent Gaussian copula model with alternative hypersphere decomposition of the covariance matrix. Correctly specifying the covariance matrix is crucial not only for the high efficiency of mean parameters but also for scientific interest. The key is to model the marginal mean and pairwise covariance, simultaneously, for spatial binary data. Two generalized estimating equations are proposed to estimate the parameters, and the asymptotic properties of the resulting estimators are investigated. To evaluate the performance of the methods, we conduct simulation studies and provide real data analysis for illustration.

Key words and phrases: Alternative hypersphere decomposition, generalized estimating equation, joint mean-angle model, latent Gaussian copula model.

1. Introduction

Binary data widely arise in a range of scientific fields such as public health, biomedicine, economics, and agriculture. The collection of binary data is often more straightforward and cost-effective than gathering accurate continuous measurements since binary data only takes on two possible values. This paper focuses on the analysis of spatially correlated binary data, where the spatial dependence that cannot be ignored in making statistical inferences relies on the distances between sampling locations. Our research was motivated by a geostatistical study of Bovine Tuberculosis (bTB) infection of cattle herds and badger setts in Ireland. In this study, the binary outcome is whether the cattle herd or badger sett is infected with bTB or not. On the one hand, the primary objective is to describe the binary outcome as a function of the available covariates. On the other hand, since bTB is an infectious disease that can spread

*Corresponding author. E-mail: jianxinpan@uic.edu.cn