Computational Statistics and Data Analysis 178 (2023) 107612

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

www.elsevier.com/locate/csda

The circular quantile residual

Ana C.C. Andrade [a], Gustavo H.A. Pereira [b,*], Rinaldo Artes [c]

[a] Department of Statistics, Federal University of São Carlos, São Carlos, Brazil
[b] Department of Statistics, Federal University of São Carlos, Rod. Washington Luís, km 235 - SP-310, São Carlos, CEP 13565-905, Brazil
[c] Insper Institute of Education and Research, São Paulo, Brazil

ARTICLE INFO

Article history:
Received 13 July 2021
Received in revised form 22 August 2022
Accepted 2 September 2022
Available online 13 September 2022

Keywords:
Circular data
Circular-linear regression models
Diagnostic analysis
Quantile residual

ABSTRACT

Circular-linear regression is often used to model the relationship between a circular dependent variable and a set of linear predictor variables. It is used in many areas such as meteorology, biology, and medicine. For checking model adequacy, it is desirable to use residuals that are approximately standard normally distributed. Most of the residuals used in circular regression models do not meet this requirement and are used especially for outlier identification. Other residuals are limited to the von Mises regression models. An asymptotically standard normally distributed residual that can be used for any parametric circular-linear regression model is introduced. Monte Carlo simulation studies suggest that the distribution of this residual is well approximated by the standard normal distribution in small samples. To study the behavior of this residual, two regression models are introduced, and two applications are used to show that the proposed residual can detect model misspecification.

© 2022 Elsevier B.V. All rights reserved.

1. Introduction

Circular-linear regression is often used to model the relationship between a circular dependent variable and a set of linear predictor variables. It is useful for modeling phenomena of a directional or periodic nature, whose observations are, in practice, commonly considered in two-dimensional space and denominated circular data (Fisher, 1995; Jammalamadaka and Sengupta, 2001; Mardia and Jupp, 2009). Some examples of these variables include wind directions (Lang et al., 2020), movement of animals in response to stimuli (Rodríguez et al., 2020), and angle of regions of the human eyes (Abuzaid, 2020).

The von Mises family of distributions is often used to model the dependent variable of regression models with a circular response (Fisher and Lee, 1992) due to the similarities with the normal distribution and its main inferential properties. Gould (1969) introduced a regression model with a von Mises distributed response variable and linear predictors, whose parameters are estimated using the maximum likelihood method. However, Johnson and Wehrly (1978) noted problems of non-identifiability in Gould's model. Thus, the authors proposed an alternative to Gould's approach based on a marginal distribution function completely specified for the explanatory variable. Finally, Fisher and Lee (1992) generalized Johnson and Wehrly's methodology using a link function that maps the real line to the circle. The models relate the mean and dispersion parameter of the response variable to a set of linear explanatory variables. An overview of advances in regression models for circular data can be found in Kim and SenGupta (2018) and Pewsey and García-Portugués (2021).

* Corresponding author.
E-mail address: gpereira@ufscar.br (G.H.A. Pereira).

https://doi.org/10.1016/j.csda.2022.107612
0167-9473/© 2022 Elsevier B.V. All rights reserved.