MATEMATIKA, 2017, Volume 33, Number 2, 159–163

On parameter estimation of a replicated linear functional relationship model for circular variables

[1]Nurkhairany Amyra Mokhtar, [2]Yong Zulina Zubairi, [3]Abdul Ghapor Hussin and [4]Rossita Mohamad Yunus

[1,3]Faculty of Defence Sciences and Technology
National Defence University of Malaysia
Kem Sungai Besi, 57000 Kuala Lumpur, Malaysia

[2]Centre for Foundation Studies in Science,
University of Malaya, 50603 Kuala Lumpur, Malaysia

[4]Institute of Mathematical Sciences, Faculty of Science,
University of Malaya, 50603 Kuala Lumpur, Malaysia

e-mail: [1]khairany.amyra@gmail.com, [2]yzulina@um.edu.my
[3]abdulghapor@gmail.com, [4]rossita@um.edu.my

Abstract Replicated linear functional relationship model is often used to describe relationships between two circular variables where both variables have error terms and replicate observations are available. We derive the estimate of the rotation parameter of the model using the maximum likelihood method. The performance of the proposed method is studied through simulation, and it is found that the biasness of the estimates is small, thus implying the suitability of the method. Practical application of the method is illustrated by using a real data set.

Keywords Replicated linear functional relationship model; circular variables; parameter estimation; von Mises distribution

AMS mathematics subject classification 47S99; 34C60

1 Introduction

Error-in-variables model (EIVM) considers error in both x and y variables [1]. The functional relationship model is part of EIVM where the variables are fixed. When the data is circular, formal analysis cannot be done with usual statistical technique due to the wrap-around nature of a circle. Circular data can be represented as angles or as points of a circle [2].Early works of circular data analysis dates back to the mid-18th century [3]. The data may be measured in the range $[0, 2\pi)$ radians or $[0°, 360°)$ and the technique of analysing circular variable differs from usual Euclidean type variables [4] due to the wrapped-around nature of a circle.

The Von Mises distribution is said to be the most practical distribution on the circle [5]. The probability distribution function of the Von Mises distribution is given by

$$g(\theta; \mu, \kappa) = \frac{1}{2\pi I_0(\kappa)} e^{\kappa \cos(\theta - \mu)} \tag{1}$$

where $I_0(\kappa)$ is the modified Bessel function of the first kind and order zero, which can be defined by

$$I_0(\kappa) = \frac{1}{2\pi} \int_0^{2\pi} e^{\kappa \cos \theta} d\theta \tag{2}$$

where μ is the mean direction and κ is the concentration parameter.

2 The replicated linear functional relationship model

The replicated linear functional relationship model is given by $Y = \alpha + X(\mathrm{mod} 2\pi)$, where α is the rotation parameter and the variables considered for this model are

$$x_{ij} = X_i + \delta_{ij} \text{ and } y_{ik} = Y_i + \varepsilon_{ik}.$$

Both variables x and y have the error term of $\delta_{ij} \sim VM(0, \kappa)$ and $\varepsilon_{ik} \sim VM(0, \nu)$, respectively. In this case, the error terms δ_{ij} and ε_{ik} are independently distributed with a von Mises distribution and they have the concentration parameter κ and ν, respectively. Concentration parameter affects the Von Mises distribution inversely as variance affects the Normal distribution [6]. In this case, it is important to consider that there may be replicated observations of X_i and Y_i occurring in p groups. Measurement $x_{ij}, (j = 1, ..., m_i)$ are made on X_i and measurements $y_{ik}, (k = 1, ..., m_i)$ are made on Y_i, where $0 \leq x_{ij}, y_{ik} < 2\pi$.

3 The maximum likelihood estimation of parameters

In this section, X and the rotation parameter α are estimated using the method of maximum likelihood. The log likelihood equation of the von Mises distribution is given by

$$\log L(\alpha, \kappa, \nu, X; x, y) = -NM \log 2\pi - N \log I_0(\kappa) - M \log I_0(\nu) + \kappa \sum_{i=1}^{p} \sum_{j=1}^{n} \cos(x_{ij} - X_i) + \nu \sum_{i=1}^{p} \sum_{k=1}^{m} \cos(y_{ik} - \alpha - X_i) \quad (3)$$

To estimate X_i we have to find the first derivative of X with respect to log L thus giving

$$\frac{\partial \log L}{\partial X_i} = \kappa \sum_{j=1}^{m} \sin(x_{ij} - X_i) + \nu \sum_{k=1}^{m} \sin(y_{ik} - \alpha - X_i) = 0. \quad (4)$$

X_i may be solved iteratively by some "initial guess". Suppose $\hat{X}_{i0}$ is an initial estimate of $\hat{X}_i$. Then,

$$x_{ij} - \hat{X}_i = x_{ij} - \hat{X}_{i0} + \hat{X}_{i0} - \hat{X}_i = \left(x_{ij} - \hat{X}_{i0}\right) + \Delta_i$$

where $\Delta_{\mathrm{i}} = \hat{X}_{i0} - \hat{X}_i$. We may also have $y_{ik} - \hat{\alpha} - \hat{X}_i = (y_{ik} - \hat{\alpha} - \hat{X}_{i0}) + \Delta_i$ Thus, the partial derivative of equation (4) becomes

$$\sin\left(x_{ij} - \hat{X}_{i0} + \Delta_i\right) + \sin(y_{ik} - \hat{\alpha} - \hat{X}_{i0} + \Delta_i) = 0$$

for small Δ, then $\cos \Delta_i \approx 1$ and $\sin \Delta_i \approx 1$.

Hence the equation is simplified (approximately) to become

$$\hat{X}_{i1} \approx \hat{X}_{10} + \frac{\sum\limits_{j=1}^{m} \sin\left(x_{ij} - \hat{X}_{i0}\right) + \frac{\hat{\nu}}{\hat{\kappa}} \sum\limits_{k=1}^{m} \sin\left(y_{ik} - \hat{\alpha} - \hat{X}_{i0}\right)}{\sum\limits_{j=1}^{m} \cos\left(x_{ij} - \hat{X}_{i0}\right) + \frac{\hat{\nu}}{\hat{\kappa}} \sum\limits_{k=1}^{m} \cos\left(y_{ik} - \hat{\alpha} - \hat{X}_{i0}\right).} \quad (5)$$

To estimate α we have to find the first derivative of α with respect to $\log L$.

$$\frac{\partial \log L}{\partial \alpha} = \nu \sum_{i=1}^{p}\sum_{k=1}^{m} \sin\left(y_{ik} - \alpha - X_i\right) = 0, \tag{6}$$

$$\frac{\cos\hat{\alpha}}{\sin\hat{\alpha}} = \frac{\sum\limits_{i=1}^{p}\sum\limits_{k=1}^{m} \cos\left(y_{ik} - \hat{X}_i\right)}{\sum\limits_{i=1}^{p}\sum\limits_{k=1}^{m} \sin\left(y_{ik} - \hat{X}_i\right)}. \tag{7}$$

Therefore, we obtain the estimate of α as given by

$$\hat{\alpha} = \tan^{-1}\left\{\frac{\sum\limits_{i=1}^{p}\sum\limits_{k=1}^{m} \cos\left(y_{ik} - \hat{X}_i\right)}{\sum\limits_{i=1}^{p}\sum\limits_{k=1}^{m} \sin\left(y_{ik} - \hat{X}_i\right)}\right\}. \tag{8}$$

4 Simulation study

To assess the accuracy and the biasness of the parameters of the proposed model, a simulation study has been carried out [7]. The number of simulation is set to be s, meanwhile the values of n and κ for the error terms have been generated. In this model, the value of $\alpha$is circular meanwhileκvalues are continuous.In this section, the value of s is set to be 5000 for each simulation. The values of X have been generated from the Von Mises distribution of VM$(2,3)$ and the true value of $\alpha = \frac{1}{4}\pi \approx 0.7854$. The values of the concentration parameters of the error term used in this study are $\kappa = 3$, 5, 10, 15 respectively. For each value of κ, the sample size n =30, 78 and 132 are considered for the simulation.

4.1 Biasness of $\hat{\alpha}$

Bias measure is used to investigate the performance. For the bias measure of $\hat{\alpha}$, the mean, circular distance and mean resultant length are used. The respective formulae are as follow:

- Mean of circular parameter $\hat{\alpha}$, $\bar{\hat{\alpha}}$:

$$\bar{\hat{\alpha}} = \begin{cases} \tan^{-1}\left(\dfrac{S}{C}\right) & \text{when } S > 0,\ C > 0, \\ \tan^{-1}\left(\dfrac{S}{C}\right) + \pi & \text{when } C < 0, \\ \tan^{-1}\left(\dfrac{S}{C}\right) + 2\pi & \text{when } S < 0,\ C > 0, \end{cases}$$

where $C = \sum_{j=1}^{s} \cos\left(\hat{\alpha}_j\right)$ and $S = \sum_{j=1}^{s} \sin\left(\hat{\alpha}_j\right)$.

- Circular distance, $d = \pi - \left|\pi \left|\bar{\hat{\alpha}} - \alpha\right|\right|$.
- Mean resultant length, $\bar{R} = \dfrac{1}{s}\sqrt{\left(\sum_{j=1}^{s} \cos\left(\hat{\alpha}_j\right)\right)^2 + \left(\sum_{j=1}^{s} \sin\left(\hat{\alpha}_j\right)\right)^2}$.

5 Simulation result

Table 1 shows the simulation result of the bias measure of $\hat{\alpha}$. From Table 1, for any fixed κ, the circular mean of $\hat{\alpha}$ becomes closer to the real value of α ≈0.7854, as we increase the value of n. The circular distance gets smaller as we increase the value of n as well. Similarly, based on mean resultant length, the estimation gets better with increasing value of n for anyfixed κ.

Table 1: Biasness of $\hat{\alpha}$

		Circular mean, $\bar{\hat{\alpha}}$	Circular distance, d	Mean resultant length, $\bar{R}$
$\kappa = 3$ $\mu = 3$	n =50	0.7857	0.0003	0.9917
	n =78	0.7853	0.0001	0.9947
	n =132	0.7867	0.0013	0.9968
$\kappa = 5$ $\mu = 5$	n =50	0.7839	0.0015	0.9955
	n =78	0.7856	0.0002	0.9971
	n =132	0.7845	0.0009	0.9983
$\kappa = 10$ $\mu = 10$	n =50	0.7857	0.0003	0.9979
	n =78	0.7855	0.0001	0.9987
	n =132	0.7857	0.0003	0.9992
$\kappa = 15$ $\mu = 15$	n =50	0.7859	0.0005	0.9986
	n =78	0.7853	0.0001	0.9991
	n =132	0.7854	0.0000	0.9995

6 Application to real data

In this paper, we illustrate the applicability of the proposed model using the wind direction data collected from the Humberside coast of the North Sea, United Kingdom. With the sample size of 49, the data of the wind direction which was measured by HF radar system, developed by UK Rutherford and Appleton Laboratories, is addressed as the variable x. Meanwhile variable y is the wind data measured by anchored wave buoy.The relationship between the variables x and y is given by $Y = 0.08570 + X \pmod{2\pi}$ where $\mathrm{var}(\hat{\alpha}) = 0.01425$ is small, where this means that the values of $\hat{\alpha}$ are tightly clustered together and this indicates a good estimation for α.

7 Conclusion

This study proposes an estimation of the rotation parameterα.The simulation study shows that the parameter estimation is good since the bias becomes smaller as the sample size and the parameter concentration are increased. Also an application of the method is givenby using a real wind direction data

Acknowledgement

We would like to thank National Defence University of Malaysia and University Malaya (research grant RG369-15AFR) for supporting this work. We also acknowledge Ministry of Higher Education (MOHE) Malaysia for financial support to this research under the STEM Grant (Vote no. AJ091002.5600.07397).

References

[1] Mokhtar N. A, Zubairi Y. Z. and Hussin A. G. A simple linear functional relationship model for circular variables and its application. In *Proceedings of the 9th International Conference on Renewable Energy Sources.* 2015. 57–63

[2] Jammaladaka, S. R. and Sengupta, A. *Topics in Circular Statistics.* World Scientific Publishing. 2001.

[3] Fisher, N. I. *Statistical Analysis of Circular Data.* Cambridge University Press. 1993.

[4] Hussin A. G, Fieller N. and Stillman E. Pseudo-replicates in the linear circular functional relationship model. *Journal of Applied Sciences.* 2005. 5(1): 138–143.

[5] Mardia, K. V. and Jupp, P. E. *Directional Statistics*. John Wiley & Sons. 2000.

[6] Mokhtar N. A, Zubairi Y. Z. and Hussin A. G. Parameter estimation of simultaneous linear functional relationship model for circular variables assuming equal error variances. *Pakistan Journal of Statistics.* 2015. 31(2): 251–265.

[7] Mokhtar N. A, Zubairi Y. Z. and Hussin A. G. Detecting multiple outliers in linear functional relationship model for circular variables using clustering technique. *AIP Conference Proceedings.* 2017. 1842 (020005).