Two-stage Circular-circular Regression with Zero-inflated Response and Covariate

Technical Report No. ASU/2017/14
Dated: 16 August, 2017

Jayant Jha
Applied Statistics Unit,
Indian Statistical Institute,
Kolkata 700108
jayantjha@gmail.com

Prajamitra Bhuyan
Applied Statistics Unit,
Indian Statistical Institute,
Kolkata 700108
bhuyan.prajamitra@gmail.com

Two-stage Circular-circular Regression with Zero-inflated Response and Covariate

Jayant Jha and Prajamitra Bhuyan
Applied Statistics Unit, Indian Statistical Institute
203 B. T. Road, Kolkata - 700 108, India
E-mails: jayantjha@gmail.com, bhuyan.prajamitra@gmail.com

Abstract

In many real life scenarios, the response and the covariate are circular in nature. The circular-circular regression model is a popular technique to analyse such relationships. In some cases, the angular response and the covariate are zero-inflated. In statistical literature, circular-circular regression models and estimation of associated parameters with only zero-inflated responses have been proposed. However, there are no models to deal with zero-inflated response as well as covariate. In this paper, the Möbius transformation based two-stage circular-circular regression model is considered for such data. We propose Bayesian estimation of the model parameters using MCMC algorithm. The method is illustrated through simulation and the analysis of a real data set arising from cataract surgery study.

Keywords : Circular dispersion; Credible interval; Latent variable; Metropolis-Hastings algorithm; Truncated wrapped Cauchy.

1 Introduction

There are several real life phenomenon where the measurements are taken in angles. Some common examples of such random variables are wind direction (Kato et al., 2008), direction of migrating birds (Busse and Trocinska, 1999), etc. There are some random variables which are cyclic in nature, however not directly measured in angles, e.g. time of the day and date of the year (Jha and Biswas, 2017b). These random variables are commonly known as circular random variables or directional random variables in the literature (Mardia and Jupp, 2000).

It is of interest to analyse the relationship of a circular response with associated covariates in various fields of sciences like meteorology (Kato et al., 2008), geoscience (Rivest, 1997), medical science (Jha and Biswas, 2017a), and other allied fields. In this context, some rotational models have been proposed where the predicted mean direction of the response is a fixed rotation of the covariate. See Mackenzie (1957) and Rivest (1997) for more details. Fisher and Lee (1992), Bhattacharya and Sengupta (2009) and Gould (1969) proposed regression models for the case of linear-circular regressions. Due to the difference in topology of circle and Euclidean space, these models can not be directly applied to the case of circular-circular regression. Downs and Mardia (2002) proposed Möbius transformation based regression link function for circular-circular regression. Later, Kato et al. (2008) also considered the Möbius Transformation based link function by reparametrizing the model of Downs and Mardia (2002). This reparametrization induced a nice geometry to the regression link function and when the angular error follows wrapped Cauchy distribution, it provides some advantages in terms of distributional properties as the wrapped Cauchy distribution is closed under rotation and Möbius transformation. Another advantage of this model is that the rotational model, where the predicted mean direction of the response is a fixed rotation of the covariate, is a subset of this model. Also, unlike the rotational model, this model is adequate when there is a high concentration of responses on a part of the unit circle.

In this article, we consider the circular-circular regression model where both the response and the covariate are zero-inflated. There are numerous studies focusing on zero-inflated random variables in the linear setup. Tobin (1958), Heckman (1974, 1979) proposed some models in the context of linear regression where the responses are zero-inflated. Lambert (1992) considered Poisson regression for count data with excess zeros in the response variable. Bhuyan et al. (2016) discussed the case when both the responses and covariates are zero-inflated and proposed estimation methodology for the model parameters under Bayesian set up. See Min and Agresti (2002) for a detailed review of the zero-inflated regression models. However, in all of these works, the response and the covariates are linear in nature where zero-inflation is caused by censorship through a selection mechanism. The case of circular random variables is considerably different from linear ones. The difference mainly arises due to the topology of the circle where the zero can not be considered to be at boundary of the sample space. Moreover, in the circular setup, a spike at zero may be interpreted in the same way like a spike at any other angle, since origin can be fixed arbitrarily without loss of generality. There are very few works on the analysis of zero-inflated circular data. Zero-inflated distribution modelling for circular random variables was studied in Biswas et al. (2016b). In the context of circular-circular regression, Jha and Biswas (2017a) proposed a model with zero-inflated response variable and discussed the estimation of associated model parameters. However, there is no such model available in the literature for the case of zero-inflated circular covariate. In order to avoid the difficulty arising from zero-inflated covariate, Jha and Biswas (2017a) analysed a subset of the dataset considering only non-zero covariate values.

In order to model such data, we define continuous latent variables and consider two-stage circular-circular regression model, which is an extension of the model proposed by Kato et al. (2008). We propose an estimation methodology under Bayesian set up using MCMC algorithm for the associated model parameters, which is described in Section 2. We carry out a real-life data analysis based on a cataract surgery dataset as an illustration of our proposed methodology in Section 3. Section 4 discusses the simulation studies. We end with some concluding remarks and possible extensions in Section 5.

2 Proposed Model and Methodology

In this section, we first describe circular-circular regression model proposed by Kato et al. (2008). Then we propose two stage circular-circular regression with Möbius transformation based link functions. Let us represent circular random variables θ_Y and θ_X as complex random variables $Y = e^{i\theta_Y}$ and $X = e^{i\theta_X}$, respectively taking values on the circumference of a unit circle. The circular-circular regression model of Kato et al. (2008) can be written in the following way:

$$y = \beta_0 \frac{x + \beta_1}{1 + \bar{\beta}_1 x} \epsilon,$$

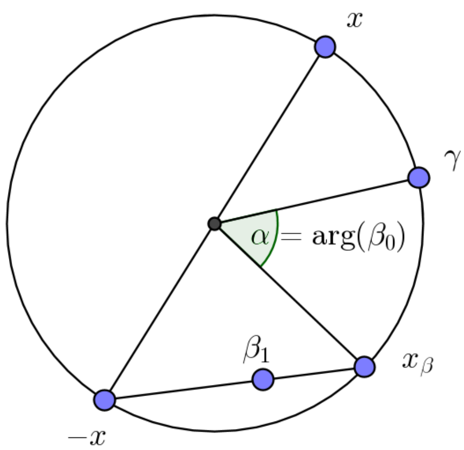

Figure 1: Circular-circular regression model.

where β_0, $\epsilon \in \{z : z \in \mathbb{C}; |z| = 1\}$, $\beta_1 \in \mathbb{C}$, and the angular error $\arg(\epsilon)$ follows a wrapped Cauchy distribution with mean direction 0. The geometry of the Möbius transformation based link function is shown in Figure 1, where γ shows the predicted mean direction of the response given the value of the covariate is x. See Kato et al. (2008) for more details.

Now, we consider the case where there is high concentration of observations in a particular direction for both the response and the covariate. One can use the rotational symmetry of a circle and consider θ_y and θ_x to be zero-inflated circular variables. In order to model zero-inflated data, we first define circular latent variables

$$\theta_Y = \begin{cases} 0, & \text{if } \theta_{Y^*} \in (-\delta_Y, \delta_Y) \\ \theta_{Y^*}, & \text{otherwise,} \end{cases}$$

and

$$\theta_X = \begin{cases} 0, & \text{if } \theta_{X^*} \in (-\delta_X, \delta_X) \\ \theta_{X^*}, & \text{otherwise,} \end{cases}$$

where δ_Y, δ_X are known constants taking values in $[0, \pi)$. We consider a circular instrumental variable θ_{X_2}, and propose the two stage circular-circular regression model as

$$Y^* = \beta_0 \frac{X^* + \beta_1}{1 + \overline{\beta_1} X^*} \epsilon_1, \tag{2.1}$$

$$X^* = b_0 \frac{X_2 + b_1}{1 + \overline{b_1} X_2} \epsilon_2 \tag{2.2}$$

where, β_0, b_0, ϵ_1, $\epsilon_2 \in \{z : z \in \mathbb{C}; |z| = 1\}$; β_1 ,$b_1 \in \mathbb{C}$, $Y^* = e^{i\theta_{Y^*}}$, $X^* = e^{i\theta_{X^*}}$ and $X_2 = e^{i\theta_{x_2}}$. We assume $\arg(\epsilon_i) \sim WC(0, \rho_i)$, where $WC(\mu, \rho)$ represents the wrapped

Cauchy distribution with parameters $\mu \in [0, 2\pi)$ and $\rho \in [0, 1]$, for $i = 1, 2$. Also, $\arg(\epsilon_1)$ and $\arg(\epsilon_2)$ are assumed to be independently distributed. Note that the above model reduces to the two-stage circular-circular regression model without zero-inflation when $\delta_X = \delta_Y = 0$.

2.1 Bayesian Estimation

We propose Bayesian estimation of the parameters involved in the model (2.1) and (2.2) using MCMC algorithm based on data augmentation. We sample the unobserved latent variables from the conditional probability distribution of $\theta_{Y^*}(\theta_{X^*})$, given $\theta_Y(\theta_X) = 0$, as well as model parameters. Note that the above conditional distribution follows the truncated wrapped Cauchy distribution with density function

$$f_{TW}(\theta; \mu, \rho, -\delta, \delta) = \begin{cases} K^{-1} f_W(\theta; \mu, \rho), & \text{if } \theta \in (-\delta, \delta) \\ 0, & \text{otherwise,} \end{cases}$$

where $K = \int_{-\delta}^{\delta} f_W(\theta; \mu, \rho) dx$, and $f_W(\theta; \mu, \rho)$ is the density of the wrapped Cauchy distribution with parameters μ and ρ. We propose an algorithm for generating samples from the truncated wrapped Cauchy distribution which is discussed in the Subsection 2.2.

Let us denote $\Theta_i = (\theta_{0i}, r_i, \theta_{1i}, \rho_i)$ for $i = 1, 2$, where $\beta_0 = e^{i\theta_{01}}$, $\beta_1 = r_1 e^{i\theta_{11}}$, $b_0 = e^{i\theta_{02}}$, $b_1 = r_2 e^{i\theta_{12}}$, $r_1, r_2 \in [0, \infty)$. The joint posterior density for the model in equation (2.1) can be written as

$$\pi(\Theta_1, \theta_{Y^*}|\theta_Y) \propto \pi(\Theta_1) \times \Pi_{i=1}^n \{f_W(\theta_{Y_i}; \mu_{1i}, \rho_1) 1(\theta_{Y_i} \neq 0) + f_{TW}(\theta_{Y_i^*}; \mu_{1i}, \theta_{X_i^*}, -\delta_Y, \delta_Y) 1(\theta_{Y_i} = 0)\},$$

where $\mu_{1i} = \arg(\beta_0 \frac{X_i^* + \beta_1}{1 + \bar{\beta}_1 X_i *})$, and $\pi(\Theta_1)$ denotes the prior density of Θ_1. Similarly, for the model in equation (2.2), the joint posterior density can be written as

$$\pi(\Theta_2, \theta_{X^*}|\theta_{X_2}) \propto \pi(\Theta_2) \times \Pi_{i=1}^n \{f_W(\theta_{X_i}; \mu_{i2}, \rho_2) 1(\theta_{X_i} \neq 0) + f_{TW}(\theta_{X_i^*}; \mu_{i2}, \rho_2, -\delta_X, \delta_X) 1(\theta_{X_i} = 0)\},$$

where $\mu_{2i} = \arg(b_0 \frac{X_{2i} + b_1}{1 + \bar{b}_1 X_{2i}})$, and $\pi(\Theta_2)$ denotes the prior density of Θ_2. The conditional densities of the model parameters can then be written as

$$\pi(\Theta_1|-) \propto \pi(\Theta_1) \Pi_{i=1}^n f_W(\theta_{Y_i^*}; \mu_{1i}, \rho_1), \tag{2.3}$$

$$\pi(\Theta_2|-) \propto \pi(\Theta_2) \Pi_{i=1}^n f_W(\theta_{X_i^*}; \mu_{2i}, \rho_2), \tag{2.4}$$

Note that the full conditional densities of latent variables $\theta_{X_i^*}$ and $\theta_{Y_i^*}$ have the following closed form expressions

$$\pi(\theta_{X_i^*}|-) \equiv \begin{cases} \theta_{X_i} \text{ with probability } 1, & \text{if } \theta_{X_i} \neq 0 \\ f_{TW}(\theta_{X_i*}; \mu_{2i}, \rho_2, -\delta_X, \delta_X), & \text{otherwise,} \end{cases} \tag{2.5}$$

and

$$\pi(\theta_{Y_i^*}|-) \equiv \begin{cases} \theta_{Y_i} \text{ with probability } 1, & \text{if } \theta_{Y_i} \neq 0 \\ f_{TW}(\theta_{Y_i^*}; \mu_{1i}, \rho_1, -\delta_Y, \delta_Y), & \text{otherwise,} \end{cases} \tag{2.6}$$

respectively. However, the full conditionals of the model parameters can not be expressed in closed form. Therefore, we employ Metropolis-Hastings algorithm for generating samples from the posterior densities of the parameters and the detailed algorithm is provided in Subsection 2.2.

2.2 Sampling Algorithms

2.2.1 Sample Generation from Truncated Wrapped Cauchy Distribution

Let θ_Z be a circular random variable following truncated wrapped Cauchy distribution with pdf $f_{TW}(\theta_z; \mu, \rho, a, b)$, where $a, b \in [-\pi, \pi)$. If $a < b$, then the support of the above density is (a, b), and $(a, \pi) \cup [-\pi, b)$, otherwise. In order to generate samples from $f_{TW}(\theta_z; \mu, \rho, a, b)$ for $a < b$, we simulate observations from $f_W(\theta_z; \mu, \rho)$ and accept the observations which lie in (a, b). Similarly, for $a > b$, we accept the observations lying in $(a, \pi) \cup [-\pi, b)$. However, the acceptance rate is very low for small values of $(b-a)1(a < b) + [2\pi - (b-a)]\, 1(a > b)$. For example, the acceptance rate is approximately 0.0003, for generating samples from $f_{TW}(\theta_z; \mu, \rho, a, b)$ with $a = \pi - 0.035$, $b = -\pi + 0.035$, $\mu = 0$, and $\rho = 0.95$. Thus, we propose a novel algorithm for generating samples from the truncated wrapped Cauchy distribution based on the geometry of the Möbius transformation from unit circle to unit circle.

If θ_Z is uniformly distributed in $[0, 2\pi)$ and $Z = e^{i\theta_z}$, then $\arg\left(\frac{\psi - Z}{1 - \bar{\psi} Z}\right)$ follows wrapped Cauchy distribution with parameters $\mu = \arg(\psi)$, and $\rho = |\psi|$, where $\psi \in \{c \in \mathbb{C} : |c| \leq 1\}$ (Kato et al., 2008). The geometry of the transformation $\eta(Z) = \frac{\psi - Z}{1 - \bar{\psi} Z}$ is provided in Figure 1. Note that, $\eta(Z)$ is the point on the circumference of the unit circle which is situated at the

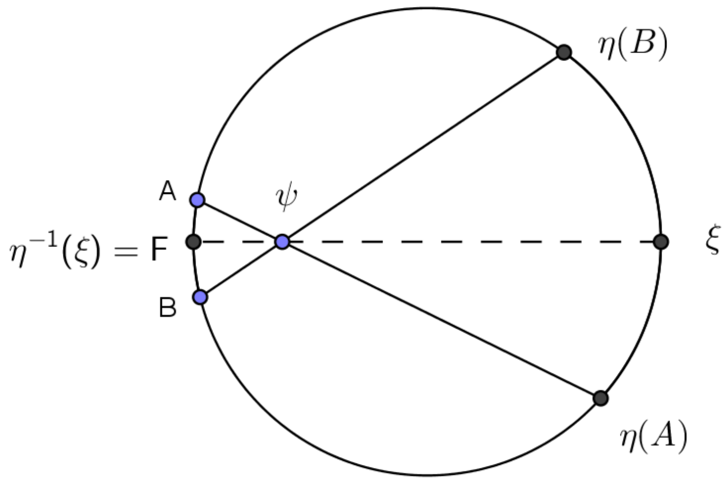

Figure 2: Sample Generation from truncated wrapped Cauchy distribution.

intersection of the line joining Z and ψ and the unit circle. It is easy to see that $\arg\{\eta(Z)\}$ follows wrapped Cauchy distribution with $\mu = \arg(\psi)$ and $\rho = |\psi|$.

In order to simulate an observation W from the truncated wrapped Cauchy distribution, one can generate a unit complex number ξ uniformly in the region between $\eta(A)$ and $\eta(B)$ and consider the argument of its inverse Möbius transformation $\eta^{-1}(\xi)$, where $A = e^{ia}$ and $B = e^{ib}$. The aforementioned explanation is represented diagrammatically in Figure 2. Also, we provide a detailed algorithm which can be implemented for generating samples from truncated wrapped Cauchy distribution below.

Algorithm 1: Sampling from Truncated Wrapped Cauchy Distribution

Step 1: Take $A = e^{ia}$, $B = e^{ib}$ and choose a point c in the support of $f_{TW}(\theta_z; \mu, \rho, a, b)$.

Step 2: Generate a random unit complex number ξ uniformly from the arc joining $\eta(A)$ and $\eta(B)$ containing $\eta(C)$, where $C = e^{ic}$.

Step 3: Take $W = \arg\{(\eta^{-1}(\xi)\}$.

2.2.2 Metropolis-Hastings Algorithm

For the purpose of Bayesian estimation, we consider the following prior distributions for the parameter vectors Θ_1 and Θ_2. The joint prior distribution $\pi(\Theta_i)$ can be expressed as the product of $\pi(\theta_{0i}) \propto 1$, $\pi(\theta_{1i}) \propto 1$, $\pi(r_i) \propto e^{-r_i^2}$ and $\pi(\rho_i) \propto \rho_i^{a_{\rho_i}-1}(1-\rho_i)^{a_{\rho_i}-1}$, for $i = 1, 2$.

It can be easily verified that the joint posterior density is proper for $a_{\rho_i} > 1$. Similar priors have been considered by Ravindran and Ghosh (2011) in the context of circular-circular and circular-linear regressions. For the purpose of implementing Metropolis-Hastings Algorithm, we consider the proposal distributions for $\theta_{01}, \theta_{11}, \theta_{02}, \theta_{12}, \rho_1, \rho_2$ to be uniform, and the proposal distributions for r_1, r_2 are chosen to be exponential.

Algorithm 2: Metropolis Hastings Algorithm

Step 1: Sample θ_{X^*} from the density $\pi(\theta_{X^*}|-)$ given in equation (2.5), and then sample θ_{Y*} from the density $\pi(\theta_{Y^*}|-)$ given in equation (2.6).

Step 2: Generate the model parameters sequentially from the corresponding proposal densities and denote it as ν_p'. Given the previous value of ν_p and the current draw ν_p', return ν_p' with probability

$$\alpha_{MH}(\nu_p, \nu_p') = \min\left\{1, \frac{\pi(\nu_p'|-)\pi(\nu_p', \nu_p)}{\pi(\nu_p|-)\pi(\nu_p, \nu_p')}\right\},$$

otherwise, repeat the previous value ν_p, where $\pi(d, w)$ denotes the proposal density at d with parameter w and $\pi(.|)$ denotes the full conditional densities given in equations (2.3) and (2.4).

Step 3: Repeat Step 1 and Step 2 until convergence.

3 Data Analysis

We consider the dataset obtained from a study on cataract surgery conducted at Disha Eye Hospital and Research Center, Barrackpore, West Bengal, India, over a period of two years from 2008 to 2010. In this study, patients were treated with different techniques of cataract surgery. See Bakshi (2010) for further details. During the cataract surgery, incision causes unwanted changes to the natural corneal shape causing an astigmatic eye. The axis of astigmatism of the eyes were measured on the 1st, 7th and 15th day after the operation. According to medical practitioners, if the axis is closer to 0, 90 or 180 degrees, then it is not a matter of serious concern. Therefore, axes of astigmatism were multiplied by 4 (modulo

360) and this transformed dataset was considered to test the treatment differences with respect to astigmatism after the cataract surgery operation. Due to this transformation, the variables become zero-inflated. See Biswas et al. (2015) and Biswas et al. (2016a) for more details. Also, it is of interest to predict the axis of astigmatism after two weeks of the surgery for the purpose of identifying patients who require frequent monitoring and additional care. Jha and Biswas (2017a) fitted a zero-inflated circular-circular regression model considering the response and the covariate as the axis of astigmatism after 15 days and 7 days of the surgery, respectively, where only the part of dataset was considered in which the responses were zero-inflated but the covariates were not.

In this analysis, we consider the dataset corresponding to SICS, Snare and Conventional Phacoemulsification etc. techniques of cataract surgery, where both the response and the covariate are zero-inflated. We consider the response (θ_y), the covariate (θ_x) and the instrumental variable (θ_{x_2}) as the axis of astigmatism after 15 days, 7 days and 1 day after the surgery, respectively. Among the 54 observations for the response and the covariate, there are 31% and 35% zeroes, respectively.

In order to apply the methodology provided in Section 2, we consider $\delta = \delta_X = \delta_Y = 0.035$ radians (2 degrees) and $a_{\rho_i} = 2$ for $i = 1, 2$. This particular choice of δ is considered as the precision of the observed values are measured up to 4 degrees. As mentioned before, when $\delta_X = \delta_Y = 0$, the proposed model reduces to conventional two-stage circular-circular regression. In order to compare the results obtained from our proposed model (Model I, say), we also apply the two-stage circular-circular regression without zero-inflation (Model II, say) for data analysis. We generate 100,000 samples from the posterior distributions of the associated model parameters using MCMC algorithm and find the posterior mean and standard deviation (s.d.) for linear parameters and posterior circular mean and circular dispersion (c.d.) for circular parameters based on every 10th iterate discarding the first 60,000 iterations as burn-in. The convergence of the chains are monitored graphically. The mean (circular mean) and s.d. (c.d.), of the linear parameters (circular parameters) are reported in Table 1. Note that, the circular dispersion for n circular observations $\phi_1, \ldots \phi_n$ is given by $1 - \bar{R}$, where $\bar{R} = \frac{|| \sum_{i=1}^{n} z_i ||}{n}$ and $z_i = (\cos \phi_i, \sin \phi_i)$ for each i. We also report the 95% highest posterior density (HPD) credible interval. The details for finding the HPD credible interval for circular parameters are provided in Jha (2017). As expected, the 95%

HPD credible interval for both r_1 and r_2 does not contain 1 (See Table 1), hence, one can conclude in a crude way that θ_y (θ_x) is dependent on θ_x (θ_{x_2}). It is evident that Model I fits the data better compared to Model II with respect to BIC=$k \log(n) - 2\log(L)$, where k is the number of parameters and L is the likelihood function (See Table 1).

In order to compare the predicted values with the observed values, we converted the predicted values in degrees and rounded off to the nearest integer divisible by 4. In general, the patients not affected by astigmatism after 7 days of the surgery remain unaffected in near future. In our dataset, there were 17 patients who remain unaffected by astigmatism during the study period. Interestingly, our fitted model predicts the same. Medical practitioners are more interested in identifying patients whose performances either improve or deteriorate. As per fitted Model I, we detect improvement in 17 out of the 21 patients. However, the deterioration is detected for only 6 out of the 13 patients. Overall, the proposed model fits the data pretty well and one can use the aforementioned results to take effective decisions during post-operative care. As per our prediction, we may conclude that the patients whose conditions have improved and seem to stay good in the future require less monitoring while rest of the patients need more frequent monitoring and care.

4 Simulation Studies

In order to study the performance of the proposed method, we generate data considering different sets of parameter values each at two different sample sizes, 50 (close to the sample size of data analysis) and 100. Similar to the data analysis, we consider $\delta = 0.035$ radians. We generate θ_{x_2} from von Mises distribution with mean 0 and $\kappa = 2$. The five different sets of parameters values of Θ_1 and Θ_2 are chosen such that the approximate proportions of zeros in the response and covariate are given by $(0.15, 0.15)$, $(0.10, 0.10)$, $(0.10, 0)$, $(0, 0.10)$ and $(0, 0)$, respectively. The choice of priors are taken to be the same as used in data analysis. Similar to the data analysis, we generate 100,000 samples from the posterior distributions of the associated model parameters using MCMC algorithm and find the posterior mean/circular mean and s.d./c.d. based on every 10th iterate discarding the first 60,000 iterations as burn-in. This is repeated 100 times and the average estimates are reported in Tables 2-6. We also report the coverage probability (CP) and 95% HPD interval corresponding to all

Table 1: Estimates of model parameters for Cataract Surgery data

Model I		
Parameters	mean/circular mean (s.d./c.d.)	95% HPD Credible Interval
θ_{01}	5.6352 (0.0217)	(5.2810,5.9387)
θ_{11}	1.6711 (0.0661)	(1.005,2.3580)
θ_{02}	6.0366 (0.0075)	(5.9220,0.0949)
θ_{12}	1.7282 (0.1958)	(5.8441,2.4084)
r_1	0.3285 (0.0913)	(0.2111,0.4895)
r_2	0.1467 (0.0490)	(0.0625,0.2070)
ρ_1	0.8640 (0.0322)	(0.7963,0.9227)
ρ_2	0.8957 (0.0339)	(0.8379,0.9468)
BIC	310.6729	
Model II		
Parameters	mean/circular mean (s.d./c.d.)	95% HPD Credible Interval
θ_{01}	5.5534 (0.0135)	(5.1538,5.8386)
θ_{11}	1.8041 (0.0265)	(1.3810,2.3290)
θ_{02}	6.1244 (0.0210)	(5.7827,0.2431)
θ_{12}	1.2530 (0.4699)	(5.6026,2.3961)
r_1	0.3592 (0.0759)	(0.2326,0.5449)
r_2	0.1505 (0.0735)	(0.0655,0.2991)
ρ_1	0.8726 (0.0332)	(0.8037,0.9297)
ρ_2	0.8978 (0.0327)	(0.8317,0.9526)
BIC	436.1317	

Table 2: Results of the simulation study with 15% zeroes in both the response and covariate

n=50					
Parameters	circular mean (c.d.)	CP	Parameters	mean (s.d.)	CP
$\theta_{01} = 0$	-0.086(0.068)	0.94	$r_1 = 0.9$	0.910(0.083)	0.93
$\theta_{02} = 0$	0.053(0.036)	0.95	$r_2 = 1.2$	1.174(0.082)	0.92
$\theta_{11} = 0$	0.086(0.071)	0.93	$\rho_1 = 0.85$	0.844(0.030)	0.97
$\theta_{12} = 0$	-0.049(0.031)	0.96	$\rho_2 = 0.85$	0.838(0.030)	0.95
n=100					
Parameters	circular mean (c.d.)	CP	Parameters	mean (s.d.)	CP
$\theta_{01} = 0$	0.000(0.014)	0.92	$r_1 = 0.9$	0.894(0.046)	0.95
$\theta_{02} = 0$	0.004 (0.012)	0.93	$r_2 = 1.2$	1.200(0.053)	0.97
$\theta_{11} = 0$	-0.001(0.015)	0.93	$\rho_1 = 0.85$	0.847(0.021)	0.98
$\theta_{12} = 0$	-0.001(0.010)	0.96	$\rho_2 = 0.85$	0.848(0.020)	0.95

the parameters. As expected, the CPs for all the cases are close to 95% and the standard deviations and circular dispersions decrease as the sample size increases (See Tables 2-6). Comparing the results of the simulation study without zero-inflation (See Table 6) with zero-inflated cases (Tables 2-5), it can be readily seen that even with the higher percentage of zeroes, our method seems to perform reasonably well.

4.1 Model Comparison and Sensitivity Analysis

We first compare the performance of parameter estimates associated with Model I and Model II when the data is generated from Model I with sample size $n = 50$ and $\delta = 0.070$ radians (4 degrees). The averages of the estimates over 100 replications are reported in Table 7. It can be observed that most of the parameter estimates for Model II are biased compared to those of Model I. Moreover, the mean BIC value for Model I is smaller compared to Model II, which clearly indicates that Model II is inadequate to deal with zero-inflated data.

In order to carry out sensitivity analysis, we consider the following mis-specified simulation model where a significant proportion of zero values for both response and covariates

Table 3: Results of the simulation study with 10% zeroes in both the response and covariate

n=50					
Parameters	circular mean (c.d.)	CP	Parameters	mean (s.d.)	CP
$\theta_{01} = 0$	-0.057(0.049)	0.92	$r_1 = 1.2$	1.185(0.109)	0.97
$\theta_{02} = 0$	-0.045(0.046)	0.93	$r_2 = 1.2$	1.188(0.079)	0.97
$\theta_{11} = 0$	0.053(0.045)	0.94	$\rho_1 = 0.85$	0.845(0.031)	0.95
$\theta_{12} = 0$	0.039(0.040)	0.95	$\rho_2 = 0.85$	0.849(0.030)	0.97

n=100					
Parameters	circular mean (c.d.)	CP	Parameters	mean (s.d.)	CP
$\theta_{01} = 0$	0.001(0.010)	0.93	$r_1 = 1.2$	1.200(0.069)	0.94
$\theta_{02} = 0$	-0.001 (0.009)	0.96	$r_2 = 1.2$	1.190(0.048)	0.97
$\theta_{11} = 0$	0.005(0.009)	0.92	$\rho_1 = 0.85$	0.849(0.020)	0.98
$\theta_{12} = 0$	-0.007(0.008)	0.96	$\rho_2 = 0.85$	0.849(0.020)	0.94

Table 4: Results of the simulation study with 10% zeroes in response only

n=50					
Parameters	circular mean (c.d.)	CP	Parameters	mean (s.d.)	CP
$\theta_{01} = 0$	-0.196(0.035)	0.93	$r_1 = 0.9$	0.880(0.061)	0.95
$\theta_{02} = \pi/2$	1.617(0.033)	0.91	$r_2 = 1.5$	1.457(0.114)	0.97
$\theta_{11} = 0$	0.187(0.039)	0.92	$\rho_1 = 0.85$	0.843(0.031)	0.96
$\theta_{12} = 0$	-0.038(0.024)	0.94	$\rho_2 = 0.85$	0.847(0.030)	0.96

n=100					
Parameters	circular mean (c.d.)	CP	Parameters	mean (s.d.)	CP
$\theta_{01} = 0$	-0.034(0.006)	0.91	$r_1 = 0.9$	0.894(0.027)	0.97
$\theta_{02} = \pi/2$	1.594 (0.012)	0.94	$r_2 = 1.5$	1.486(0.078)	0.96
$\theta_{11} = 0$	0.036(0.006)	0.92	$\rho_1 = 0.85$	0.846(0.021)	0.98
$\theta_{12} = 0$	-0.022(0.008)	0.95	$\rho_2 = 0.85$	0.845(0.021)	0.96

Table 5: Results of simulation study with 10% zeroes in covariate only

n=50					
Parameters	circular mean (c.d.)	CP	Parameters	mean (s.d.)	CP
$\theta_{01} = \pi/2$	1.646(0.053)	0.92	$r_1 = 0.9$	0.905(0.086)	0.96
$\theta_{02} = 0$	-0.053(0.033)	0.93	$r_2 = 1.5$	1.472(0.115)	0.96
$\theta_{11} = 0$	-0.081(0.057)	0.91	$\rho_1 = 0.85$	0.846(0.032)	0.95
$\theta_{12} = 0$	0.040(0.023)	0.97	$\rho_2 = 0.85$	0.845(0.030)	0.97
n=100					
Parameters	circular mean (c.d.)	CP	Parameters	mean (s.d.)	CP
$\theta_{01} = \pi/2$	1.579(0.012)	0.94	$r_1 = 0.9$	0.895(0.050)	0.97
$\theta_{02} = 0$	-0.001 (0.010)	0.92	$r_2 = 1.5$	1.485(0.077)	0.96
$\theta_{11} = 0$	-0.010(0.013)	0.95	$\rho_1 = 0.85$	0.846(0.021)	0.95
$\theta_{12} = 0$	0.010(0.007)	0.97	$\rho_2 = 0.85$	0.846(0.020)	0.94

Table 6: Results of the simulation study without zero-inflation

n=50					
Parameters	circular mean (c.d.)	CP	Parameters	mean (s.d.)	CP
$\theta_{01} = \pi/2$	1.646(0.007)	0.92	$r_1 = 0.3$	0.307(0.049)	0.96
$\theta_{02} = \pi/2$	1.553(0.009)	0.93	$r_2 = 0.3$	0.319(0.047)	0.96
$\theta_{11} = 0$	-0.081(0.022)	0.91	$\rho_1 = 0.85$	0.841(0.030)	0.95
$\theta_{12} = 0$	0.040(0.043)	0.97	$\rho_2 = 0.85$	0.848(0.030)	0.97
n=100					
Parameters	circular mean (c.d.)	CP	Parameters	mean (s.d.)	CP
$\theta_{01} = \pi/2$	1.574(0.001)	0.93	$r_1 = 0.3$	0.300(0.030)	0.95
$\theta_{02} = \pi/2$	1.577 (0.002)	0.92	$r_2 = 0.3$	0.310(0.028)	0.97
$\theta_{11} = 0$	-0.012(0.004)	0.94	ρ_1=0.85	0.848(0.020)	0.96
$\theta_{12} = 0$	0.010(0.013)	0.96	$\rho_2 = 0.85$	0.845(0.021)	0.95

Table 7: Comparison of Model I and Model II when data is generated from Model I with more than 40% zeros in both the response and covariate

Model I					
Parameters	circular mean (c.d.)	CP	Parameters	mean (s.d.)	CP
$\theta_{01} = 0$	0.007(0.006)	0.97	$r_1 = 0.9$	0.907(0.088)	0.91
$\theta_{02} = 0$	0.005 (0.009)	0.97	$r_2 = 1.2$	1.199(0.026)	0.93
$\theta_{11} = 0$	-0.004(0.002)	0.98	$\rho_1 = 0.93$	0.929(0.015)	0.94
$\theta_{12} = 0$	-0.003(0.001)	0.98	$\rho_2 = 0.95$	0.947(0.012)	0.93
BIC=333.1662					

Model II					
Parameters	circular mean (c.d.)	CP	Parameters	mean (s.d.)	CP
$\theta_{01} = 0$	0.007(0.002)	0.66	$r_1 = 0.9$	0.944(0.048)	0.92
$\theta_{02} = 0$	-0.004(0.002)	0.47	$r_2 = 1.2$	1.180(0.030)	0.93
$\theta_{11} = 0$	-0.007(0.002)	0.88	$\rho_1 = 0.93$	0.960(0.012)	0.76
$\theta_{12} = 0$	0.005(0.001)	0.99	$\rho_2 = 0.95$	0.948(0.011)	0.75
BIC=431.9582					

Table 8: Mean BIC values for Model I and Model II under mis-specified simulation model

p	δ	BIC-Model I	BIC-Model II
0.10	0.035	184.260	185.007
0.20	0.035	196.024	197.891
0.10	0.070	204.397	223.793
0.20	0.070	206.470	281.461

are allocated randomly.

$$\pi(\theta_{X_i}|-) \equiv \begin{cases} 0 & \text{with probability } p, \\ f_{TW}(\theta_{X_i}; \mu_{2i}, \rho_2, -\delta, \delta), & \text{with probability } 1-p. \end{cases}$$

and

$$\pi(\theta_{Y_i}|-) \equiv \begin{cases} 0 & \text{with probability } p, \\ f_{TW}(\theta_{Y_i}; \mu_{1i}, \rho_1, -\delta, \delta) & \text{with probability } 1-p. \end{cases}$$

We generate data of sample size $n = 50$ with $\beta_0 = b_0 = 1$, $\beta_1 = 0.9$, $b_1 = 1.2$, $\rho_1 = \rho_2 = 0.9$, for 100 replications. We compare Model I and Model II based on mean BIC values for four different combinations of δ and p and the results are presented in Table 8. It can be observed that Model I fits the data as good as Model II for small values of δ. As expected, the performance of Model I is much better compared to Model II even for moderately large values of δ.

5 Discussion

In this paper, a Bayesian methodology has been developed for estimation of parameters of circular-circular regression model with zero-inflated covariate and response unlike the existing frequentist methods that only model the cases with zero-inflated response. We provide a simple estimation methodology using MCMC algorithm based on data augmentation technique. We considered a two-stage circular-circular regression model with the Möbius transformation-based link function. As a special case, the methodology is applicable for conventional circular-circular regression with or without zero-inflated response and/or covariate. Also, one can implement the proposed methodology with other choices of link function and/or choices of angular error distributions. Moreover, as the latent variables involved in the modelling are continuous, one can easily modify the proposed methodology for two-stage circular-circular regression with discrete response variable. Although we have not explicitly considered missing data in our current analysis, but if the missingness is ignorable (i.e. missing at random) then a simple data augmentation technique has to be incorporated with the proposed methodology. For data with non-ignorable missingness, modelling and the estimation of the associated parameters are technically challenging, which will be addressed in future.

References

Bakshi, P. (2010). Evaluation of various surgical techniques in brunescent cataracts. *Unpublished thesis, Disha Eye Hospital, India.*

Bhattacharya, S. and Sengupta, A. (2009). Bayesian analysis of semiparametric linear-circular models. *Journal of Agricultural, Biological, and Environmental Statistics*, 9:14–33.

Bhuyan, P., Biswas, J., Ghosh, P., and Das, K. (2016). Bayesian two-stage regression model for zero-inflated longitudinal outcomes. *Technical Report No. ASU/2016/5, Applied Statistics Unit, Indian Statistical Institute, Kolkata, URL: http://www.isical.ac.in/ asu/TR/TechRepASU201605.pdf.*

Biswas, A., Dutta, S., Laha, A. K., and Bakshi, P. K. (2015). Response-adaptive allocation for circular data. *Journal of Biopharmaceutical Statistics*, 25:830–842.

Biswas, A., Dutta, S., Laha, A. K., and Bakshi, P. K. (2016a). Comparison of treatments in a cataract surgery with circular response. *Statistical Methods in Medical Research*, 5:2238–2249.

Biswas, A., Jha, J., and Dutta, S. (2016b). Modelling of circular random variables with a spike at zero. *Statistics and Probability Letters*, 109:194–201.

Busse, P. and Trocinska, A. (1999). Evaluation of orientation experiment data using circular statistic doubts and pitfalls in assumptions. *Ring*, 21:107–130.

Downs, T. D. and Mardia, K. V. (2002). Circular regression. *Biometrika*, 89:683–697.

Fisher, N. I. and Lee, A. J. (1992). Regression models for an angular response. *Biometrika*, 48:665–677.

Gould, A. L. (1969). A regression technique for angular variates. *Biometrics*, 25:683–700.

Heckman, J. (1974). Shadow prices, market wages and labor supply.. *Econometrica*, 42:679–694.

Heckman, J. (1979). Sample selection bias as a specification error. *Journal of the Royal statistical Society, Series B, Methodological*, 49:127–145.

Jha, J. (2017). Best approach direction for spherical random variables. *Technical Report No. ASU/2017/5, Applied Statistics Unit, Indian Statistical Institute, Kolkata, URL: http://www.isical.ac.in/ asu/TR/TechRepASU201705.pdf.*

Jha, J. and Biswas, A. (2017a). Circular-circular regression model with a spike at zero. *Technical Report No. ASU/2017/5, Applied Statistics Unit, Indian Statistical Institute, Kolkata, URL: http://www.isical.ac.in/ asu/TR/TechRepASU201705.pdf.*

Jha, J. and Biswas, A. (2017b). Multiple circular-circular regression. *Statistical Modelling*, 17:142–171.

Kato, S., Shimizu, K., and Shieh, G. S. (2008). A circular-circular regression model. *Statistica Sinica*, 18:633–645.

Lambert, D. (1992). Zero-inflated poisson regression, with an application to defects in manufacturing. *Technometrics*, 34:1–14.

Mackenzie, J. K. (1957). The estimation of an orientation relationship. *Acta Cryst*, 10:61–62.

Mardia, K. V. and Jupp, P. E. (2000). *Directional Statistics*. Wiley, London.

Min, Y. and Agresti, A. (2002). Modeling nonnegative data with clumping at zero: A survey. *Journal of Iranian Statistical Society*, 1:7–33.

Ravindran, P. and Ghosh, S. K. (2011). Bayesian analysis of circular data using wrapped distributions. *Journal of Statistical Theory and Practice*, 5(4):547–561.

Rivest, L. P. (1997). A decentred predictor for circular-circular regression. *Biometrika*, 84:318–324.

Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica*, 26:24–36.

Acknowledgement

The authors are thankful to Mr. Jayabrata Biswas for many helpful comments and suggestions.