SOME OUTLIER PROBLEMS IN A CIRCULAR REGRESSION MODEL

SAFWATI BINTI IBRAHIM

FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2013

SOME OUTLIER PROBLEMS IN A CIRCULAR REGRESSION MODEL

SAFWATI BINTI IBRAHIM

THESIS SUBMITTED IN FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

INSTITUTE OF MATHEMATICAL SCIENCES
FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2013

ABSTRAK

Kajian ini melihat kepada tiga masalah berkaitan dengan model regrasi bulatan JS dengan lima objektif untuk dicapai. Dua objektif pertama adalah berkaitan dengan masalah titik terpencil dalam model. Objektif pertama adalah penyiasatan keteguhan kaedah anggaran model regresi bulatan JS dengan kehadiran titik terpencil dalam set data. Objektif kedua adalah menggunakan tiga ujian berangka yang berasaskan kepada pendekatan penghapusan baris untuk mengesan titik terpencil yang mungkin dalam model regresi bulatan JS. Ujian pertama adalah mengambilkira versi statistik *COVRATIO* yang diubahsuai dengan menggunakan matrik kovarians reja dalam model regresi bulatan JS. Ujian berikutnya adalah menggunakan perbezaan purata ralat statistik dengan menggunakan fungsi kosinus dan sinus. Untuk setiap ujian, penjanaan nilai genting dan kuasa prestasi dipersembahkan melalui simulasi. Secara umum, ketiga-tiga ujian berangka menunjukkan prestasi yang baik dalam mengesan titik terpencil dalam model regresi bulatan JS.

Dua objektif seterusnya ialah melihat kepada pembangunan satu model regresi bulatan JS teritlak. Objektif ketiga adalah memperluaskan model regresi bulatan JS untuk memasukkan lebih dari satu pembolehubah bulatan tidak bersandar. Formula umum model regrasi bulatan JS teritlak dan anggaran parameter regresi menggunakan kaedah kuasa dua terkecil dibentangkan. Prestasi kaedah anggaran disiasat menggunakan simulasi dan secara umumnya adalah bagus. Objektif keempat membincangkan masalah multikolinearan dalam model regresi bulatan teritlak. Satu kaedah yang diubahsuai untuk mengesan kewujudan multikolinearan berdasarkan faktor inflasi varians dicadangkan untuk disesuaikan dengan sifat model regresi bulatan JS teritlak. Jika multikolinearan wujud, kami menggunakan idea analisis regresi rabung

bagi mendapatkan anggaran parameter di dalam model. Prosedur yang dicadangkan berfungsi dengan baik apabila dilaksanakan pada data set simulasi dan data set sebenar.

Objektif yang terakhir ialah membangunkan satu rangka kerja model hubungan fungsian dengan menggunakan model regresi bulatan JS dalam pembangunan tersebut. Di sini, kami menganggap kedua-dua pembolehubah bulatan bersandar dan pembolehubah bulatan tidak bersandar mengandungi ralat. Penganggaran parameter diperoleh secara berangka meggunakan kaedah pengganggaran pelelaran kebolehjadian maksimum. Disebabkan kerumitan penganggar parameter, ralat piawai bagi penganggar di dapati dengan menggunakan kaedah cangkuk.

Untuk ilustrasi, tiga data set bulatan yang sebenar dipertimbangkan, iaitu, set data arah pergerakan angin, set data mata dengan dua pembolehubah dan set data mata dengan empat pembolehubah.

ABSTRACT

This study looks at three problems related to the JS circular regression model with five objectives in mind. The first two objectives are concerned with the problem of outliers in the model. The first is the investigation of the robustness of the JS circular regression estimation method in the presence of outliers in the data set. The second is the use of three numerical tests based on row deletion approach to detect possible outliers in the JS circular regression model. The first test considered is a modified version of the *COVRATIO* statistic by utilizing the covariance matrix of residuals of the JS circular regression model. The other tests are based on the difference mean circular error statistics using cosine and sine functions. For each test, the generation of cut-off points and the power of performance are presented via simulation. In general, the three numerical tests perform well in detecting outliers in JS circular regression model.

The next two objectives look at the development of a new generalized JS circular regression model. The third looks at extending the JS circular regression model to include more than one circular explanatory variable. The general formulation of the generalized JS circular regression model and the estimation of the regression parameters using the least squares method are presented. The performance of the estimation method is investigated via simulation and is generally good. The fourth looks at the problem of multicollinearity in the generalized model. A new modified procedure to detect the presence of multicollinearity based on the variance inflation factor is proposed to suit the nature of the generalized JS circular regression model. If the multicollinearity does exist, we use the idea of the ridge regression analysis to find the parameter estimates of the model. The proposed procedure works well when implemented on simulated and real data sets.

The last objective is to develop a new functional relationship model framework by using the JS circular regression model in the setup. Here, we assume both the circular dependent and explanatory circular variables are subject to errors. The parameter estimates may be obtained numerically using iterative procedure on the maximum likelihood estimators. Due to the complexity of the estimators, the standard errors of the estimates are obtained using bootstrap method.

For illustration, three real circular data sets are considered, namely, wind direction data set, eye data set with two variables and another multivariate eye data set with four variables.

ACKNOWLEDGEMENTS

Firstly, I would like to express my sincere gratitude to my supervisors, Assoc. Prof. Dr. Ibrahim Mohamed and Prof. Dr. Abdul Ghapor Hussin (UPNM), for their guidance, support, invaluable help, encouragement and supervision throughout my research. Their understanding, patience and valuable advice have been the keys to the success of this study. Also a special thank is extended to Professor S. R. Jammalamadaka for his constructive suggestions during his visit to the university.

With deep sense of gratitude, I would like to thank my husband, Mohd Irwan bin Yusoff, and my son, Muhammad Afiq Ikhwan, for their love, understanding, support and encouragement at all stage of this study. Also to my parents and family members who have always supported me throughout the period. Without their tremendous love and support, I would not have been able to concentrate on my study and endure some difficult times through all these years.

I am grateful to the Ministry of Higher Education Malaysia and University of Malaya for providing the scholarship. I also would like to gratefully acknowledge the support given by the staff members of the Institute of Mathematical Sciences, University of Malaya especially Puan Budiyah. Special thank also goes to my friends; Adzhar, Ikhwan, Adia, Mardziah, Dharini and others whom have always directly or indirectly motivate me all along this journey. And Alhamdulillah, Praise to Allah for His Blessings and without His Will, this study will never be completed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATION

κ	Concentration parameter
$VM(\mu,\kappa)$	The von Mises distribution with mean direction μ and concentration parameter κ
R	Resultant length
$\overline{R}$	Mean resultant length
$I_0(\kappa)$	The modified Bessel function of the first kind and order zero.
$\overline{\theta}$	Sample mean direction
ϕ	Sample median direction
ν	Resistant constant for circular boxplot
$CIQR$	Circular interqurtile range
$\boldsymbol{r}_A$	Circular residuals
$A(\kappa)$	Ratio of Bessel functions, a measure of goodness-of-fit for circular regression model
d_{ij}	Circular distance between two circular observations θ_i and θ_j
r_c	Circular correlation coefficient
$COVRATIO_{(-i)}$	The determinantal ratio of covariance matrices for reduced and full data.
MCEc	Mean circular error in terms of cosine function
MCEs	Mean circular error in terms of sine function
DMCEc	The maximum absolute difference between *MCEc* for full and reduced data
DMCEs	The maximum absolute difference between *MCEs* for full and reduced data
λ	Contamination level

CHAPTER ONE

INTRODUCTION

1.1 Background of the Study

Circular or directional statistics is a branch of statistics that deals with data points distributed on a circle. It uses angles as the measurements of directions ranging from 0° to 360° or in radians $(0,2\pi]$. It can be displayed on the circumference of a unit circle. Circular data arise in various ways including those corresponding to two circular measuring instruments, for instance the compass and the clock, and broadly used in different areas such as

(i) **Natural science**: Rivest (1997) predicted the direction of ground movement during an earthquake, while Downs & Mardia (2002) had applied their proposed circular regression models on circular data.

(ii) **Medical sciences**: Downs *et al*. (1970) studied the correlations among circadian biological rhythms wherein a 24-hour clock is considered as a circle (Binkley, 1990; Downs, 1974; Moore-Ede *et al*., 1982) and the angle of knee flexion as a measure of recovery of orthopaedic patients (Jammalamadaka *et al.*, 1986).

(iii) **Meteorology**: Data include wind and wave directions (Mardia, 1972; Johnson & Wehrly, 1977; Hussin *et al.*, 2004 and Gatto & Jammalamdaka, 2007), the number of times a day at which thunderstorms occur and the frequencies of heavy rain in a year (Mardia & Jupp, 2000).

(iv) **Biology**: The bird orientation in homing or migration (Mardia, 1972), animal navigation (Batschelet, 1981) and spawning times of a particular fish (Lund, 1999).

(v) **Physics**: Fractional part of atomic weights (von Mises, 1918) and source of signals in the case of airplane crashes (Lenth, 1981).

(vi) **Psychology**: Studies of mental maps to represent the surroundings of respondents (Gordon *et al.*, 1989) and time pattern in crime incidence (Brunsdon & Corcoran, 2006).

(viii) **Geology**: Modelling the cross-bedding data (Jones & James, 1969), the orientations of fractures and fabric elements in deformed rocks (Mardia, 1972) and the direction of earthquake displacement (Rivest, 1997).

(ix) **Geography:** Orientation data appear naturally when readings consist of longitudes and latitudes such as the longitude and latitude of each shock, the variation of the number of earthquakes (Mardia, 1972).

Due to the bounded property of circular observations, we need to consider special statistical methods in analyzing such data, that is, the descriptive and inferential analysis of circular data cannot be carried out using standard methods for observations on Euclidean space (Agostinelli, 2007). For example, let us consider two angles 10° and 350° as illustrated in Figure 1.1. The arithmetic mean by treating the data as linear observation is 180°. However, the mean direction of the two directions has to be 0°. Therefore, special statistical methods and techniques are needed to analyse circular data.

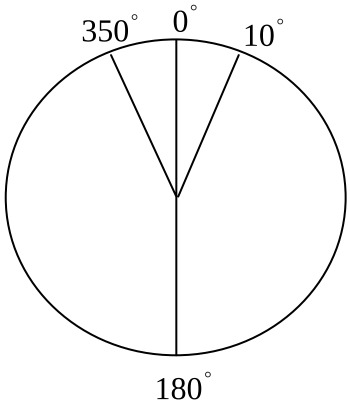


Figure 1.1: Arithmetic and Geometric mean

Currently, the analysis of circular data attracts the interest of statisticians and researchers from different scientific fields due to the accessibility to such data. As a result, new statistical methods for circular data have been developed and statistical softwares have been produced that provide tools for circular data analysis including Axis, Oriana, DDSTP and MATLAB. In addition, Jammalamadaka & SenGupta (2001) provided some routines for circular data analysis written in R/S-Plus language. More routines are expected to be available in the future with more problems related to circular data analysis being studied including the occurrence of outliers in circular samples and circular regression models.

Strong interests on circular regression model have also been shown (see Gould, 1969; Mardia, 1972; Laycock, 1975; Down & Mardia, 2002; Hussin *et al.*, 2004 and Kato *et al.*, 2008). Another model of our interest is proposed by Jammalamadaka & Sarma (1993) for the case when both response variable v and explanatory variable u are circular. They used the conditional expectation of the vector e^{iv} given u to represent the relationship between v and u. The properties of the models for the case of a single explanatory variable have been studied (see Sec. 8.6 of Jammalamadaka & SenGupta, 2001). We refer the model as the *JS circular regression model*. We will extend the model by introducing p circular explanatory variables in the model.

The challenge to detect the existence of outliers in circular regression models has not received enough attention yet. So far, very few published papers focusing on the detection of outliers in circular regression models can be found in the literature such as Hussin *et al.* (2004) and Abuzaid *et al*. (2008). Here, the problem of our interest is the detection of outliers on the JS circular regression model. The circular residual is used for the above purpose. Recently, Abuzaid *et al*. (2008, 2011, 2013) discussed the issues and proposed several statistics for detecting outliers in Hussin's circular regression model. We employ the statistics and investigate their performance when applied on JS circular regression model.

Multiple linear regressions are used to model linear relationship between a dependent variable and one or more independent variables. However, when the independent variables are highly correlated, then we have a problem of multicollinearity. The problem can make it impossible or difficult to assess the relative importance of individual predictors from the estimated coefficients of the regression equation (Farrar & Glauber, 1967). It can also have severe effects on the estimated coefficients in a multiple regression analysis. A number of studies have been carried out on overcoming the problem of multicollinearity in the linear regression model (see Farrar & Glauber, 1967; Lemieux, 1978; Mansfield & Helms, 1982; Montgomery & Peck, 1992 and Haan, 2002). The presence of multicollinearity problem can be detected by looking at the correlation structure of the independent variables. If the variables are orthogonal, then the problem does not exist. A more objective way is by looking at the values of variance inflation factor (*VIF*), see for example, Lemieux (1978) and Mansfield & Helms (1982). Any *VIF* values different from one indicates the presence of multicollinearity. Now, the main goal is to deal with the multicollinearity when estimating the parameters of the regression model. Hoerl (1962) and Hoerl & Kennard (1968) suggested a noble approach by introducing a constant k in the LS estimates of the regression model which is believed to be able to control the effect of the problem. But, to the best knowledge of the author, both issues of a multiple circular regression and the problem of multicollinearity in the circular regression model has not been considered yet. The interest here is to develop a generalized circular regression model with more than one explanatory variable in the model and to propose a modified procedure of dealing with multicollinearity problem in the JS circular regression models.

The study of linear “measurement error” or “errors-in-variables” models (EIVM) began more than a century ago. Since then this area has gained importance for the study of relationships between variables. If errors in the explanatory variables are ignored, it is well known that the estimators obtained by classical or ordinary linear regression are biased and inconsistent (Buonaccorsi, 1996). In practice most studies, for example, the life sciences, biology, ecology and economics involve variables that cannot be recorded exactly. When the purpose is to estimate relationships between groups or populations, errors arise mostly as experimental and observational errors or as errors representing variability of individual subjects.

The functional model is a part of EIVM where it refers to the study of the relationship between variables which are subjected to errors (Ramsay & Silverman, 2005). The functional relationship models in linear case have been extensively developed in the literature. The interest here is to extend the linear functional relationship model to the case when both variables are circular and subjected to errors. Works on functional relationship models for circular variables have also been reported in the past (see Hussin, 1997; Bowtell & Patefield, 1999 and Caires & Wyatt, 2003). We will consider the circular regression model proposed by Jammalamadaka & Sarma (1993) in the setup. Herewith, we refer the model as the JS circular functional relationship model. Note that we only consider unreplicated data in this study.

1.2 Statement of the Problem

So far, the study on circular regression is limited to a single independent variable only. One of the models is the JS circular regression model which is known to have very interesting properties closely related to the theory of multiple linear regression. In this study, we look at three main problems which have not been explored yet related to this model; firstly, we use three methods of identifying outliers in the model; secondly, we generalize the JS circular regression model to multiple case and then we look at the problem of multicollinearity in the model; and thirdly, we utilize the model in the functional relationship framework.

1.3 Objectives

Based on the statement of problem above, the researcher has outlined the following objectives for this study:

1. To investigate the robustness of the JS circular regression estimation method toward the existence of outliers in the data.
2. To extend the use of three different methods of detecting outliers in JS circular regression model, i.e., *COVRATIO*, *DMCEc* and *DMCEs* statistics.
3. To extend the JS circular regression model to generalized JS circular regression model with two or more independent variables.
4. To propose a new procedure of dealing with multicollinearity problem in the JS circular regression models.
5. To develop a new functional relationship model framework based on the JS circular regression models.
6. To apply the methods on real life data.

1.4 Significance of Study

The findings from this study will be beneficial in the following ways:

1. Contribute to the knowledge regarding the modelling of circular regression and detection of outliers.
2. Optimize the estimation of parameters in circular models by the identification of outliers.
3. Contribute to the new method of dealing with multicollinearity problem and optimize the estimation.
4. Contribute to the functional model framework in circular regression model.

1.5 Research Outline

This research attempts to handle the problem of outliers in circular regression by proposing three statistical methods, then, look at the problem of multicollinearity in the model using Hoerl and Kennard's method; and lastly, propose the functional relationship framework for JS circular regression model. The research is outlined as follows:

Chapter two gives a literature review on the circular regression models and the problem of outliers in circular regression models. We present a review on different methods of identification of outliers in linear regression which has the possibility to be extended to the model proposed by Jammalamadaka & Sarma (1993). Special discussions are also presented on the problem of multicollinearity in the multiple linear regression model and on the setup of the functional relationship model involving circular regression models.

Chapter three discusses the theory of JS circular regression model and the general effect of outliers on the model as well as the robustness property of the model. We illustrate the application of the model on two real data sets.

Chapter four presents a numerical statistic based on the idea of *COVRATIO* statistic in linear regression that can be used to detect possible outliers in the JS circular regression models. Via simulation, the cut-off points are obtained and the power of performance is investigated. The statistic is then applied on wind direction data.

Chapter five presents another two numerical statistics to detect possible outliers in the JS circular regression models by considering the difference mean circular error using cosine and sine functions. The cut-off points and the power of performance are investigated. The statistics are then applied on local eye data.

Chapter six presents the development of the generalized JS circular regression models and the treatment of multicollinearity problem in the multiple JS circular regression models using ridge regression approach. We use the multivariate eye data to illustrate the problem.

Chapter seven presents the development of functional relationship model involving JS circular regression models. The parameters are estimated using maximum likelihood estimation method. Due to the large number of parameters to be estimated, the asymptotic variances of the parameters are obtained using bootstrap procedure. Extensive simulation studies are used to show the performance of the maximum likelihood estimation method. For illustration, we apply the method on bivariate eye data.

Chapter eight presents the summary of the study and the suggestion for further research.

Lists of appendices are attached including the eye data sets, wind direction data, simulation results, and the S-Plus subroutines used in this thesis.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

This chapter reviews the theory on circular statistics, circular regression model and the detection of outliers in linear and circular regression models. We also look at the problem of multicollinearity and the theory on linear functional relationship model with the intention to be extended to the case involving JS circular regression models.

2.2 Circular Statistics

There has been a great deal of interest for the past few decades in the study of data of circular type. For describing circular data, development of descriptive measures and the investigation of special characteristics of circular data have been studied, see for example, Rao (1969), Lenth (1981), He & Simpson (1992), Jammalamadaka & Sarma (1993), Kato *et al.* (2008) and Abuzaid *et al.* (2012). We present some numerical and graphical techniques that can be used to describe circular data.

2.2.1 Numerical Statistics

Let $\theta_1,...,\theta_n$ be observations in a random circular sample of size n from a circular population. Some of the circular descriptive measures are as follows:

(i) The mean direction

To find the mean direction, we consider each observation θ_i as a unit vector and calculate the corresponding values of $\cos\theta_i$ and $\sin\theta_i$. The resultant length R is then given by

$$R = \sqrt{C^2 - S^2}\,, \tag{2.1}$$

where $C = \sum_{i=1}^{n} \cos\theta_i$ and $S = \sum_{i=1}^{n} \sin\theta_i$. The mean direction, denoted by $\bar{\theta}$, is given by the argument of the vector $z = C + iS$ giving

$$\bar{\theta} = \begin{cases} \tan^{-1}(S/C) & \text{if } C \geq 0 \\ \tan^{-1}(S/C) + \pi & \text{if } C < 0. \end{cases} \tag{2.2}$$

One of the characteristics of mean direction is that $\sum_{i=1}^{n} \sin(\theta_i - \bar{\theta}) = 0,$ which is similar to that of the linear case.

(ii) The median direction

Mardia & Jupp (2000) defined the median as any point ϕ, where half of the data lie in the arc $[\phi, \phi + \pi)$ and the other points are nearer to ϕ than to $\phi + \pi$. For any circular sample, Fisher (1993) defined the median direction as the observation ϕ which minimizes the summation of circular distances to all observations,

$$d(\phi) = \pi - \sum_{i=1}^{n} \left|\pi - |\theta_i - \phi|\right| \text{ for } i = 1, ..., n. \tag{2.3}$$

Fisher's definition is used to obtain the circular median in the Oriana statistical software package.

(iii) The mean resultant length

Mean resultant length $\bar{R}$ is defined as the length of the centre of vector $z = C + iS$ and is useful for unimodal data to measure the concentration of the circular data towards the centre. It is defined by $\bar{R} = R/n$ and lies in the range [0,1] where R is given in (2.1) above.

(iv) The sample circular variance

The sample circular variance is defined by the quantity $V = 1 - \bar{R}$, where $0 \le V \le 1$. The smaller value of circular variance refers to a more concentrated sample. However, this measure is rarely used compared to the other measures of circular concentration, in particular, the concentration parameter κ to be described later.

(v) The sample circular standard deviation

The quantity $\nu = \sqrt{-2\log(1-V)}$ is defined as the sample circular standard deviation with $0 < \nu < \infty$, where V is the sample circular variance. The reason for defining the circular standard deviation in this way rather than as the square root of the sample circular variance is to obtain some reasonable approximations for proportion of von Mises distribution provided the distribution is not too dispersed (see Fisher (1993, p.54)).

(vi) The concentration parameter

The concentration parameter, denoted by κ, is a standard measure of dispersion for circular data. Best and Fisher (1981) gave the maximum likelihood estimates of the concentration parameter κ as follows

$$\kappa = \begin{cases} 2\bar{R} + \bar{R}^3 + \frac{5}{6}\bar{R}^5, & \text{if} \quad \bar{R} < 0.53 \\ -0.4 + 1.39\bar{R} + \frac{0.43}{(1-\bar{R})}, & \text{if} \quad 0.53 \le \bar{R} < 0.85 \\ (\bar{R}^3 - 4\bar{R}^2 + 3\bar{R})^{-1}, & \text{if} \quad \bar{R} \ge 0.85, \end{cases}$$

where $\bar{R}$ is mean resultant length. The values of κ lies in the range $[0, \infty)$. The larger the value of κ suggests the circular data are more concentrated in the direction of the mean direction $\bar{\theta}$. When κ is closer to zero, the circular data is more uniformly distributed around the unit circle.

(vii) Circular distance between circular observations

Jammalamadaka & SenGupta (2001) defined the circular distance between any two circular observations to be the smaller of the two arc lengths between the two points along the circumference of a unit circle. For any two angles ϕ and θ, the circular distance is given by

$$d_{\circ}(\phi,\theta) = \min(\phi-\theta, 2\pi-(\phi-\theta)) = \pi - \left|\pi - |\phi-\theta|\right| \qquad (2.4)$$

where $d_{\circ}(\theta,\hat{\theta}) \in [0,\pi]$. If the circular distances between observation θ_i and its neighbours on both sides are relatively larger than the distance between other successive observations, then θ_i may be considered as an outlier.

2.2.2 Graphical Techniques

Several graphical techniques can be used to explore circular samples, in particular the von Mises sample, and will be described here. These graphical techniques can also be used for the purpose of detecting outliers in the data set.

(i) P-P plot

P-P plot can be obtained by finding the best fitting of cumulative von Mises distribution $\hat{F}(\theta;\hat{\mu},\hat{\kappa})$ for the circular sample. Then the plot is obtained by plotting the pairs of $\left(i/(n+1), \hat{F}(\theta_{(i)};\hat{\mu},\hat{\kappa})\right)$, $i=1,\ldots,n$, where $\theta_{(i)}$ are the ordered observations with respect to origin and n is the sample size. Any point in P-P plot that seems not to be close enough to the diagonal line is suspected to be outliers.

(ii) Q-Q plot

Q-Q plot is obtained by plotting $(\sin(q_i/2), z_{(i)})$, where $q_i = F^{-1}\left(i/(n+1);0,\hat{\kappa}\right)$ and $z_i = \dfrac{\sin(\theta_i - \hat{\mu})}{2}$, $i = 1,...,n$ and $z_{(1)},...,z_{(n)}$ are the ordered values of z_i's. Any points in Q-Q plot that are relatively far from the diagonal line are candidate to be outliers.

(iii) Spoke Plot

The spoke plot is introduced by Hussin *et al.* (2007). It consists of inner and outer rings for θ_i and ϕ_i, $0^\circ \le \theta_i, \phi_i < 360^\circ$, respectively and straight lines are used to connect the pair of points (θ_i, ϕ_i) between the two circular variables. The lesser number of lines crossing the inner ring indicates higher correlation between the two variables.

(iv) Circular Boxplot

Boxplot has been widely used in the linear exploratory data analysis. One of its applications is to identify extreme values and outliers in a univariate data set. It was developed by Tukey (1977) for linear univariate samples. This type of boxplot, however, is not suitable for a circular data due to the fact that there is no natural ordering in circular observations. This motivates Abuzaid *et al*. (2011) to develop a special boxplot for circular variables, called the circular boxplot.

The circular boxplot is also useful to identify possible outliers in circular samples. Five circular summary statistics are used. The circular median is obtained using the definition given by Fisher (1993) and in the case of prior knowledge about the circular distribution, Mardia (1972) defined the median direction ϕ as the solution of

$$\int_{\phi}^{\phi+\pi} f(\theta)\,d\theta = \int_{\phi+\pi}^{\phi+2\pi} f(\theta)\,d\theta = 0.5\,,$$

where $f(\theta)$ is the probability density function of θ. Meanwhile, the first and third quartile directions Q_1 and Q_3 are defined as

$$\int_{\phi-Q_1}^{\phi} f(\theta)d\theta = 0.25 \text{ and } \int_{\phi_1}^{\phi+Q_3} f(\theta)d\theta = 0.25\,,$$

respectively. In finding the circular interquartiles range $CIQR$**,** the observations are transformed using the formula $\theta_i - \bar{\theta}$ giving the new first and third quartiles, Q_1' and Q_3', respectively. Hence, the *CIQR* is obtained using the formula

$$CIQR = 2\pi - Q_3' + Q_1'.$$

Next, we find the upper and lower fences of the circular boxplot. Using the idea in linear boxplot, the lower fence of the circular boxplot is given by $L_F = Q_1 + \nu \times CIQR$ and the upper fence is $U_F = Q_3 - \nu \times CIQR$, where ν is a suitable resistant constant. Abuzaid *et al*. (2012) proposed the resistant constant to be $\nu = 1.5$ for $2 \le \kappa \le 3$ and $\nu = 2.5$ for $\kappa > 3$. Examples of the circular boxplot for symmetric simulated circular data is given in Figure 2.1.

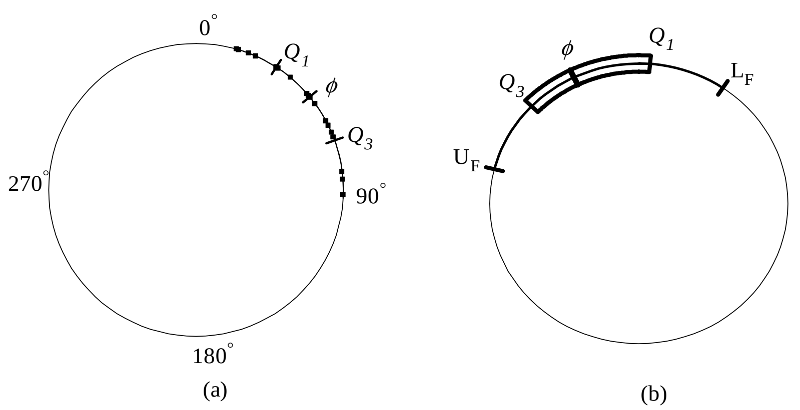


Figure 2.1: Circular boxplot

2.3 Circular Distributions

A circular distribution is a probability distribution which total probability is concentrated on the circumference of a unit circle. Each point on the circumference represents a direction. Circular distributions are essentially of two types; they may be discrete, assigning probability masses only to a countable number of directions, or may be absolutely continuous with respect to the measure on the circumference of a unit circle. There are a number of important circular distributions that have been developed in the past (see Mardia, 1972; Fisher, 1993; Ko, 1992; Jammalamadaka & SenGupta, 2001 and Agostinelli, 2007) including the von Mises (Normal Circular) distribution, Wrapped Normal distribution, Wrapped Cauchy distribution and Cardioid distribution.

The most common distribution used for circular data is the von Mises distribution. The distribution can be approximated by normal distribution for large concentration parameter. Besides, Jammalamadaka & SenGupta (2001) reviewed the wrapped α stable distribution with the wrapped Cauchy and the wrapped normal distributions as the special cases. We present some of these distributions below.

2.3.1 The von Mises (*VM*) Distribution

The von Mises distribution is introduced by von Mises (1918) to study the deviations of measured atomic weight from integral values. It is the most common distribution considered for unimodal samples of circular data. The von Mises distribution has been extensively discussed where many inferential techniques have been developed. It is denoted by $VM(\mu,\kappa)$, where μ is the mean direction and κ is the concentration parameter. The probability density function for the von Mises distribution is given by

$$f(\theta;\mu,\kappa)=\frac{1}{2\pi\, I_0(\kappa)}\exp\{\kappa\cos(\theta-\mu)\}, \qquad 0<\theta,\mu\le 2\pi \text{ and } \kappa\ge 0,$$

where $I_0(\kappa)$ is the modified Bessel function of the first kind and order zero, and it is given by Fisher (1993) where

$$I_0(\kappa)=\frac{1}{2\pi}\int_0^{2\pi}\exp(\kappa\cos\theta)\,d\theta=\sum_{r=0}^{\infty}\left(\frac{\kappa}{2}\right)^{2r}\left(\frac{1}{r!}\right)^{2}.$$

Some of the von Mises density properties are:

(i) it is symmetrical about the mean direction μ,

(ii) it has a mode at μ, and

(iii) it has antimode at $(\mu\pm\pi)$.

Best & Fisher (1981) gave the maximum likelihood estimates of the concentration parameter κ as follows:

$$\hat{\kappa}=\begin{cases}2\bar{R}+\bar{R}^3+\frac{5}{6}\bar{R}^5, & \text{if} \quad \bar{R}<0.53,\\ -0.4+1.39\bar{R}+\frac{0.43}{(1-\bar{R})}, & \text{if} \quad 0.53\le\bar{R}<0.85,\\ (\bar{R}^3-4\bar{R}^2+3\bar{R})^{-1}, & \text{if} \quad \bar{R}\ge 0.85.\end{cases}$$

2.3.2 The Wrapped Normal (*WN*) Distribution

A wrapped normal distribution is obtained by wrapping a normal distribution around a unit circle. The normal distribution is denoted by $N(\mu_L,\sigma_L^2)$ where μ_L is the mean and σ_L^2 is the variance while the *WN* distribution is denoted by $WN(\mu,\rho)$, where μ is the mean direction and ρ is the measure of concentration parameter. Its probability distribution function is given by

$$f(\theta)=\frac{1}{\sigma\sqrt{2\pi}}\sum_{\kappa=-\infty}^{\infty}\exp\left[\frac{-(\theta-\mu-2\kappa\pi)^2}{2\sigma^2}\right],$$

where σ^2 is the circular variance.

From Whittaker and Watson (1944), an alternative and more useful representation of this density is

$$f(\theta)=\frac{1}{\sqrt{2\pi}}\left[1+2\sum_{\kappa=-\infty}^{\infty}\rho^{\kappa^2}\cos\kappa(\theta-\mu)\right], \qquad 0\le\theta<2\pi, \qquad 0\le\rho<1.$$

The distribution is unimodal and symmetric about the value $\theta=\mu$. Unlike the von Mises distribution, the *WN* distribution possesses the additive property, that is, the convolution of two *WN* variables is also *WN*. Specifically, if $\theta_1 \sim WN(\mu_1,\rho_1)$, $\theta_2 \sim WN(\mu_2,\rho_2)$, and are independent, then $\theta_1+\theta_2 \sim WN(\mu_1+\mu_2,\rho_1+\rho_2)$.

2.3.3 The Wrapped Cauchy (*WC*) Distribution

A wrapped Cauchy distribution is obtained by wrapping the Cauchy distribution on the real line with density

$$f(\theta)=\left(\frac{1}{\pi}\right)\frac{\sigma}{\sigma^2+(\theta-\mu)^2}, \qquad -\infty<\theta<\infty,$$

around the circle. The probability density function is as follows

$$\begin{aligned} f(\theta)&=\frac{1}{2\pi}\left(1+2\sum_{k=1}^{\infty}\rho^{k}\cos k(\theta-\mu)\right)\\ &=\frac{1}{2\pi}\frac{1-\rho^2}{1+\rho^2-2\rho\cos(\theta-\mu)}, \qquad 0\le\theta<2\pi, \end{aligned}$$

where $\rho=e^{-\sigma}$. The equality of the two expressions above is verified by equating the real parts of the geometric series identity

$$\sum_{k=1}^{\infty} a^k = \frac{a}{1-a}, \qquad 0 < a < 1,$$

with $a = \rho e^{-i(\theta-\mu)}$. The distribution is unimodal and symmetric.

2.4 Circular Regression Models

A number of circular regression models have been proposed by a number of authors. Lund (1999) proposed a regression model where the independent variables consist of one circular variable and a set of linear variables. For a circular response $\boldsymbol{V}$, a circular predictor ϕ and a set of linear covariates $\boldsymbol{X}$, the least circular distance regression model is given by

$$v = \mu(\phi, \boldsymbol{X}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2) + \varepsilon,$$

where $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ are vectors of parameters and $\boldsymbol{\varepsilon}$ is the random circular error with mean direction $\boldsymbol{0}$. The parameter estimates are obtained by maximizing the average cosine residuals of the model given by

$$\hat{L}(\beta, \alpha, r) = \frac{1}{n} \sum_{i=1}^{n} \cos\{v_i - \mu(\phi, \boldsymbol{X}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2)\}.$$

Earlier, Mardia (1972) considered a model by assuming each of the response circular variable θ_i, $i = 1,...,n$, to be independently distributed from von Mises distribution with mean direction μ_i and unknown concentration parameter κ. A slightly different model was also proposed by Fisher & Lee (1992). The model is originally proposed by Gould (1969) in predicting the mean direction of a circular response variable Θ from a vector of linear covariates $\boldsymbol{X} = (x_1,..., x_k)$ given by

$$\mu = \mu_0 + \sum_{j=1}^{p} \beta_j x_j ,$$

where μ_0 and β's are unknown parameters and x_j is a linear covariate, for *j*= 1, ..., *p*.

For the case when both response and explanatory variables are circular, say ***U*** and ***V*** respectively, a few circular regression models have been proposed using different approaches. The earliest model is proposed by Laycock (1975) who expressed the model as a multiple linear regression model with complex entries. On the other hand, Rivest (1997) proposed a circular–circular regression model to predict the *y*-direction based on the rotation of the decentred *x*-angle. The model is given by

$$y = \theta(x; \beta, \alpha, r) + \varepsilon ,$$
$$\theta(x; \beta, \alpha, r) = \beta + \tan^{-1}\{\sin(x-\alpha), r + \cos(x-\alpha)\} \pmod{2\pi} ,$$

where β and α are angles belonging to $[0,2\pi)$, r is real number and ε has a distribution with mean *0*. The parameters are then estimated by maximizing the average cosine residuals.

In cases when ***U*** and ***V*** are circular variables with mean directions α' and β' respectively, Down & Mardia (2002) applied the following mapping to relate *u* and *v* such that

$$\tan\frac{1}{2}(v - \beta') = \omega \tan\frac{1}{2}(u - \alpha') ,$$

where ω is a slope parameter in the closed interval [-1,1]. The mapping defines a one-to-one relationship with a unique solution given by

$$v = \beta' + 2\tan^{-1}\left\{\omega \tan\frac{1}{2}(u - \alpha')\right\} .$$

Later, Hussin *et al*. (2004) proposed a simple circular-circular regression model involving one independent variable only given by

$$v_i = \alpha'' + \beta'' u_i + \varepsilon_i \pmod{2\pi},$$

where ε_i is circular random error having a von Mises distribution with circular mean *0* and concentration parameter κ and α'' and β'' are the coefficients of the model. The model is useful when we are interested to find a direct relationship between the two circular variables, for example, in modelling a relationship for calibration between two instruments.

Another interesting model is proposed by Jammalamadaka & Sarma (1993) who considered the conditional expectation of the vector e^{iv} given u to represent the relationship between u and v and utilized the definition of characteristic function of a complex number. The model is given by

$$E(e^{iv} \mid u) = \rho(u)e^{i\mu(u)} = g_1(u) + ig_2(u),$$

where $\mu(u)$ is the conditional mean direction of v given u with conditional concentration $0 \le \rho(u) \le 1$. Due to the difficulty of estimating $g_1(u)$ and $g_2(u)$, the functions are expressed in terms of their trigonometric polynomial expansions.

In this study, we consider the model proposed by Jammalamadaka & Sarma (1993) which is already referred to in Chapter one as the *JS circular regression model*. We will describe the model in detail in Chapter three. The adequacy of the model will also be investigated and procedures of detecting outliers in the model will be developed in the subsequent chapters.

2.5 Outliers and Influential Observations in Regression Models

Outlier is a common problem in the statistical analysis. It is defined as an observation that is very different to the other observations in a set of data. Beckman & Cook (1983) and Barnett & Lewis (1994) defined an outlier in a set of data to be an observation (or subset of observations) which appears to be inconsistent with the remainder of that set of data. Different approaches to deal with the outliers in various areas can be found in the literature. However, there is few published work discussing the problem of outliers in regression models for circular variables. Overview on outliers in linear and circular regression analysis is given in the following section.

2.5.1 Outliers in Linear Regression Model

This section reviews some of the techniques used to identify outliers in linear regression based on row deletion approach, see Belsley *et al.* (1980). It investigates the impact of deleting one row at a time from the design matrix $\boldsymbol{X}$ and vector $\boldsymbol{Y}$ on the fitted values, residuals and the estimated parameter. Here, the interest is in identifying suitable methods that can be extended to the circular regression case.

Regression analysis concerns with fitting models to data in which there is a single continuous response variable whose expected values depend on the values of the explanatory variables. Linear regression model is given by

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}, \qquad (2.5)$$

where $\boldsymbol{Y}$ is n–vector of response, $\boldsymbol{X}$ is $n \times p$ full rank matrix of known constants, $\boldsymbol{\beta}$ is p-vector of unknown parameters and $\boldsymbol{\varepsilon}$ is n-vector of errors with the assumptions that $E(\boldsymbol{\varepsilon}) = \boldsymbol{0}$ and $V(\boldsymbol{\varepsilon}) = \sigma^2 \boldsymbol{I}_n$. The least squares estimation of $\boldsymbol{\beta}$ is given by

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}, \qquad (2.6)$$

where $E(\widehat{\boldsymbol{\beta}}) = \boldsymbol{\beta}$ and $\operatorname{cov}(\widehat{\boldsymbol{\beta}}) = \sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}$. The residual sum of squares about the fitted model is given by $\text{SSE} = (\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}})'(\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}})$ while the least squares estimator of σ^2 is an unbiased estimator defined by $s^2 = \text{SSE}/(n-p-1)$. In assessing the goodness-of-fit of the regression model, we consider partitioning the total sum of squares into two components due to the regression and the residuals given by

$$\text{SST} = \text{SSR} + \text{SSE}.$$

Table 2.1: Analysis of Variance (ANOVA) table

Source	*df*	SS	MS
Total	n-1	SST	MST = SST/(n-1)
Regression	p	SSR	MSR = SSR/p
Residual	n- p -1	SSE	MSE = SSE/ n- p -1

Meanwhile, the *coefficient of determination*, denoted by R^2, is given by

$$R^2 = \frac{\text{SSR}}{\text{SST}} = 1 - \frac{\text{SSE}}{\text{SST}}$$

which describes the proportion of variation "accounted for" or "explained" by the regression. The relative sizes of the sum of squares terms indicate how good the regression is in terms of fitting the data set. If the regression is a perfect fit, all residuals are zero, and the value of R^2 is 1. But if the regression is a total failure, the sum of squares of residuals equals the total sum of squares, then no variation is accounted for by the regression, and the value of R^2 is zero. The sum of squares terms can be summarized in an Analysis of Variance table as shown in Table 2.1, where *df* refers to the degrees of freedom, while the mean squared (MS) terms are the sum of squares terms divided by the degrees of freedom. The residual mean square (MSE) is the sample estimate of the variance of the regression residuals. The population value of the error term is sometimes written as σ_e^2 while the sample estimate is given by

$$\text{MSE} = s^2 .$$

The square root of the residual mean square is called the residual mean square error (*RMSE*), or the standard error (*SE*) of the estimate given by

$$RMSE = \sqrt{s^2} = \sqrt{\text{MSE}} .$$

The ordinary residual vector is defined as

$$\boldsymbol{e} = \boldsymbol{Y} - \hat{\boldsymbol{Y}} = (1 - \boldsymbol{H})\boldsymbol{Y},$$

where $\hat{\boldsymbol{Y}}$ is the vector of the fitted values and $\boldsymbol{H} = \boldsymbol{X}(\boldsymbol{X}^{'}\boldsymbol{X})^{-1}\boldsymbol{X}^{'}$ is the hat matrix which is a symmetric and idempotent matrix. The matrix $\boldsymbol{H}$ contains the information on the influence of the response value $\boldsymbol{Y}_i$ on the corresponding fitted value $\hat{\boldsymbol{Y}}_i = \boldsymbol{H}_i^{'}\boldsymbol{Y}$, where $\boldsymbol{H}_i^{'}$ is the *i*th row of matrix $\boldsymbol{H}$. The h_{ii} is the diagonal elements of the hat matrix $\boldsymbol{H}$. Huber (1981) suggested that h_{ii} with values less than 0.2 appearing to be safe, values between 0.2 and 0.5 as being risky and values greater than 0.5, if possible, be avoided by the control of design matrix. Belsley *et al.* (1980) suggested an approximation cut-off value at 0.05 level of significant to be $(2p/n)$, where p is the number of model coefficients.

The effect of deleting one row on the estimation of parameters and their covariance, residual sum of squares and fitted values can be used to identify outliers in the data set. First, we look at the effect of outliers on the parameter estimation of $\boldsymbol{\beta}$. Let $\widehat{\boldsymbol{\beta}}_{(-i)}$ be the least square estimate of $\boldsymbol{\beta}$ when the *i*th observation is deleted. Then

$$\widehat{\boldsymbol{\beta}}_{(-i)} = (\boldsymbol{X}_{(-i)}^{'}\boldsymbol{X}_{(-i)})^{-1}\boldsymbol{X}_{(-i)}^{'}\boldsymbol{Y}_{(-i)},$$

where $\boldsymbol{X}_{(-i)}$ and $\boldsymbol{Y}_{(-i)}$ are obtained by removing the *i*th row in $\boldsymbol{X}$ and $\boldsymbol{Y}$, respectively. The change in the estimate of the parameter vector $\boldsymbol{\beta}$ when the *i*th observation is deleted is given by

$$\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_{(-i)} = \frac{(\boldsymbol{X}^{'}\boldsymbol{X})^{-1}\boldsymbol{X}_i^{'}e_i}{1-h_{ii}} ,$$

where $\boldsymbol{X}_i$ is the *i*th row of the $\boldsymbol{X}$ matrix. Cook (1977,1979) considered a statistic based on the confidence ellipsoids for investigating the contribution of each data point i to the least squares estimate of the parameter, $\boldsymbol{\beta}$, which is given by

$$\frac{(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})^{'}\boldsymbol{X}^{'}\boldsymbol{X}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})}{ps^2} \sim F_{p.n-p} .$$

In order to determine the degree of influence of the ith data point on the estimated parameter vector, $\boldsymbol{\beta}$, Cook suggested the measure of the critical nature of each data point to be

$$D_{(-i)} = \frac{(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_{(-i)})' \boldsymbol{X}' \boldsymbol{X} (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_{(-i)})}{ps^2}$$

$$= \frac{e_i^2}{ps^2} \left\{ \frac{h_{ii}}{(1 - h_{ii})^2} \right\}.$$

A large value of $D_{(-i)}$ indicates that the associated observation has a strong influence on the estimate of parameter vector $\hat{\boldsymbol{\beta}}$.

Another technique is to compare the estimated covariance matrix of $\boldsymbol{\beta}$ using all available data, $\sigma^2 (\boldsymbol{X}' \boldsymbol{X})^{-1}$, with the estimated covariance matrix when the ith observation is deleted, $\sigma^2 (\boldsymbol{X}'_{(-i)} \boldsymbol{X}_{(-i)})^{-1}$. Belsley *et al.* (1980) suggested to compare the two matrices using a determinantal ratio which is given by

$$COVRATIO_{(-i)} = \frac{\det\{s^2_{(-i)} [\boldsymbol{X}'_{(-i)} \boldsymbol{X}_{(-i)}]^{-1}\}}{\det\{s^2 (\boldsymbol{X}' \boldsymbol{X})^{-1}\}}$$

$$= \left(\frac{s_{(-i)}}{s} \right)^{2p} \frac{1}{1 - h_{ii}}.$$

A value of $COVRATIO_{(-i)}$ which is not near unity indicates that the ith observation is possibly influential. They further proposed that any data point with $|COVRATIO_{(-i)} - 1|$ close to or larger than $(3p/n)$ is identified as an outlier. In this study, we will used the idea of $COVRATIO$ to the JS circular regression models, but based on the covariance matrix of the errors due to the two observational regression-like models found in the JS circular regression framework.

2.5.2 Outliers in Circular Regression Models

Detection and investigation of outliers in circular regression is also important since outliers also have influence on the inferential results of the model. There are few published work related to the outliers in the regression for circular variables or circular regression models. The occurrence of outliers can be detected through some diagnostic tools such as P-P plot and circular correlation measures.

Fisher & Lee (1992) discussed the diagnostics checking for their proposed model. They used some diagnostic plots like the plot of residuals direction against the independent variable and the Q-Q plot when presenting the analysis of the distance and direction taken by small blue periwinkles.

Lund (1999) used the von Mises Q-Q plot and proposed the Akaike information criterion (*AICC*) statistic by assuming that the error has a von Mises distribution with concentration parameter κ when proposing the least circular distance regression (LCD) model for circular data. The model with minimum *AICC* is deemed to be the best fit. Moreover, he assessed the goodness of fit by using $A(\hat{\kappa})$ given by the equation

$$A(\hat{\kappa}) = \frac{1}{n}\sum_{i=1}^{n}\cos\left[v_i - \mu\left(\phi_i, \boldsymbol{U}_i, \hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_2\right)\right],$$

where $\hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_2$ are the regression coefficients of the LCD model, as an analogue of the residual sums of squares (*SSE*) in linear regression. Further, Lund (1999) touched on the available circular correlation measures (see Fisher, 1993; Mardia & Jupp, 2000) which could be applied to the observed and fitted values of the model. Consequently, squaring these measures gives an analogue to the coefficient of determination, R^2, of the linear regression. For a random sample $(u_1, v_1), \ldots, (u_n, v_n)$, the simplest measure is proposed by Jammalamadaka & Sarma (1988) and is given by

$$r_c = \frac{\sum_{i=1}^{n} \sin(u_i - \bar{u}) \sin(v_i - \bar{v})}{\sqrt{\sum_{i=1}^{n} \sin^2(u_i - \bar{u}) \sin^2(v_i - \bar{v})}},$$

where $\bar{u}$ and $\bar{v}$ are the sample mean directions.

So far, only Abuzaid *et al*. (2008) looked at the possibility of identifying outliers in Hussin's circular regression model via residual analysis using a new definition of circular residuals based on circular distance. The analysis is done using graphical and numerical methods. Besides, Abuzaid *et al*. (2011, 2012) also proposed the *COVRATIO* technique and residual-based statistics in detecting outliers in the same circular regression model.

However, there is no published work can be found on the problem of outliers in JS circular regression models. In this study, the investigation on the occurrence of outliers in the JS circular regression model will be carried out using two approaches; firstly, via the row deletion approach using the *COVRATIO* statistic and, secondly, using the new statistics first proposed by Abuzaid *et al*. (2013) for the Hussin's circular regression model.

2.6 Multicollinearity in Multiple Linear Regression

Multiple linear regression (MLR) model is a method used to model the linear relationship between a dependent variable and one or more independent variables. The full multiple linear regression model is given in equation (2.5). In some cases, the independent variables in the model might be near-linear dependence, leading to a problem of *multicollinearity.* The problem will cause difficulty to assess the relative importance of individual predictors from the estimated coefficients of the regression equation. In some extreme cases, we may fail to obtain the estimates as the matrix $\boldsymbol{X'X}$ is close to being singular. Perfect multicollinearity occurs when the correlation between two independent variables is equal to 1 or -1. Mansfield & Helms (1982) presented several indication of multicollinearity problem including:

1) High correlation between pairs of independent variables,
2) Statistically nonsignificant regression coefficients on important predictors, and
3) Extreme effect on the changes of sign or magnitude of regression coefficients when an independent variable is included or excluded.

2.6.1 Effect of Multicollinearity

In studying the effect of multicollinearity on regression modeling, Hoerl & Kennard (1970a, 1970b) and Swindel (1976) considered the unbiased linear estimation with minimum variance or maximum likelihood estimation when the random vector, $\boldsymbol{\varepsilon}$, is normally distributed giving $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}$ as the estimate of $\boldsymbol{\beta}$. This gives the minimum sum of squares of the residuals $\mathrm{SSE} = \left(\boldsymbol{Y} - \boldsymbol{X}\hat{\beta}\right)'\left(\boldsymbol{Y} - \boldsymbol{X}\hat{\beta}\right)$. The properties of $\hat{\boldsymbol{\beta}}$ can be found in Scheffe (1960) for the case $\boldsymbol{X}'\boldsymbol{X}$ is not nearly a unit matrix.

Hoerl & Kennard (1970a, 1970b) demonstrated the effects of the multicollinearity on the estimation of $\boldsymbol{\beta}$ by considering the variance-covariance matrix $\mathrm{COV}\left(\hat{\boldsymbol{\beta}}\right) = \sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}$ and the distance of $\boldsymbol{\beta}$ from its expected value, say, $L_1 \equiv \hat{\boldsymbol{\beta}} - \boldsymbol{\beta}$ giving

$$L_1^2 = \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)' \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) \tag{2.7}$$

with

$$E\left[L_1^2\right] = \sigma^2 \text{Trace}\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}$$

or equivalently

$$E\left[\hat{\boldsymbol{\beta}}'\hat{\boldsymbol{\beta}}\right] = \boldsymbol{\beta}'\boldsymbol{\beta} + \sigma^2 \text{Trace}\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}.$$

When the error $\boldsymbol{\varepsilon}$ is normally distributed, then

$$\text{COV}\left(L_1^2\right) = 2\sigma^4 \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-2}. \tag{2.8}$$

Using these properties, we attempt to show the uncertainty in $\hat{\boldsymbol{\beta}}$ when $\boldsymbol{X}'\boldsymbol{X}$ moves from a unit matrix to an ill-conditioned one. If the eigenvalues of $\boldsymbol{X}'\boldsymbol{X}$ are denoted by

$$\lambda_{\max} = \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{\boldsymbol{p}} = \lambda_{\min} > 0,$$

then

$$E\left[L_1^2\right] = \sigma^2 \sum_{i=1}^{p} \frac{1}{\lambda_i} \tag{2.9}$$

and the variance when the error is normally distributed is given by

$$\text{VAR}\left[L_1^2\right] = 2\sigma^4 \sum_{i=1}^{\boldsymbol{p}} \left(\frac{1}{\lambda_{\boldsymbol{i}}}\right)^2. \tag{2.10}$$

Note that when the matrix $\boldsymbol{X}'\boldsymbol{X}$ is ill-conditioned due to multicollinearity, then some of the λ_j will be small. Hence, from (2.9), the least squares estimates $\hat{\boldsymbol{\beta}}$ is farther away from true parameter $\boldsymbol{\beta}$ and, from (2.10), the variances of the least squares estimator of the regression coefficient have larger values. Hence, proper handling of multicollinearity problem is greatly needed.

2.6.2 Multicollinearity Diagnostics

Farrar & Glauber (1967) published a now well-known article on the problem of multicollinearity in regression analysis. Their article presents the possibility of making misleading inferences by only examining the simple bivariate correlations among the variables in the presence of multicollinearity. Therefore, the multiple correlations of each variable on all of the others are needed to be examined in order to assess the extent of collinearity in the data. They also showed that if the variables are found to be orthogonal, then there is no multicollinearity problem in the data. But if the variables are not orthogonal, then, the multicollinearity may be present in the data. Later, Lemieux (1978) extended the work and developed an alternative method of computing the correlation using algorithm that is available from common multiple regression algorithms.

Some authors have suggested a formal detection-tolerance or the Variance Inflation Factor (*VIF*) for measuring multicollinearity in the multiple linear regression models. “Variance Inflation” refers to the effect of multicollinearity on the variance of estimated regression coefficients. Multicollinearity depends not just on the bivariate correlations between pairs of predictors, but on the multivariate predictability of any one predictor from the other predictors. Accordingly, the *VIF* is obtained based on the value of the coefficient of determination R_i^2 by regressing X_i on the other independent variables and is given by

$$VIF_i = \frac{1}{1 - R_i^2}, \tag{2.11}$$

where the VIF_i is associated with the *i*th predictor, X_i. Note that if the *i*th predictor is independent of the other predictors, the *VIF* will take value one, while if the *i*th predictor is almost perfectly predicted from the other predictors, the *VIF* approaches infinity. In this case, the variance of the estimated regression coefficients is unbounded. A good model should have R_i^2 closed to 1 (see Lemieux, 1978; Mansfield & Helms, 1982).

Mansfield & Helms (1982) also presented the need of examining latent roots and latent vectors of the correlation matrix and the *VIF*. When the data is orthogonal then all the *VIF* equals unity. Otherwise, the *VIF* could be a good indicator that provides the user with a measure of how many times larger the $\mathrm{VAR}\left(\hat{\beta}_j\right)$ will be for multicollinear data than for orthogonal data. If the *VIF*'s are not unusually larger than 1.0, then the multicollinearity problem is not a problem.

Meanwhile, Haan (2002) noted that some researchers use a *VIF* of 5 and others use a *VIF* of 10 as critical thresholds. These values correspond, respectively, to R_i^2 values of 0.80 and 0.90. Some compute an average *VIF* for all predictors and declare that multicollinearity problem exists when the average is “considerably” larger than one. The *VIF* is closely related to a statistic called the tolerance, which is $\frac{1}{VIF}$. Some statistics packages report the *VIF* and some report the tolerance. Once the problem has been identified, we may then use the ridge regression model to study the relationship between the variables and is described in the following section.

2.7 Ridge Regression

Hoerl (1962) and Hoerl & Kennard (1968) suggested an approach to control the inflation and general instability caused by multicollinearity associated with the least squares estimates by introducing a constant k in the LS estimates of a multiple linear regression model as follows:

$$\begin{aligned}\hat{\boldsymbol{\beta}}^* &= \left(\boldsymbol{X'X} + k\boldsymbol{I}\right)^{-1}\boldsymbol{X'Y}^{(j)};\ k \geq 0 \\ &= \boldsymbol{WX'Y}^{(j)}, \qquad (2.12)\end{aligned}$$

where $\boldsymbol{W} = \left(\boldsymbol{X'X} + k\boldsymbol{I}\right)^{-1}$. The estimation and analysis built for the equation (2.12) is labelled as *ridge regression analysis*. We can show that the relationship between the LS and ridge estimates is given by

$$\hat{\boldsymbol{\beta}}* = \left[\boldsymbol{I}_p + k\left(\boldsymbol{X'X}\right)^{-1}\right]^{-1}\hat{\boldsymbol{\beta}}$$

$$= \boldsymbol{Z}\hat{\boldsymbol{\beta}} ,$$

where $\boldsymbol{Z} = \left[\boldsymbol{I}_p + k\left(\boldsymbol{X'X}\right)^{-1}\right]^{-1}$. The resulting residual sum of squares is

$$\phi^*(k) = \left(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}^*\right)'\left(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}^*\right)$$

which can also be written as

$$\phi^*(k) = \boldsymbol{Y'Y} - \left(\hat{\boldsymbol{\beta}}^*\right)'\boldsymbol{X'Y} - k\left(\hat{\boldsymbol{\beta}}^*\right)'\left(\hat{\boldsymbol{\beta}}^*\right) \ . \qquad (2.13)$$

The expression (2.13) shows that $\phi^*(k)$ is the total sum of squares minus the regression sum of squares for $\hat{\boldsymbol{\beta}}^*$ with a modification depending upon the squared length of $\hat{\boldsymbol{\beta}}^*$.

Hoerl & Kennard (1970a and 1970b) mentioned that ridge regression for multiple linear regression model has two important aspects to be considered; the ridge trace and the determination of a value of *k* that gives a stable estimate of $\boldsymbol{\beta}$. The ridge trace is a two-dimensional plot of the $\hat{\boldsymbol{\beta}}^*(k)$ and the residual sum of squares, $\phi^*(k)$, for a number of values of *k* in the interval [0,1]. They have suggested some guideline to choose an appropriate value of *k* as follow:

(i) The system should stabilize at a certain value of *k* and should follow the general characteristics of an orthogonal system.

(ii) Coefficients should not have unreasonable absolute values with respect to the factors for which they represent.

(iii) Least squares coefficients with incorrect signs will eventually have changed to the proper sign when *k* get closer to one.

(iv) The SSE should not have inflated to an unreasonable value. It will not be large relative to the minimum residual sum of squares or too large to be a reasonable variance for the process generating the data.

Meanwhile, Swindel (1976) proposed modified ridge regression estimators based on prior information and Liu (1993) proposed a new estimator that combines the Stein estimator with the ordinary ridge regression estimator. Then, Li and Yang (2012) proposed a modified Liu estimator based on prior information and the Liu estimator. Here, we will focus on applying the ridge regression model proposed by Hoerl & Kennard (1970) to the generalized JS circular regression models with more than one explanatory variables.

2.8 Functional Model

The functional model is part of the general class of error-in-variables model (EIVM), in which the underlying variables are deterministic (or fixed) where EIVM refers to the case when both variables are subject to errors (see Hussin, 1998). The fitting of a linear relationship with errors in the continuous linear variables or EIVM had been explored, see Madansky (1959), Adcock (1877, 1878), Moran (1971), Kendall & Stuart (1973) and Fuller (1987). Adcock (1877, 1878) investigated the estimation properties under some realistic assumptions in ordinary linear regression models and obtained the least squares solution for the slope parameters by assuming that both variables have equal error variances. Since then, several authors have worked on the problem of estimating the parameters in the setup. Besides, Kendall (1951, 1952) formally made a distinction between functional and structural relationship between the two variables.

In the problem of linear functional relationship model, assume that we have a sample $\{(X_{1,}Y_1), ..., (X_{n,}Y_n)\}$, where the $\boldsymbol{X_i}$'s are independent and identically distributed as a random function X and the $\boldsymbol{Y_i}$'s are generated by the following regression model;

$$x_i = X_i + e_{1i} \quad \text{and} \quad y_i = Y_i + e_{2i}, \text{ where}$$

$$Y_i = a + bX_i, \quad \text{for} \quad i = 1, 2, ..., n, \tag{2.14}$$

where a is a constant, b is the slope function and e_{1i} , e_{2i} are the errors following Normal distribution. Various estimation methods of model (2.14) have been developed in the past, see Hussin (1998), Caries & Wyatt (2002) and Cai & Hall (2006).

As for circular functional relationship model, Hussin *et al*. (1997) proposed a model which follows exactly the form of model (2.14), but the variables are now circular. Later, Caries & Wyatt (2002) considered the same model but the parameter β is fixed as unity. Suppose u_j and v_j are the observed values of the circular variables U and V respectively, where $0 \le u_j, v_j < 2\pi$ for j = 1, 2, ..., n. For any fixed U_j, they assume that the observations u_j and v_j are measured with errors δ_j and ε_j, respectively. Thus, the circular functional relationship model is given by

$$u_j = U_j + \delta_{ij} \quad \text{and} \quad v_j = V_j + \varepsilon_j, \text{ where}$$

$$Y_j = a + bX_j \quad (\text{mod } 2\pi), \text{ for} \quad j = 1, 2, ..., n, \tag{2.15}$$

where a is a constant parameter and b is a real value close to unity. The complex linear functional relationship model for circular variables is given by

$$\left(\cos x_j + i \sin x_j\right) = \left(\cos X_j + i \sin X_j\right) + \delta_j$$

and

$$\left(\cos y_j + i \sin y_j\right) = \left(\cos Y_j + i \sin Y_j\right) + \varepsilon_j, \tag{2.16}$$

where δ_j and ε_j are independently distributed with bivariate complex Gaussian distribution according to Goodman (1963), with zero mean and variance σ_1^2 and σ_2^2, respectively, and $0 \le x_j, y_j < 2\pi$ for j = 1, 2, ..., n denoted by a series of a complex number. By using equations (2.15) and (2.16), Hussin (1998) provided the maximum likelihood estimators for the model and then applied on the analysis of wind direction data measured by two different methods: the anchored wave buoy and HF radar system, with the objective is to compare the calibration of both measurement techniques. Recently, Hassan *et al*. (2010) considered some issues with regard to model (2.16) including a new method of estimating the concentration parameter of errors. In this study, we will extend the idea to accommodate the JS circular regression models in the circular functional relationship set-up.

2.9 Summary

We have reviewed the theory on circular statistics, circular regression model and the detection of outliers in linear and circular cases in this chapter. We have also looked at the problem of multicollinearity in multiple linear regression and the theory on linear functional relationship model and circular functional relationship model. We intend to expand these theories to the case involving JS circular regression models in the subsequent chapters.

CHAPTER THREE

JS CIRCULAR REGRESSION MODEL

3.1 Introduction

Regression analysis is a statistical technique used for investigating the relationship between variables. Discussion on the development of circular regression models has begun dated back to Gould (1969) in predicting the mean direction of a circular response variable Θ from a vector of linear covariates $\boldsymbol{X} = (x_1, \ldots, x_k)$. Mardia (1972) extended the model by assuming each of the response variables θ_i, $i = 1, \ldots, n$, to be independently distributed from von Mises distribution with mean direction μ_i and unknown concentration parameter κ.

For the case when both response and explanatory variables are circular, say *u* and *v* respectively, a few circular regression models have been proposed using different approach. The earliest model is proposed by Laycock (1975) who expressed the model as a multiple linear regression model with complex entries. Meanwhile, Downs & Mardia (2002) proposed a model based on a one-to-one mapping between the independent angle u and the mean of dependent angle v such that the locus of the points (u,v) is a continuous closed curve winding once around a toroidal surface. Other models include those proposed by Hussin *et al.* (2004) and Kato *et al*. (2008). Another interesting model of our interest is proposed by Jammalamadaka & Sarma (1993) who utilized the theory on the conditional expectation of the vector e^{iv} given u and again is referred as the *JS circular regression model*. The properties of the models for single explanatory variable are presented next.

3.2 JS Circular Regression Model

Jammalamadaka & Sarma (1993) proposed a regression model for two circular random variables ***U*** and ***V***. To predict *v* for a given *u*, consider the conditional expectation of the vector e^{iv} given *u* such that

$$\begin{aligned} E(e^{iv} \mid u) &= \rho(u)\, e^{i\mu(u)} \\ &= \rho(u)\cos\mu(u) + i\rho(u)\sin\mu(u) \\ &= g_1(u) + i\, g_2(u), \end{aligned} \tag{3.1}$$

where $e^{iv} = \cos v + i \sin v$, $\mu(u)$ represents the conditional mean direction of *v* given *u* and $\rho(u)$ represents the conditional concentration parameter. Equivalently, we may write

$$\begin{aligned} E(\cos v \mid u) &= g_1(u) \\ E(\sin v \mid u) &= g_2(u). \end{aligned} \tag{3.2}$$

Then, we may predict *v* such that

$$\mu(u) = \hat{v} = \arctan^* \frac{g_2(u)}{g_1(u)} = \begin{cases} \tan^{-1} \dfrac{g_2(u)}{g_1(u)} & \text{if } g_1(u) \geq 0 \\ \pi + \tan^{-1} \dfrac{g_2(u)}{g_1(u)} & \text{if } g_1(u) \leq 0 \\ \text{undefined} & \text{if } g_1(u) = g_2(u) = 0. \end{cases} \tag{3.3}$$

Due to the difficulty in estimating $g_1(u)$ and $g_2(u)$, they are approximated using suitable functions by taking into account the fact that both are periodic function with period 2π. Jammalamadaka & Sarma (1993) considered the trigonometric polynomials of function of one variable to approximate $g_1(u)$ and $g_2(u)$, see Kufner & Kadlec (1971). For a suitable degree *m*, we have

$$g_1(u) \approx \sum_{k=0}^{m} (A_{\boldsymbol{k}} \cos ku + B_{\boldsymbol{k}} \sin ku)$$
$$g_2(u) \approx \sum_{k=0}^{m} (C_{\boldsymbol{k}} \cos ku + D_{\boldsymbol{k}} \sin ku). \tag{3.4}$$

Consequently, we have the following two observational regression-like models:

$$\cos v = \sum_{k=0}^{m} (A_k \cos ku + B_k \sin ku) + \varepsilon_1$$
$$\sin v = \sum_{k=0}^{m} (C_k \cos ku + D_k \sin ku) + \varepsilon_2, \tag{3.5}$$

where $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2)$ is the vector of random errors following the bivariate normal distribution with mean vector $\boldsymbol{0}$ and unknown dispersion matrix $\boldsymbol{\Sigma}$. The parameters A_k, B_k, C_k, and D_k, $k = 0,1,\ldots,m$, the standard errors as well as the matrix $\boldsymbol{\Sigma}$ can then be estimated.

3.3 Estimation of JS Circular Regression Parameters

We consider two methods of estimating the parameters of the JS circular regression model, namely, the least squares and maximum likelihood estimation method.

3.3.1 Least Squares Method

Jammalamadaka & Sarma (1993) had described the estimation method for the JS circular regression model based on the generalized least squares (LS) approach. Let $(u_1, v_1), (u_2, v_2), \ldots, (u_n, v_n)$ be a random circular sample of size n. From (3.5), we now have the observational regression-like equations given by

$$V_{1j} = \cos v_j = \sum_{k=0}^{m} (A_k \cos ku_j + B_k \sin ku_j) + \varepsilon_{1j}$$
$$V_{2j} = \sin v_j = \sum_{k=0}^{m} (C_k \cos ku_j + D_k \sin ku_j) + \varepsilon_{2j} \tag{3.6}$$

for $j = 1, \dots, n$. Assume that $B_0=D_0=0$ to ensure identifiability. Therefore, the observational equations (3.6) can be summarized as

$$\begin{aligned} \boldsymbol{V}^{(1)} &= \left(V_{11},\dots,V_{1n}\right)' \\ \boldsymbol{V}^{(2)} &= \left(V_{21},\dots,V_{2n}\right)' \\ \boldsymbol{\varepsilon}^{(1)} &= \left(\varepsilon_{11},\dots,\varepsilon_{1n}\right)' \\ \boldsymbol{\varepsilon}^{(2)} &= \left(\varepsilon_{21},\dots,\varepsilon_{2n}\right)' \end{aligned} \tag{3.7}$$

$$\underset{n\times(2m+1)}{\boldsymbol{U}} = \begin{bmatrix} 1 & \cos u_1 & \cdots & \cos mu_1 & \sin u_1 & \cdots & \sin mu_1 \\ 1 & \cos u_2 & \cdots & \cos mu_2 & \sin u_2 & \cdots & \sin mu_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & \cos u_n & \cdots & \cos mu_n & \sin u_n & \cdots & \sin mu_n \end{bmatrix} \tag{3.8}$$

and

$$\begin{aligned} \boldsymbol{\lambda}^{(1)} &= \left(A_0, A_1, \dots, A_m, B_1, \dots, B_m\right)' \\ \boldsymbol{\lambda}^{(2)} &= \left(C_0, C_1, \dots, C_m, D_1, \dots, D_m\right)'. \end{aligned} \tag{3.9}$$

The observational equations (3.6) can be written in matrix form

$$\begin{aligned} \boldsymbol{V}^{(1)} &= \boldsymbol{U}\boldsymbol{\lambda}^{(1)} + \boldsymbol{\varepsilon}^{(1)} \\ \boldsymbol{V}^{(2)} &= \boldsymbol{U}\boldsymbol{\lambda}^{(2)} + \boldsymbol{\varepsilon}^{(2)}. \end{aligned} \tag{3.10}$$

The least squares estimates turn out to be

$$\begin{aligned} \hat{\boldsymbol{\lambda}}^{(1)} &= \left(\boldsymbol{U}'\boldsymbol{U}\right)^{-1}\boldsymbol{U}'\boldsymbol{V}^{(1)} \\ \hat{\boldsymbol{\lambda}}^{(2)} &= \left(\boldsymbol{U}'\boldsymbol{U}\right)^{-1}\boldsymbol{U}'\boldsymbol{V}^{(2)}. \end{aligned} \tag{3.11}$$

Then, the covariance matrix $\boldsymbol{\Sigma}$ resulting from equation (3.5) can be estimated using the least squares theory. Let

$$R_0(p,q) = \boldsymbol{V}^{(p)'}\boldsymbol{V}^{(q)} - \boldsymbol{V}^{(p)'}\boldsymbol{U}\left(\boldsymbol{U}'\boldsymbol{U}\right)^{-1}\boldsymbol{U}'\boldsymbol{V}^{(q)}, \tag{3.12}$$

where $R_0 = \left(R_0(p,q)\right)_{p,q=1,2}$, then

$$\hat{\boldsymbol{\Sigma}} = \left[n - 2(2m+1)\right]^{-1} R_0 \tag{3.13}$$

is an unbiased estimate of $\boldsymbol{\Sigma}$, and hence the standard errors of the estimators can then be found. The estimated covariance matrix $\hat{\boldsymbol{\Sigma}}$ will be used in the Chapter 4 to identify possible outliers in the data. We can also estimate $\mu(u)$ by using equation (3.3) and ρ using the following equation:

$$\rho(u)=\sqrt{\frac{1}{n}\sum_{j=1}^{n}\rho^2(u_j)}=\sqrt{\frac{1}{n}\sum_{j=1}^{n}\left[\, g_1^2(u_j)+g_2^2(u_j)\,\right]}, \tag{3.14}$$

where $0\le\rho(u)\le 1$.

3.3.2 Maximum Likelihood Estimation Method

An alternative estimation method is the maximum likelihood estimation (MLE) method. This estimation method will be used in the development of the functional form of the JS regression models in Chapter 7. For simplicity, we consider the case when m=1. Hence, from equation (3.6), we expand the error term and obtain

$$\varepsilon_j=\left(\cos\ v_j-A_0-A_1\cos u_j-B_1\sin u_j\right)$$
$$+\,i\,\left(\sin v_j\ -C_0-C_1\cos\ u_j-D_1\sin u_j\right)$$

and

$$\left|\varepsilon_j\right|^2=\left(\cos\ v_j-A_0-A_1\cos u_j-B_1\sin u_j\right)^2$$
$$+\left(\sin v_j\ -C_0-C_1\cos\ u_j-D_1\sin u_j\right)^2.$$

Therefore, the log-likelihood function is given by

$$\log L\left(A_0,A_1,B_1,C_0,C_1,D_1,\sigma^2;u_j,v_j\right)=$$
$$-\,n\log\left(\pi\sigma^2\right)-\frac{1}{\sigma^2}\sum_j\left(\cos\ v_j-A_0-A_1\cos u_j-B_1\sin u_j\right)^2 \tag{3.15}$$
$$-\frac{1}{\sigma^2}\sum_j\left(\sin v_j\ -C_0-C_1\cos\ u_j-D_1\sin u_j\right)^2.$$

The function log L is then differentiated with respect to each parameters and equated to zero. Hence, we obtain the following estimates of the parameters:

$$\hat{A}_0 = \frac{1}{n}\sum_j \left(\cos\ v_j - \hat{A}_1 \cos u_j - \hat{B}_1 \sin u_j\right) \tag{3.16}$$

$$\hat{A}_1 = \frac{\sum_j \left(\cos\ v_j - \hat{A}_0 - \hat{B}_1 \sin u_j\right)}{\sum_j \left(\cos u_j\right)} \tag{3.17}$$

$$\hat{B}_1 = \frac{\sum_j \left(\cos\ v_j - \hat{A}_0 - \hat{A}_1 \cos\ u_j\right)}{\sum_j \left(\sin u_j\right)} \tag{3.18}$$

$$\hat{C}_0 = \frac{1}{n}\sum_j \left(\sin v_j - \hat{C}_1 \cos\ u_j - \hat{D}_1 \sin u_j\right) \tag{3.19}$$

$$\hat{C}_1 = \frac{\sum_j \left(\sin v_j - \hat{C}_0 - \hat{D}_1 \sin u_j\right)}{\sum_j \left(\cos u_j\right)} \tag{3.20}$$

$$\hat{D}_1 = \frac{\sum_j \left(\sin v_j - \hat{C}_0 - \hat{C}_1 \cos\ u_j\right)}{\sum_j \left(\sin u_j\right)}. \tag{3.21}$$

As for the estimation $\hat{\sigma}^2$ for σ^2, we first let $\sigma^2 = w$. Using the log-likelihood function given by (3.15), we differentiate $\log L$ with respect to *w* and then set to zero giving

$$\hat{w} = \frac{1}{n}\left[\begin{array}{l}\sum_j \left(\cos\ v_j - \hat{A}_0 - \hat{A}_1 \cos u_j - \hat{B}_1 \sin u_j\right)^2 \\ \quad + \sum_j \left(\sin v_j\ - \hat{C}_0 - \hat{C}_1 \cos\ u_j - \hat{D}_1 \sin u_j\right)^2\end{array}\right]. \tag{3.22}$$

Both methods of LS and MLE should give similar estimates of the parameters $A_0, A_1, B_1, C_0, C_1, D_1$, under the assumption that the error terms are normally distributed. We now consider the effect of outlier on the LS estimates only in the next section.

3.4 Effect of Outliers on LS Estimation Method

In statistical analysis, the existence of outlying values from the others in the data set should raise some concern. The study on the existence of outliers in linear data sets and linear regression has been carried out extensively in the past; see for examples, Beckman & Cook (1983), Barnett & Lewis (1978), Belsley *et al*. (1980), Montgomery & Peck (1992) and Barnett & Lewis (1994). The effect of outliers on the parameter estimation and data modelling are known to be severe. Here, it is useful to be able to identify outliers first in the data before taking the next course of action in treating the presence of outliers in the data set. Next, we first investigate the robustness of JS circular regression model by introducing outliers in the data set.

3.4.1 Simulation Procedure

A simulation study was carried out to investigate the effect of outlier on the parameter estimates of JS circular regression models. For simplicity, we consider the case when m=1. Hence, we have the following set of parameters to be estimated:

$$\boldsymbol{\lambda} = \left(\boldsymbol{\lambda}^{(1)}, \boldsymbol{\lambda}^{(2)}\right) = \left(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6\right)$$

$$= \left(A_0, A_1, B_1, C_0, C_1, D_1\right). \tag{3.23}$$

We consider the set of uncorrelated random errors $(\varepsilon_1, \varepsilon_2)$ from the bivariate Normal distribution with mean vector $\boldsymbol{0}$ and variances (σ_1, σ_2) to be (0.03,0.03). For simplicity, we set the true values of A_0 and C_0 of the JS model to be zero, while A_1, B_1, C_1 and D_1 are obtained by using the standard additive trigonometric polynomial equations $\cos(a+u)$ and $\sin(a+u)$. For example, $\cos(2+u) = -0.4161 \cos u - 0.9093 \sin u$ and $\sin(2+u) = 0.9093 \cos u - 0.4161 \sin u$ when a = 2. Then by comparing with equation (3.5), the true values of A_1, B_1, C_1 and D_1 are -0.4161, -0.9093, 0.9093 and -0.4161 respectively. Similarly, we can also get different sets of true values by choosing different values of a. Here, we consider the values of $a = -6, -2, 2$ and 6. We then introduce outliers into the data such that the percentages of contamination used is c%=10%, 20%, 30%, 40% and 50% from the sample size n=100. The full steps of the simulation are described below:

(i) Generate fixed variable $\boldsymbol{U}$ of size n from $VM(\pi, 2)$.

(ii) Generate $\boldsymbol{\varepsilon}_1$ and $\boldsymbol{\varepsilon}_2$ of size n from $\boldsymbol{N}\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}0.03 & 0\\0 & 0.03\end{pmatrix}\right)$. For a fixed a, obtain the true values of A_1, B_1, C_1 and D_1. We let the true values of A_0 and C_0 being zero. Then, we calculate V_{1j} and V_{2j}, $j = 1, \ldots, n$ using equation (3.6).

(iii) Obtain the variable $v_j = \arctan\left(\dfrac{V_{2j}}{V_{1j}}\right)$, $j = 1, \ldots, n$.

(iv) For uncontaminated model, the generated circular data (u_j, v_j) above is fitted to the JS circular regression model to give the parameter estimates $\hat{\lambda} = \left(\hat{A}_0, \hat{A}_1, \hat{B}_1, \hat{C}_0, \hat{C}_1, \hat{D}_1\right)$.

(v) For c% contaminated data, we replace the last $c \times n/100$ observations v in (iv) by the newly generated values v^* such that the errors $\boldsymbol{\varepsilon}_1^*$ and $\boldsymbol{\varepsilon}_2^*$ are now generated from $N\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}0.1 & 0\\0 & 0.1\end{pmatrix}\right)$. Then, the generated contaminated circular data are fitted using the JS circular regression model to give the parameter estimates $\hat{\boldsymbol{\lambda}}^* = \left(\hat{A}_0^*, \hat{A}_1^*, \hat{B}_1^*, \hat{C}_0^*, \hat{C}_1^*, \hat{D}_1^*\right)$ using equations (3.16) - (3.21).

(vi) Finally, the steps (i) – (v) above are repeated for *simu*=1000 times. For each parameter estimates $(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6) = (A_0, A_1, B_1, C_0, C_1, D_1)$, the estimated mean, bias, standard error (*SE*) and root mean squared error (*RMSE*) for outlier-free data set are calculated using the following formulas:

- Mean of the estimates is given by

$$\overline{\lambda}_i = \frac{\sum_{j=1}^{simu} \hat{\lambda}_{i,j}}{simu}, \; i=1, 2, ..., 6. \tag{3.24}$$

- Standard error (*SE*) of the estimates is given by

$$SE\left(\hat{\lambda}_i\right) = \sqrt{\frac{\sum_{j=1}^{simu} \left(\hat{\lambda}_{i,j} - \overline{\lambda}_i\right)^2}{simu}}, \; i=1, 2, ..., 6. \tag{3.25}$$

- Bias of the estimates is given by

$$\text{bias}\left(\hat{\lambda}_i\right) = \overline{\lambda}_i - \lambda_i, \; i=1, 2, ..., 6. \tag{3.26}$$

- Residual mean square error (*RMSE*) of the estimates is given by

$$RMSE\left(\hat{\lambda}_i\right) = \sqrt{\frac{\sum_{j=1}^{simu} \left(\hat{\lambda}_{i,j} - \lambda_i\right)^2}{simu}}, \; i=1, 2, ..., 6. \tag{3.27}$$

Similarly, we calculate the mean, *SE*, bias and *RMSE* of the estimates for contaminated data using equation (3.24) - (3.27) by replacing $\hat{\lambda}_i$ with $\hat{\lambda}_i^*$ for *simu*=1000 times.

3.4.2 Discussion

The results are tabulated in Tables 3.1 - 3.4 for each value of *a* considered. Several results are observed as follows:

1. For outlier-free data set, the estimated mean for all parameter estimates are consistently close to the true values. When the data are contaminated with outliers, the bias is generally larger than the uncontaminated data set.

2. When the percentage of contamination increases from 10% to 50%, the value of the bias increases.
3. The standard errors (*SE*) for all parameters estimates are generally small for uncontaminated data but gets larger as the percentages of contamination increase.
4. The value for root mean squares error (*RMSE*) of each parameter estimates increase when the percentages of contamination increase.

By looking at the results for uncontaminated data, the least squares estimation method performs well in estimating the parameters of the JS circular regression model. However, the method is affected by the presence of outliers in the data. The effect is worsen with the present of higher percentage of contaminated observations in the data.

Table 3.1: Parameter estimates for contaminated and uncontaminated data when a = -6

Estimates	True value	**Uncontaminated**				**Contaminated**							
		0%				**10%**				**20%**			
		mean	*SE*	bias	*RMSE*	mean	*SE*	bias	*RMSE*	mean	*SE*	bias	*RMSE*
$\hat{A}_0$	0.0000	-0.0001	0.0053	-0.0001	0.0022	-0.0577	0.0699	-0.0577	1.8241	-0.1030	0.0929	-0.1030	3.2571
$\hat{A}_1$	0.9602	0.9597	0.0059	-0.0005	0.0161	0.8761	0.0835	-0.0841	2.6600	0.8083	0.1096	-0.1519	4.8024
$\hat{B}_1$	-0.2794	-0.2798	0.0042	-0.0004	0.0134	-0.2563	0.0337	0.0231	0.7298	-0.2335	0.0451	0.0459	1.4512
$\hat{C}_0$	0.0000	-0.0006	0.0041	-0.0006	0.0193	-0.0184	0.0684	-0.0184	0.5810	-0.0339	0.0938	-0.0339	1.0721
$\hat{C}_1$	0.2794	0.2780	0.0059	-0.0015	0.0461	0.2552	0.0765	-0.0242	0.7643	0.2347	0.1063	-0.0447	1.4150
$\hat{D}_1$	0.9602	0.9610	0.0036	0.0008	0.0254	0.8763	0.0460	-0.0839	2.6521	0.8027	0.0603	-0.1574	4.9783

Estimates	True value	**Contaminated**											
		30%				**40%**				**50%**			
		mean	*SE*	bias	*RMSE*	mean	*SE*	bias	*RMSE*	mean	*SE*	bias	*RMSE*
$\hat{A}_0$	0.0000	-0.1576	0.1106	-0.1576	4.9831	-0.2189	0.1291	-0.2189	6.9210	-0.2579	0.1360	-0.2579	8.1558
$\hat{A}_1$	0.9602	0.7276	0.1284	-0.2326	7.3544	0.6389	0.1463	-0.3213	10.1601	0.5809	0.1552	-0.3793	11.9945
$\hat{B}_1$	-0.2794	-0.2105	0.0515	0.0689	2.1784	-0.1899	0.0583	0.0896	2.8320	-0.1649	0.0664	0.1145	3.6209
$\hat{C}_0$	0.0000	-0.0524	0.1129	-0.0524	1.6580	-0.0705	0.1244	-0.0705	2.2309	-0.0799	0.1421	-0.0799	2.5258
$\hat{C}_1$	0.2794	0.2105	0.1228	-0.0689	2.1796	0.1901	0.1363	-0.0893	2.8239	0.1744	0.1487	-0.1050	3.3214
$\hat{D}_1$	0.9602	0.7234	0.0727	-0.2368	7.4888	0.6483	0.0836	-0.3119	9.8632	0.5746	0.0913	-0.3856	12.1941

Table 3.2: Parameter estimates for contaminated and uncontaminated data when *a* = -2

Estimates	True value	Uncontaminated				Contaminated							
		0%				10%				20%			
		mean	*SE*	bias	*RMSE*	mean	*SE*	bias	*RMSE*	mean	*SE*	bias	*RMSE*
$\hat{A}_0$	0.0000	-0.0001	0.0043	-0.0001	0.0021	0.0236	0.0664	0.0236	0.7475	0.0530	0.0962	0.0530	1.6762
$\hat{A}_1$	-0.4161	-0.4161	0.0058	0.0000	0.0013	-0.3846	0.0762	0.0315	0.9964	-0.3462	0.1044	0.0700	2.2131
$\hat{B}_1$	0.9093	0.9089	0.0037	-0.0004	0.0127	0.8345	0.0446	-0.0748	2.3650	0.7565	0.0624	-0.1528	4.8311
$\hat{C}_0$	0.0000	0.0001	0.0054	0.0001	0.0020	0.0548	0.0704	0.0548	1.7333	0.0991	0.0923	0.0991	3.1336
$\hat{C}_1$	-0.9093	-0.9088	0.0062	0.0005	0.0155	-0.8290	0.0825	0.0803	2.5396	-0.7648	0.1068	0.1445	4.5702
$\hat{D}_1$	-0.4161	-0.4159	0.0043	0.0003	0.0092	-0.3803	0.0356	0.0358	1.1332	-0.3496	0.0466	0.0665	2.1040

Estimates	True value	Contaminated											
		30%				40%				50%			
		mean	*SE*	bias	*RMSE*	mean	*SE*	bias	*RMSE*	mean	*SE*	bias	*RMSE*
$\hat{A}_0$	0.0000	0.0702	0.1082	0.0702	2.2191	0.0988	0.1219	0.0988	3.1250	0.1299	0.1336	0.1299	4.1063
$\hat{A}_1$	-0.4161	-0.3227	0.1165	0.0934	2.9536	-0.2828	0.1345	0.1334	4.2175	-0.2447	0.1406	0.1715	5.4225
$\hat{B}_1$	0.9093	0.6874	0.0730	-0.2219	7.0172	0.6125	0.0807	-0.2968	9.3867	0.5372	0.0909	-0.3721	11.7663
$\hat{C}_0$	0.0000	0.1504	0.1079	0.1504	4.7573	0.2007	0.1206	0.2007	6.3475	0.2581	0.1417	0.2581	8.1624
$\hat{C}_1$	-0.9093	-0.6871	0.1228	0.2222	7.0254	-0.6158	0.1345	0.2935	9.2821	-0.5339	0.1554	0.3754	11.8705
$\hat{D}_1$	-0.4161	-0.3127	0.0577	0.1035	3.2714	-0.2782	0.0613	0.1380	4.3629	-0.2474	0.0698	0.1688	5.3368

Table 3.3: Parameter estimates for contaminated and uncontaminated data when *a* = 2

Estimates	True value	Uncontaminated				Contaminated							
		0%				10%				20%			
		mean	*SE*	bias	*RMSE*	mean	*SE*	bias	*RMSE*	mean	*SE*	bias	*RMSE*
$\hat{A}_0$	0.0000	0.0000	0.0013	0.0001	0.0001	0.0274	0.0646	0.0274	0.8663	0.0511	0.0905	0.0511	1.6173
$\hat{A}_1$	-0.4161	-0.4161	0.0018	-0.0004	-0.0004	-0.3810	0.0730	0.0352	1.1123	-0.3501	0.0995	0.0661	2.0901
$\hat{B}_1$	0.9093	0.9093	0.0012	-0.0001	-0.0001	-0.8307	0.0434	0.0786	2.4841	-0.7593	0.0586	0.1500	4.7428
$\hat{C}_0$	0.0000	0.0000	0.0017	-0.0003	-0.0003	-0.0570	0.0658	-0.0570	1.8010	-0.0972	0.0928	-0.0972	3.0731
$\hat{C}_1$	-0.9093	-0.9093	0.0020	-0.0010	-0.0010	0.8269	0.0771	-0.0824	2.6062	0.7643	0.1071	-0.1450	4.5849
$\hat{D}_1$	-0.4161	-0.4161	0.0014	-0.0010	-0.0010	-0.3790	0.0339	0.0372	1.1761	-0.3470	0.0475	0.0692	2.1874

Estimates	True value	Contaminated											
		30%				40%				50%			
		mean	*SE*	bias	*RMSE*	mean	*SE*	bias	*RMSE*	mean	*SE*	bias	*RMSE*
$\hat{A}_0$	0.0000	0.0705	0.1122	0.0705	2.2288	0.0977	0.1246	0.0977	3.0883	0.1246	0.1414	0.1246	3.9392
$\hat{A}_1$	-0.4161	-0.3215	0.1200	0.0947	2.9932	-0.2837	0.1335	0.1325	4.1897	-0.2479	0.1467	0.1682	5.3201
$\hat{B}_1$	0.9093	-0.6843	0.0721	0.2250	7.1142	-0.6128	0.0800	0.2965	9.3766	-0.5385	0.0907	0.3708	11.7258
$\hat{C}_0$	0.0000	-0.1529	0.1105	-0.1529	4.8349	-0.2010	0.1255	-0.2010	6.3559	-0.2541	0.1386	-0.2541	8.0346
$\hat{C}_1$	-0.9093	0.6865	0.1257	-0.2228	7.0464	0.6154	0.1407	-0.2939	9.2934	0.5379	0.1492	-0.3714	11.7441
$\hat{D}_1$	-0.4161	-0.3133	0.0575	0.1028	3.2517	-0.2783	0.0649	0.1379	4.3601	-0.2440	0.0709	0.1721	5.4436

Table 3.4: Parameter estimates for contaminated of outliers when a = 6

Estimates	True value	Uncontaminated				Contaminated							
		0%				10%				20%			
		mean	*SE*	bias	*RMSE*	mean	*SE*	bias	*RMSE*	mean	*SE*	bias	*RMSE*
$\hat{A}_0$	0.0000	0.0007	0.0057	0.0007	0.0219	-0.0552	0.0672	-0.0552	1.7446	-0.1104	0.0953	-0.1104	3.4926
$\hat{A}_1$	0.9602	0.9599	0.0062	-0.0003	0.0090	0.8788	0.0801	-0.0813	2.5720	0.7992	0.1113	-0.1610	5.0902
$\hat{B}_1$	0.2794	0.2800	0.0042	0.0006	0.0192	0.2559	0.0334	-0.0235	0.7440	0.2358	0.0451	-0.0436	1.3778
$\hat{C}_0$	0.0000	-0.0008	0.0043	-0.0008	0.0242	0.0183	0.0678	0.0183	0.5794	0.0401	0.0977	0.0401	1.2681
$\hat{C}_1$	-0.2794	-0.2817	0.0062	-0.0023	0.0712	-0.2557	0.0766	0.0237	0.7508	-0.2294	0.1077	0.0500	1.5823
$\hat{D}_1$	0.9602	0.9607	0.0036	0.0005	0.0173	0.8810	0.0445	-0.0792	2.5042	0.8016	0.0613	-0.1586	5.0145
Estimates	True value	Contaminated											
		30%				40%				50%			
		mean	*SE*	bias	*RMSE*	mean	*SE*	bias	*RMSE*	mean	*SE*	bias	*RMSE*
$\hat{A}_0$	0.0000	-0.1566	0.1111	-0.1566	4.9519	-0.2145	0.1198	-0.2145	6.7836	-0.2717	0.1368	-0.2717	8.5917
$\hat{A}_1$	0.9602	0.7280	0.1275	-0.2322	7.3417	0.6456	0.1357	-0.3146	9.9482	0.5628	0.1545	-0.3974	12.5668
$\hat{B}_1$	0.2794	0.2097	0.0521	-0.0697	2.2035	0.1865	0.0596	-0.0929	2.9371	0.1629	0.0646	-0.1165	3.6847
$\hat{C}_0$	0.0000	0.0503	0.1122	0.0503	1.5895	0.0662	0.1291	0.0662	2.0944	0.0899	0.1412	0.0899	2.8413
$\hat{C}_1$	-0.2794	-0.2113	0.1245	0.0681	2.1534	-0.1913	0.1409	0.0881	2.7872	-0.1653	0.1483	0.1141	3.6090
$\hat{D}_1$	0.9602	0.7215	0.0730	-0.2386	7.5466	0.6475	0.0810	-0.3127	9.8876	0.5662	0.0903	-0.3940	12.4591

3.5 Practical Example

We now apply the JS circular regression model on two real data sets; the local eye data and the wind direction. The eye data are collected from University Malaya Medical Centre while the wind direction data are obtained from Hussin *et al*. (2004).

3.5.1 Eye Data

We consider the eye data consisting of 23 observations of glaucoma patients (unit in radians) recorded using Optical coherence tomography (OCT) at the University Malaya Medical Centre (UMMC). OCT technology originally is used in ophthalmology to image the posterior segment, and has also been used to image anterior segment structures such as the cornea. The angle imaging of the anterior segment OCT in UMMC patients' eyes were obtained with Anterior Segment OCT (AS-OCT). The measurements selected are the angle of the posterior corneal curvature (u) and the angle of the eye (between posterior corneal curvature to iris) (v). The interest here is to model the data using JS circular regression model and to check the goodness of fit of the model.

Estimation of the model parameters

We fit the JS regression model on the data set using equations (3.6). The parameter estimates are given by $\hat{A}_0 = 1.0822$, $\hat{A}_1 = -0.1497$, $\hat{B}_1 = -0.3837$, $\hat{C}_0 = 0.0986$, $\hat{C}_1 = 0.2534$, $\hat{D}_1 = 0.5935$, $\hat{\kappa} = 22.51$, $\hat{\sigma}_1 = 0.16$ and $\hat{\sigma}_2 = 0.16$ and thus the fitted model with respect to $\hat{g}_1(u)$ and $\hat{g}_2(u)$ are as follows:

$$\hat{g}_1(u) = 1.0822 - 0.1497 \cos u - 0.3837 \sin u$$

$$\hat{g}_2(u) = 0.0986 + 0.2534 \cos u + 0.5935 \sin u .$$

Further, the prediction of $\hat{v}_j$ is given by

$$\mu(u) = \hat{v}_j = \arctan^* \frac{0.0986 + 0.2534 \cos u_j + 0.5935 \sin u_j}{1.0822 - 0.1497 \cos u_j - 0.3837 \sin u_j}, \; j = 1, \ldots, n,$$

and the concentration parameter ρ toward $\mu(u)$ using equation (3.10) is given by

$$\hat{\rho}(u) = \sqrt{\frac{1}{n} \sum_{j=1}^{n} \hat{\rho}^2 \left(u_j\right)} = \sqrt{\frac{1}{n} \sum_{j=1}^{n} \left[g_1^2(u_j) + g_2^2(u_j) \right]} = 0.9774$$

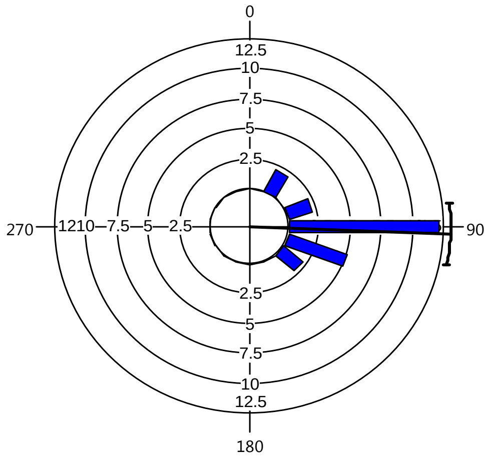
which suggest that the data seem to be highly concentrated since the value closer to one.

.

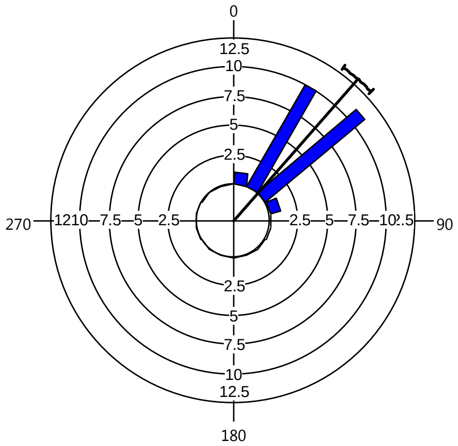
Goodness of fit of the JS model

The goodness of fit test is performed by using the Akaike information criteria (*AICC*) given by Lund (1999). The *AICC* value for the JS circular regression model is then compared with the values obtained for Hussin's circular regression model (Abuzaid *et. al.*, 2010) and Down & Mardia's circular regression model (Rambli, 2012). The *AICC* value for the JS circular regression model is -116.27, which is lower than that of Hussin circular regression model considered in Abuzaid *et al*. (2008), which is -112.91 and Down & Mardia model considered in Rambli (2011), which is -110.24. Hence, the JS circular regression model provides a better fit to the data.

This is supported by the results of the diagnostic plot. Figure 3.1 shows the plot of simple circular histograms for eye data measured by two different angles, i.e., the posterior corneal curvature and angel of the eye. The posterior corneal curvatures angles concentrated around 90° while the angel of the eye is more concentrated around 45°. Figure 3.2 shows the spoke plot of eye data. The inner ring represents the measurements by posterior corneal curvature while the outer ring represents the measurements of the angle of the eye. It can be observed that the lines do cross each other suggesting the data are not highly correlated with estimated correlation parameter $\hat{r}_c = -0.2791$. Figure 3.3 (a) and (b) show the Q-Q plots for the residuals resulting from the two observational regression-like equations of the JS circular regression model. The plot for ε_1 shows that most of the points are closer to the straight line except two points at the top right and lower left of the plot. Meanwhile, plot of ε_2 shows the points are relatively closer to the straight line. The outlying point might correspond to the outliers that may exist in the data. They will be dealt with in the next two chapters.



(a) Posterior corneal curvature



(b) Angel of the eye

Figure 3.1: Circular histograms for eye data

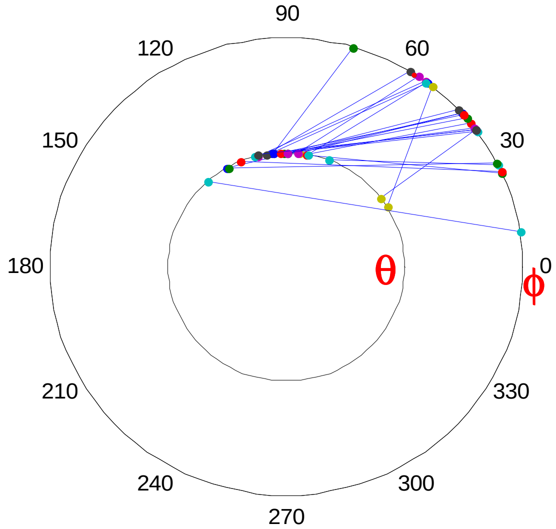
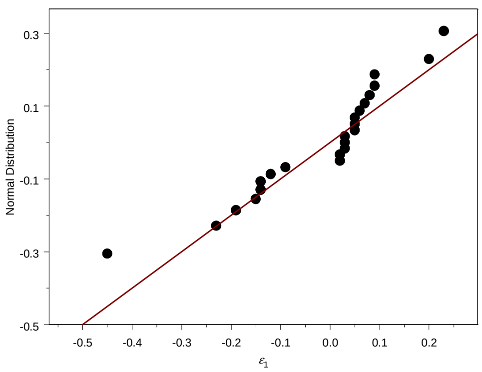
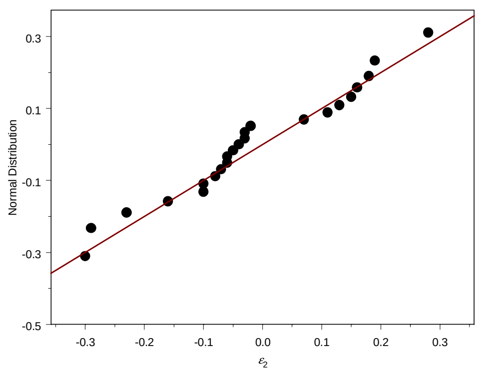


Figure 3.2 Spoke plot of eye data



(a) ε_1



(b) ε_2

Figure 3.3: Q-Q plot for residuals

3.5.2 Wind Direction Data

The wind direction data are given in Hussin *et al*. (2004) and are measured by using two different instruments; HF (High frequency) radar system and anchored wave buoy techniques for measuring the ocean wind direction. The data were collected along the Holderness coastline (the Humberside coast of the North Sea, United Kingdom). The wind direction is the direction of the local wind which blows across the sea surface and along the coast where the HF radar system and anchored wave buoys are deployed. The full data set is obtained from Hussin (1997) and is given in Appendix 1 which consists of time (in days) and the directions (in radians) being recorded. There were 129 measurements recorded by HF radar and anchored wave buoy respectively over the period of 22.7 days. Since both measurements of variables are circular, we fit the wind direction data using model (3.1).

Estimation of the model parameters

The least squares estimates of the parameters are obtained and given by $\hat{A}_0 = 0.0674$, $\hat{A}_1 = 0.7559$, $\hat{B}_1 = -0.0948$, $\hat{C}_0 = -0.047$, $\hat{C}_1 = 0.1049$, $\hat{D}_1 = 0.9762$, $\hat{\sigma}_1 = 0.3$ and $\hat{\sigma}_2 = 0.3$ and thus the fitted model gives $g_1(u)$ and $g_2(u)$ to be

$$\hat{g}_1(u) = 0.0674 + 0.7559 \cos u - 0.0948 \sin u$$

$$\hat{g}_2(u) = -0.047 + 0.1049 \cos u + 0.9762 \sin u \,.$$

Thus, model (3.1) is obtained such that

$$\mu(u) = \hat{v}_j = \arctan^* \frac{-0.047 + 0.1049 \cos u_j + 0.9762 \sin u_j}{0.0674 + 0.7559 \cos u_j - 0.0948 \sin u_j}, \quad j = 1, \ldots, n,$$

and the concentration parameter $\rho(u)$ is obtained using (3.10). The estimated concentration parameter is $\hat{\rho}(u) = 0.9322$.

Goodness of fit of the JS model

Figures 3.4 and 3.5 give the plots for diagnostic purpose. The spoke plot of wind direction data is shown in Figure 3.4. The correlation value is 0.8317 which implies a positive and strong correlation between the readings of the two instruments. Only two pairs of observations result in straight lines crossing the inner circle of the plot. One line cut diametrically across the inner circle (observation 38) while the other only cuts it in a short chord (observation 111). These observations are candidates to be outliers.

Figure 3.5 shows the Q-Q plot for residuals. The corresponding plot of ε_1 shows that most of the points are closer to the straight line except one point at the upper right of the plot. Meanwhile, plot of ε_2 also shows that most of the points are closer to the straight line except one point at the bottom left of the plot. Those points are candidate to be outliers.

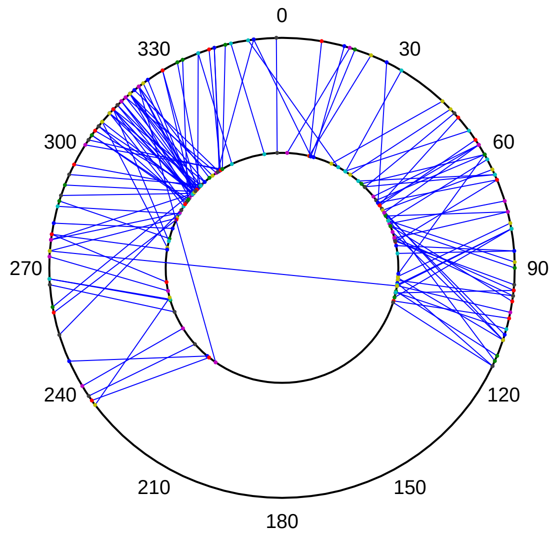
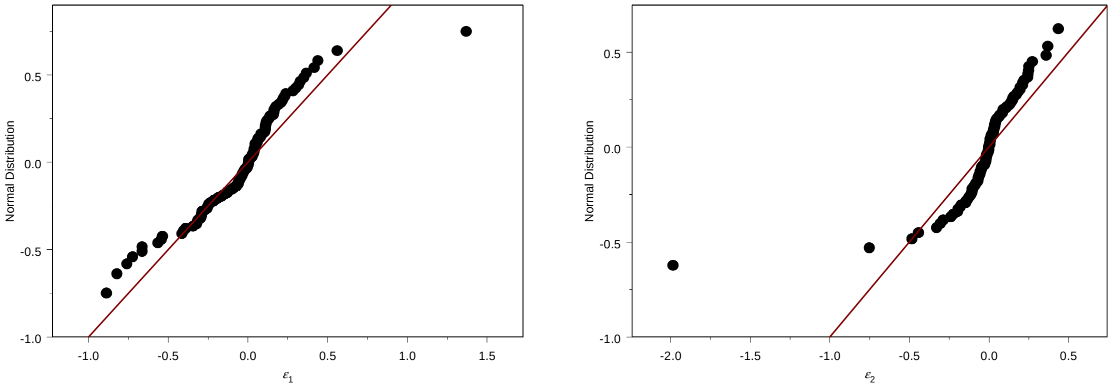


Figure 3.4: The spoke plot of wind direction data



(a) ε_1 (b) ε_2

Figure 3.5: Q-Q plot for residuals

3.6 Summary

We have considered the JS circular regression model for the case when both response and explanatory variables are circular. The theory of the model and parameter estimation method for the JS circular regression model based on the generalized least squares approach has been discussed. The simulation study has also been carried out to see the effect of outliers on the LS method. The application on two real data sets is presented.

CHAPTER FOUR

OUTLIER DETECTION IN A CIRCULAR REGRESSION MODEL USING *COVRATIO* STATISTIC

4.1 Introduction

In Section 2.5.1, we looked at some available tests to identify outliers in linear regression including *COVRATIO* statistic which is the ratio of the estimated covariance matrix of the estimated coefficients using all available data with estimated covariance matrix that results when the jth observation is deleted. We look at the background of the statistic in detail.

Belsley *et al*. (1980) used row deletion approach to investigate the impact of deleting one row at a time on estimated coefficients, fitted values, residuals and covariance matrix of linear regression models. In particular, they suggested a measure of influence based on the determinantal ratio given by

$$COVRATIO_{(-j)} = \frac{|COV_{(-j)}|}{|COV|},$$

where $|COV|$ is the determinant covariance matrix of coefficients for full data set and $|COV_{(-j)}|$ is for the reduced data set by excluding the jth row. If the ratio is close to the unity, then there is no significant difference between the covariance matrices. In other words, the jth observation is consistent with the other observations. Further, they proposed a statistic of the form $\left|COVRATIO_{(-j)} - 1\right|$ and established the cut-off point for testing the existence of outliers. That is, if the value of $\left|COVRATIO_{(-j)} - 1\right|$ is closer or larger than $(3p/n)$ then it indicates that the *j*th observation is a candidate to be an outlier, where *p* is the number of estimated coefficients and *n* is the sample size.

This chapter discusses the possibility of extending the idea of *COVRATIO* statistic to the JS circular regression models. It is motivated by the fact that the JS circular regression models have a closed form covariance matrix of the residuals ε_1 and ε_2 involved. In other word, instead of working with the covariance matrix of the parameters, we used the covariance matrix of the residuals of the two observational regression-like models given by equation (3.5). Here, we subsequently define a modified *COVRATIO* statistic in the present case. Simulation is carried out to obtain the cut-off points and to investigate the power of performance of the modified *COVRATIO* statistic.

4.2 Covariance matrix of JS circular regression model

We have discussed the covariance for the residuals of the JS circular regression model in Section 3.3.1. The covariance matrix has been given in equation (3.13) as stated below:

$$\hat{\boldsymbol{\Sigma}} = [n-2(2m+1)]^{-1} R_0, \tag{4.1}$$

where $R_0(p,q) = \boldsymbol{V}^{(\boldsymbol{p})'}\boldsymbol{V}^{(\boldsymbol{q})} - \boldsymbol{V}^{(\boldsymbol{p})'}\boldsymbol{U}(\boldsymbol{U}'\boldsymbol{U})^{-1}\boldsymbol{U}'\boldsymbol{V}^{(\boldsymbol{q})}$, $R_0 = (R_0(p,q))_{p,q=1,2}$ and

$$\underset{n\times(2m+1)}{\boldsymbol{U}} = \begin{bmatrix} 1 & \cos u_1 & \cdots & \cos mu_1 & \sin u_1 & \cdots & \sin mu_1 \\ 1 & \cos u_2 & \cdots & \cos mu_2 & \sin u_2 & \cdots & \sin mu_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & \cos u_n & \cdots & \cos mu_n & \sin u_n & \cdots & \sin mu_n \end{bmatrix},$$

$\boldsymbol{V}^{(\boldsymbol{p})}$ and $\boldsymbol{V}^{(\boldsymbol{q})}$ are the variable $\boldsymbol{V}^{(1)}$ and $\boldsymbol{V}^{(2)}$ respectively as defined in Section 3.3.1. Here, we discuss the extension of *COVRATIO* statistic to the circular case. That is the *COVRATIO* statistic in our case here is given by

$$COVRATIO_{(-j)} = \frac{|\hat{\boldsymbol{\Sigma}}_{(-j)}|}{|\hat{\boldsymbol{\Sigma}}|}, \qquad (4.2)$$

where $\hat{\boldsymbol{\Sigma}}$ is the covariance of the residuals of the full data and $\hat{\boldsymbol{\Sigma}}_{(-j)}$ is the corresponding covariance after the *j*th observation is deleted. Any observation with $\left|COVRATIO_{(-j)} - 1\right|$ exceeds the cut-off points will be identified as an outlier. The cut-off points are obtained via simulation in the following section.

4.3 Cut-off Points of Test Statistics

A simulation study is carried out to obtain the cut-off points of the test statistic for different sample sizes *n* and different values of standard deviation σ_1 and σ_2. Specifically, we generate sets of random errors from the bivariate Normal distribution with mean vector ***0*** for various combination of (σ_1, σ_2) in the range of [0.03,0.3] and *n* in the range [10,150]. The values of *a* = -3, -2, -1, 1, 2, 3 and 6 are also considered. We first find the critical values of the statistic $\left|COVRATIO_{(-j)} - 1\right|$. The complete steps to obtain the cut-off points are described below:

Step 1. Generate a variable ***U*** of size *n* from $VM(\pi, 2)$.

Step 2. Generate ε_1 and ε_2 of size *n* from $\boldsymbol{N}\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}\sigma_1 & 0\\ 0 & \sigma_2\end{pmatrix}\right)$. For a fixed *a*, obtain the true values of A_1, B_1, C_1 and D_1 with A_0 and C_0 are zero.

Step 3. Obtain the variable ***V*** using equation (3.6).

Step 4. Fit the generated circular data to the JS circular regression model to give the parameter estimates $\hat{A}_0, \hat{A}_1, \hat{B}_1, \hat{C}_0, \hat{C}_1$ and $\hat{D}_1$.

Step 5. Calculate |*COV*| from equation (3.13).

Step 6. Exclude the *j*th row from the generated sample, where $j = 1, ..., n$. For each j, repeat steps 4-5 for the reduced data set to obtain $|COV_{(-j)}|$.

Step 7. Compute $COVRATIO_{(-j)}$ and then obtain $\left|COVRATIO_{(-j)} - 1\right|$ for each j.

Step 8. Specify the maximum value of $\left|COVRATIO_{(-j)} - 1\right|$.

The process is carried out 500 times for each combination of sample size n and standard deviations (σ_1, σ_2). Then the 1%, 5% and 10% upper percentiles of the maximum values of $\left|COVRATIO_{(-j)} - 1\right|$ are calculated and used as the cut-off points of the proposed procedure. Tables 4.1 - 4.3 give the cut-off points of 1%, 5% and 10% percentiles for different n and standard deviations (σ_1, σ_2) at $a = 2$. The result shows that, for fixed σ_1 and $\sigma_2 \geq \sigma_1$, the cut-off points present an increasing trend as σ_2 gets larger. The same trend is seen when σ_2 is fixed and $\sigma_1 \geq \sigma_2$. On the other hand, the cut-off points are a decreasing function of the sample size *n*. Similar results are obtained for other values of *a* and are given in Appendix 2.

4.4 The Power of Performance of *COVRATIO* Statistic

A simulation study is carried out to investigate the performance of the $\left|COVRATIO_{(-j)} - 1\right|$ statistic for detecting outliers in the JS circular regression model (3.1) based on equation (4.2). Four different sample sizes are considered, n = 20, 30, 50, 70, 100 and 130.. The same procedure employed in Section 4.3 is used here to generate the data set. Then, the observation at position *d*, say v_d, is contaminated as follows:

$$v_d^* = v_d + \lambda\pi \quad mod\,(2\pi),$$

where v_d^* is the value after contamination and λ is the degree of contamination in the range $0 \le \lambda \le 1$. The generated data of $\boldsymbol{U}$ and $\boldsymbol{V}$ are then fitted to give the parameter estimate of $\hat{A}_0, \hat{A}_1, \hat{B}_1, \hat{C}_0, \hat{C}_1$ and $\hat{D}_1$ and $|COV|$ is calculated using equation (4.1). Consequently, exclude the *j*th row from the sample, for *j* = 1, …, *n* and refit the remaining data using equation (3.1). Then, the $COVRATIO_{(-j)}$ is then calculated. If the values of $\left| COVRATIO_{(-d)} - 1 \right|$ is maximum and greater than the cut-off point that obtained from the previous generated of cut-off points, then we say that the procedure has correctly detected the outlier in the data. The process is carried out 500 times. The power performance of the procedure is then examined by computing the percentage of the correct detection of the contaminated observation at position *d*.

The simulation results are plotted in Figures 4.1 – 4.2. Figure 4.1 illustrates the power of performance of the *COVRATIO* detection method for *n* = 70 and three different values of (σ_1, σ_2) = (0.03, 0.03), (0.05, 0.05) and (0.1, 0.1). It can be seen that the performance of the procedure is increasing as σ_1 and σ_2 get smaller. This is expected as V_{1j} and V_{2j} in equation (3.6) will fluctuate closer to the horizontal axis when σ_1 and σ_2 are closer to zero, and hence, better chance to detect the outlier even when λ is small.

On the other hand, Figure 4.2 gives the plot of power of performance of the *COVRATIO* detection method for fixed (σ_1, σ_2) = (0.1,0.1) and different values of *n* = 20, 30, 50, 70, 100 and 130. We observe that the performance is an increasing function of *n* but the curves are very close to each other when *n* is large enough. Similar patterns are also observed for the other cases.

Table 4.1: The 1% upper percentiles of the $\left| COVRATIO_{(-j)} - 1 \right|$ statistic at a = 2

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
10	**0.03**	0.9693	0.9786	0.9871	0.9793	0.9872	0.9895
	0.05	0.9655	0.9765	0.9743	0.9750	0.9815	0.9923
	0.08	0.9631	0.9763	0.9710	0.9700	0.9806	0.9902
	0.1	0.9601	0.9784	0.9763	0.9666	0.9795	0.9835
	0.3	0.9819	0.9745	0.9780	0.9635	0.9786	0.9748
	0.6	0.9794	0.9664	0.9763	0.9712	0.9704	0.9682
20	**0.03**	0.7791	0.7990	0.8435	0.8670	0.9318	0.9160
	0.05	0.7477	0.7774	0.8120	0.8294	0.9420	0.8530
	0.08	0.7862	0.7729	0.7781	0.7989	0.8966	0.8669
	0.1	0.8070	0.8061	0.8662	0.7764	0.8854	0.8341
	0.3	0.9003	0.9176	0.9556	0.9661	0.8528	0.8455
	0.6	0.8396	0.8333	0.8538	0.8591	0.7998	0.7609
30	**0.03**	0.4479	0.7677	0.8306	0.8481	0.8995	0.7614
	0.05	0.6979	0.7278	0.7881	0.8069	0.9063	0.7439
	0.08	0.7062	0.7167	0.7457	0.7815	0.8956	0.7586
	0.1	0.7195	0.7199	0.7240	0.7474	0.8890	0.7107
	0.3	0.8139	0.8158	0.8143	0.8165	0.8818	0.9134
	0.6	0.8785	0.8795	0.8773	0.9043	0.9321	0.9768
40	**0.03**	0.5554	0.6233	0.6938	0.7155	0.8383	0.8629
	0.05	0.5591	0.5996	0.6760	0.7008	0.8360	0.8228
	0.08	0.5900	0.6181	0.6678	0.6737	0.8313	0.6163
	0.1	0.6106	0.6041	0.6602	0.6754	0.8202	0.8086
	0.3	0.7117	0.7117	0.7042	0.6880	0.7670	0.8277
	0.6	0.7797	0.7899	0.7967	0.7875	0.8187	0.8285
50	**0.03**	0.5100	0.5863	0.6283	0.6515	0.7574	0.7747
	0.05	0.5611	0.5274	0.6010	0.6149	0.7314	0.7725
	0.08	0.5983	0.5913	0.5578	0.5745	0.7279	0.7805
	0.1	0.6210	0.6274	0.5894	0.5697	0.7016	0.7713
	0.3	0.6780	0.5748	0.6725	0.6703	0.7038	0.8285
	0.6	0.6899	0.7138	0.6972	0.6964	0.7890	0.7029
60	**0.03**	0.4479	0.5228	0.5900	0.6210	0.7291	0.6206
	0.05	0.4566	0.4814	0.5116	0.5640	0.7268	0.7199
	0.08	0.5127	0.5010	0.5095	0.5210	0.7028	0.6798
	0.1	0.5156	0.4758	0.5354	0.5276	0.7196	0.7363
	0.3	0.6165	0.6157	0.6194	0.6148	0.6387	0.6921
	0.6	0.6093	0.6526	0.6668	0.6536	0.7160	0.6465
70	**0.03**	0.4553	0.5258	0.5622	0.5677	0.6971	0.5939
	0.05	0.4245	0.4653	0.5334	0.5446	0.7403	0.5975
	0.08	0.4765	0.4461	0.4746	0.5042	0.7099	0.5760
	0.1	0.4764	0.4753	0.4681	0.4891	0.7032	0.5809
	0.3	0.5688	0.5823	0.5829	0.5748	0.6452	0.5946
	0.6	0.5648	0.5556	0.5496	0.5609	0.6615	0.6006

Table 4.1, continued.

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
80	**0.03**	0.3678	0.4199	0.4791	0.4892	0.5992	0.5704
	0.05	0.4042	0.3684	0.4176	0.4402	0.5868	0.5701
	0.08	0.4326	0.4177	0.3880	0.4038	0.5589	0.5475
	0.1	0.4421	0.4320	0.4061	0.3982	0.5492	0.5471
	0.3	0.5647	0.5476	0.5219	0.5365	0.5455	0.4948
	0.6	0.5718	0.5624	0.5624	0.5601	0.5829	0.4732
90	**0.03**	0.3265	0.3710	0.4390	0.4532	0.6383	0.5105
	0.05	0.3403	0.3267	0.3920	0.4264	0.6400	0.5156
	0.08	0.3794	0.3424	0.3476	0.3887	0.6294	0.5128
	0.1	0.3933	0.3654	0.3495	0.3554	0.6138	0.5166
	0.3	0.5145	0.4860	0.4675	0.4445	0.5359	0.4965
	0.6	0.5278	0.5403	0.5310	0.5165	0.5688	0.4381
100	**0.03**	0.3180	0.3712	0.4232	0.4502	0.6280	0.4684
	0.05	0.3398	0.3437	0.3812	0.4015	0.6132	0.4674
	0.08	0.3681	0.3629	0.3656	0.3813	0.6418	0.4578
	0.1	0.3734	0.3702	0.3497	0.3699	0.6203	0.4547
	0.3	0.4669	0.4571	0.4160	0.4069	0.5672	0.4434
	0.6	0.4666	0.4570	0.4715	0.4755	0.4796	0.3781
110	**0.03**	0.2960	0.3514	0.3709	0.3854	0.5716	0.4102
	0.05	0.3063	0.2964	0.3518	0.3612	0.5653	0.4303
	0.08	0.3269	0.3088	0.3049	0.3233	0.5606	0.4565
	0.1	0.3468	0.3335	0.3160	0.3076	0.5445	0.4635
	0.3	0.4292	0.4193	0.4224	0.4237	0.5064	0.3855
	0.6	0.3981	0.4100	0.4358	0.4328	0.5192	0.3392
130	**0.03**	0.2896	0.3077	0.3606	0.3931	0.5533	0.3753
	0.05	0.3147	0.2986	0.3046	0.3511	0.5424	0.3654
	0.08	0.3434	0.3276	0.3011	0.3213	0.5279	0.3604
	0.1	0.3464	0.3307	0.3257	0.3028	0.5149	0.3590
	0.3	0.4563	0.4612	0.4728	0.4449	0.4882	0.3292
	0.6	0.3541	0.3697	0.3905	0.3949	0.3716	0.3179
150	**0.03**	0.2251	0.2644	0.3200	0.3469	0.5387	0.3560
	0.05	0.2306	0.2261	0.2797	0.3020	0.5428	0.3638
	0.08	0.2552	0.2374	0.2518	0.2671	0.5432	0.3539
	0.1	0.2578	0.2432	0.2296	0.2572	0.5402	0.3412
	0.3	0.3890	0.4005	0.3930	0.3908	0.4140	0.3258
	0.6	0.3637	0.3778	0.3857	0.3793	0.3841	0.2322

Table 4.2: The 5% upper percentiles of the $\left| COVRATIO_{(-j)} - 1 \right|$ statistic at $a = 2$

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
10	**0.03**	0.9345	0.9392	0.9467	0.9495	0.9636	0.9639
	0.05	0.9208	0.9203	0.9397	0.9379	0.9614	0.9626
	0.08	0.9302	0.9253	0.9340	0.9366	0.9579	0.9665
	0.1	0.9311	0.9301	0.9293	0.9416	0.9630	0.9582
	0.3	0.9385	0.9434	0.9382	0.9226	0.9534	0.9588
	0.6	0.9431	0.9163	0.9409	0.9376	0.9418	0.9264
20	**0.03**	0.6741	0.7186	0.7608	0.7760	0.8404	0.8114
	0.05	0.6856	0.6754	0.7297	0.7603	0.8455	0.7865
	0.08	0.6996	0.6862	0.7053	0.7185	0.8099	0.7712
	0.1	0.7075	0.6967	0.8137	0.6887	0.7912	0.7591
	0.3	0.8268	0.8379	0.8594	0.8503	0.7110	0.7379
	0.6	0.7176	0.7417	0.7359	0.7399	0.7120	0.6659
30	**0.03**	0.5813	0.6445	0.7026	0.7360	0.8107	0.6876
	0.05	0.5957	0.5974	0.6818	0.6847	0.8000	0.6501
	0.08	0.6137	0.6213	0.6604	0.6662	0.7930	0.6458
	0.1	0.6162	0.6289	0.6262	0.6472	0.6856	0.6111
	0.3	0.6915	0.6850	0.6933	0.7131	0.7850	0.8246
	0.6	0.7978	0.7735	0.8022	0.7893	0.8295	0.8703
40	**0.03**	0.4918	0.5334	0.5789	0.5961	0.7315	0.7586
	0.05	0.4705	0.5139	0.5505	0.5711	0.7046	0.7163
	0.08	0.5058	0.5073	0.5170	0.5536	0.6984	0.5216
	0.1	0.5184	0.5169	0.5212	0.5421	0.6832	0.7319
	0.3	0.5846	0.6060	0.5849	0.5901	0.6516	0.7168
	0.6	0.6755	0.6700	0.6562	0.6759	0.7197	0.7274
50	**0.03**	0.4124	0.4735	0.5244	0.5431	0.6548	0.6535
	0.05	0.4240	0.4388	0.4974	0.5120	0.6453	0.6494
	0.08	0.4503	0.4390	0.4678	0.4917	0.6226	0.6476
	0.1	0.4615	0.4517	0.4618	0.4706	0.6094	0.6469
	0.3	0.5730	0.4613	0.5694	0.5632	0.6184	0.6279
	0.6	0.5645	0.5913	0.5928	0.5978	0.6337	0.6157
60	**0.03**	0.3791	0.4118	0.4607	0.4809	0.5825	0.5480
	0.05	0.3829	0.3939	0.4281	0.4516	0.6097	0.5866
	0.08	0.4090	0.4018	0.4099	0.4215	0.6090	0.5666
	0.1	0.4126	0.4059	0.4169	0.4232	0.5762	0.6010
	0.3	0.4988	0.5059	0.5000	0.4942	0.5446	0.5814
	0.6	0.5341	0.5557	0.5469	0.5451	0.5611	0.5328
70	**0.03**	0.3553	0.3751	0.4191	0.4459	0.6030	0.5002
	0.05	0.3496	0.3656	0.3803	0.4043	0.5994	0.5043
	0.08	0.3841	0.3631	0.3841	0.3940	0.5655	0.4990
	0.1	0.3934	0.3799	0.3821	0.3877	0.5578	0.5065
	0.3	0.4528	0.4397	0.4326	0.4350	0.5194	0.5179
	0.6	0.4725	0.4754	0.4784	0.4829	0.5070	0.4711

Table 4.2, continued.

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
80	**0.03**	0.3053	0.3417	0.3775	0.4044	0.5146	0.4611
	0.05	0.3188	0.3183	0.3456	0.3680	0.5132	0.4795
	0.08	0.3572	0.3373	0.3303	0.3428	0.4915	0.4856
	0.1	0.3626	0.3589	0.3418	0.3449	0.4882	0.4706
	0.3	0.4510	0.4414	0.4355	0.4331	0.4321	0.4328
	0.6	0.4529	0.4549	0.4583	0.4519	0.4647	0.3868
90	**0.03**	0.2762	0.3098	0.3569	0.3755	0.4973	0.4077
	0.05	0.2869	0.2866	0.3177	0.3379	0.4939	0.4054
	0.08	0.3052	0.3028	0.2934	0.3070	0.4995	0.4171
	0.1	0.3131	0.3102	0.2958	0.3076	0.4900	0.4275
	0.3	0.3993	0.4010	0.3966	0.3926	0.4422	0.3793
	0.6	0.4096	0.4075	0.4131	0.4188	0.4533	0.3308
100	**0.03**	0.2650	0.2995	0.3354	0.3571	0.5032	0.3685
	0.05	0.2683	0.2731	0.3141	0.3361	0.5155	0.3689
	0.08	0.2837	0.2846	0.2891	0.3013	0.5051	0.3855
	0.1	0.2881	0.2930	0.2872	0.2961	0.4892	0.3856
	0.3	0.3699	0.3661	0.3603	0.3545	0.4106	0.3738
	0.6	0.3822	0.3854	0.3897	0.3795	0.4007	0.3287
110	**0.03**	0.2379	0.2665	0.3096	0.3320	0.4460	0.3604
	0.05	0.2328	0.2423	0.2742	0.2981	0.4501	0.3685
	0.08	0.2531	0.2467	0.2460	0.2609	0.4205	0.3715
	0.1	0.2641	0.2566	0.2549	0.2550	0.4133	0.3718
	0.3	0.3338	0.3372	0.3428	0.3426	0.3797	0.3355
	0.6	0.3482	0.3489	0.3657	0.3681	0.3860	0.2762
130	**0.03**	0.2015	0.2381	0.2681	0.2850	0.4484	0.3051
	0.05	0.2150	0.2144	0.2440	0.2570	0.4529	0.3160
	0.08	0.2284	0.2186	0.2320	0.2414	0.4360	0.3173
	0.1	0.2363	0.2273	0.2288	0.2406	0.4144	0.3135
	0.3	0.3406	0.3450	0.3360	0.3265	0.3596	0.2923
	0.6	0.3051	0.3288	0.3321	0.3329	0.3228	0.2354
150	**0.03**	0.1806	0.2137	0.2497	0.2637	0.4221	0.2749
	0.05	0.1818	0.1877	0.2260	0.2405	0.4176	0.2729
	0.08	0.2031	0.1889	0.1944	0.2125	0.4068	0.2791
	0.1	0.2177	0.1962	0.1938	0.2021	0.3841	0.2831
	0.3	0.2735	0.2775	0.2742	0.2610	0.3392	0.2534
	0.6	0.2982	0.3057	0.3070	0.3117	0.3065	0.2076

Table 4.3: The 10% upper percentiles of the $\left| COVRATIO_{(-j)} - 1 \right|$ statistic at a = 2

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
10	**0.03**	0.8997	0.8992	0.9191	0.9230	0.9428	0.9445
	0.05	0.8917	0.8944	0.9030	0.9037	0.9395	0.9358
	0.08	0.8924	0.8927	0.8888	0.9022	0.9359	0.9416
	0.1	0.8913	0.8968	0.8894	0.8994	0.9322	0.9358
	0.3	0.8993	0.9062	0.8983	0.8923	0.9126	0.9233
	0.6	0.9068	0.8831	0.9007	0.8986	0.9042	0.8880
20	**0.03**	0.6330	0.6434	0.6937	0.7022	0.7699	0.7477
	0.05	0.6249	0.6328	0.6585	0.6771	0.7556	0.7287
	0.08	0.6415	0.6395	0.6493	0.6568	0.7488	0.7229
	0.1	0.6498	0.6344	0.7497	0.6392	0.7221	0.6986
	0.3	0.7750	0.7855	0.8000	0.8095	0.6571	0.6780
	0.6	0.6606	0.6702	0.6898	0.6823	0.6585	0.5891
30	**0.03**	0.5202	0.5710	0.6355	0.6536	0.7514	0.6126
	0.05	0.5356	0.5426	0.6066	0.6352	0.7494	0.5707
	0.08	0.5625	0.5566	0.5773	0.6084	0.7387	0.5881
	0.1	0.5729	0.5703	0.5770	0.5967	0.7301	0.5528
	0.3	0.6468	0.6430	0.6468	0.6437	0.7136	0.7540
	0.6	0.7078	0.7035	0.7445	0.7279	0.7760	0.7896
40	**0.03**	0.4308	0.4903	0.5204	0.5385	0.6544	0.6880
	0.05	0.4424	0.4537	0.4987	0.5216	0.6305	0.6671
	0.08	0.4565	0.4574	0.4836	0.4994	0.6338	0.4513
	0.1	0.4704	0.4621	0.4768	0.4928	0.6231	0.6711
	0.3	0.5180	0.5308	0.5249	0.5307	0.5803	0.6427
	0.6	0.6035	0.6078	0.6013	0.6113	0.6500	0.6582
50	**0.03**	0.3878	0.4242	0.4662	0.4824	0.5990	0.5811
	0.05	0.3769	0.4013	0.4360	0.4510	0.5926	0.5699
	0.08	0.3994	0.3933	0.4041	0.4267	0.5851	0.5798
	0.1	0.4061	0.4075	0.4088	0.4119	0.5676	0.5780
	0.3	0.4889	0.3841	0.4933	0.4959	0.5359	0.5726
	0.6	0.5110	0.5399	0.5282	0.5263	0.5821	0.5484
60	**0.03**	0.3385	0.3755	0.4075	0.4337	0.5393	0.5109
	0.05	0.3388	0.3487	0.3747	0.4020	0.5359	0.5359
	0.08	0.3621	0.3473	0.3589	0.3832	0.5278	0.5171
	0.1	0.3759	0.3548	0.3564	0.3688	0.5123	0.5166
	0.3	0.4312	0.4379	0.4474	0.4434	0.4769	0.5262
	0.6	0.4755	0.4896	0.4726	0.4762	0.5016	0.4588
70	**0.03**	0.3072	0.3330	0.3694	0.3899	0.5095	0.4388
	0.05	0.3052	0.3148	0.3453	0.3601	0.5091	0.4462
	0.08	0.3322	0.3214	0.3291	0.3384	0.4941	0.4613
	0.1	0.3384	0.3310	0.3320	0.3363	0.4892	0.4606
	0.3	0.3764	0.3832	0.3822	0.3810	0.4472	0.4464
	0.6	0.4218	0.4353	0.4376	0.4338	0.4473	0.4115

Table 4.3, continued.

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
80	**0.03**	0.2694	0.2969	0.3297	0.3434	0.4530	0.4125
	0.05	0.2748	0.2759	0.3069	0.3280	0.4551	0.4184
	0.08	0.3038	0.2876	0.2907	0.2965	0.4442	0.4215
	0.1	0.3140	0.3022	0.2939	0.3007	0.4305	0.4239
	0.3	0.3688	0.3894	0.3764	0.3733	0.3919	0.3776
	0.6	0.3959	0.3979	0.4008	0.4085	0.4223	0.3475
90	**0.03**	0.2437	0.2820	0.3218	0.3389	0.4343	0.3624
	0.05	0.2544	0.2558	0.2896	0.3090	0.4410	0.3627
	0.08	0.2740	0.2695	0.2682	0.2771	0.4294	0.3757
	0.1	0.2871	0.2746	0.2726	0.2717	0.4180	0.3813
	0.3	0.3434	0.3458	0.3515	0.3475	0.3717	0.3385
	0.6	0.3667	0.3687	0.3791	0.3685	0.3998	0.2959
100	**0.03**	0.2291	0.2564	0.2979	0.3155	0.4126	0.3409
	0.05	0.2293	0.2408	0.2585	0.2806	0.4169	0.3491
	0.08	0.2519	0.2374	0.2478	0.2601	0.4054	0.3480
	0.1	0.2588	0.2475	0.2399	0.2567	0.4040	0.3469
	0.3	0.3073	0.3210	0.3189	0.3133	0.3474	0.3288
	0.6	0.3451	0.3477	0.3469	0.3511	0.3625	0.2808
110	**0.03**	0.2085	0.2352	0.2737	0.2851	0.3919	0.3253
	0.05	0.2103	0.2130	0.2457	0.2672	0.3841	0.3397
	0.08	0.2228	0.2180	0.2181	0.2408	0.3740	0.3429
	0.1	0.2258	0.2277	0.2175	0.2285	0.3561	0.3403
	0.3	0.2875	0.2923	0.2929	0.2885	0.3151	0.3073
	0.6	0.3039	0.3056	0.3272	0.3196	0.3315	0.2523
130	**0.03**	0.1844	0.2113	0.2333	0.2453	0.3661	0.2713
	0.05	0.1890	0.1890	0.2193	0.2237	0.3708	0.2728
	0.08	0.1978	0.1938	0.2023	0.2101	0.3624	0.2720
	0.1	0.2062	0.2029	0.2044	0.2078	0.3515	0.2677
	0.3	0.2838	0.2832	0.2796	0.2750	0.3123	0.2587
	0.6	0.2674	0.2707	0.2724	0.2761	0.2861	0.2171
150	**0.03**	0.1583	0.1893	0.2173	0.2315	0.3611	0.2450
	0.05	0.1566	0.1689	0.1940	0.2105	0.3658	0.2470
	0.08	0.1779	0.1676	0.1743	0.1881	0.3424	0.2459
	0.1	0.1826	0.1763	0.1732	0.1801	0.3355	0.2477
	0.3	0.2496	0.2425	0.2340	0.2251	0.2783	0.2311
	0.6	0.2624	0.2712	0.2738	0.2730	0.2765	0.1866

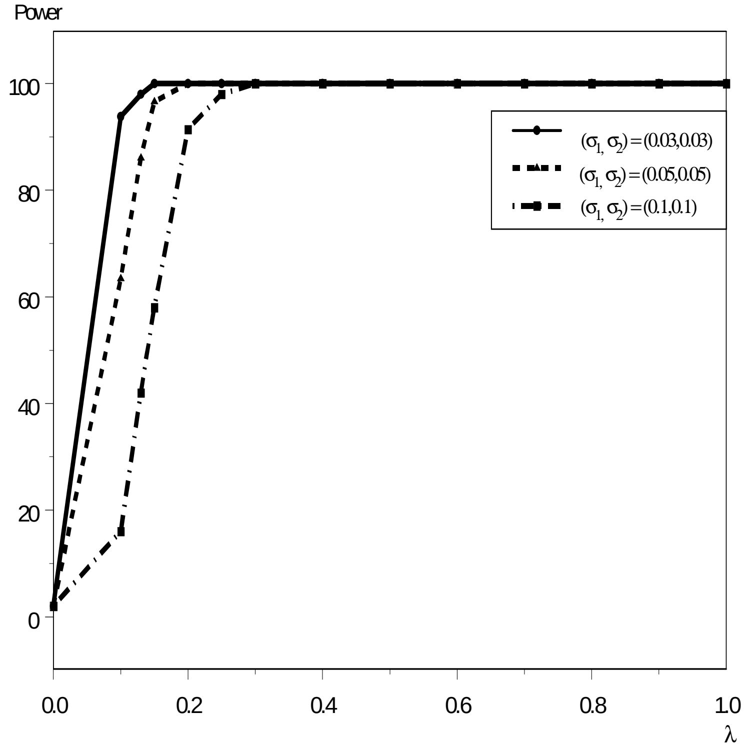


Figure 4.1: Graph of power performance for $| COVRATIO_{(-j)} - 1 |$ statistic, n=70 a = 2.

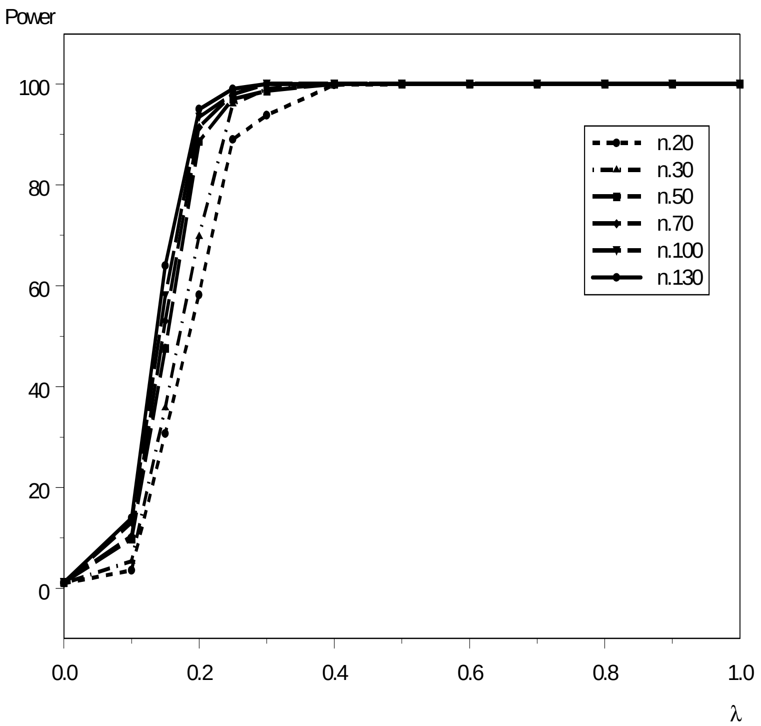


Figure 4.2: Graph of the power performance of the $\left| COVRATIO_{(-j)} - 1 \right|$ statistic, $(\sigma_1, \sigma_2) = (0.1, 0.1)$ at a = 2.

4.5 Practical Example: Wind Direction

We consider the wind direction data which have been described in Section 3.5.2. The least squares estimates are $\hat{A}_0 = 0.0674$, $\hat{A}_1 = 0.7559$, $\hat{B}_1 = -0.0948$, $\hat{C}_0 = -0.047$, $\hat{C}_1 = 0.1049$, $\hat{D}_1 = 0.9762$, $\hat{\sigma}_1 = 0.3$ and $\hat{\sigma}_2 = 0.3$ giving the fitted models of $g_1(u)$ and $g_2(u)$ as follows, respectively:

$$\hat{g}_1(u) = 0.0674 + 0.7559 \cos u - 0.0948 \sin u$$

$$\hat{g}_2(u) = -0.047 + 0.1049 \cos u + 0.9762 \sin u .$$

In addition, the Q-Q plots of the residuals associated with $\hat{g}_1(u)$ and $\hat{g}_2(u)$ suggest the occurrence of outliers in the data set. Hence, we apply the outlier detection procedure proposed in this chapter on the data set.

4.5.1 *COVRATIO* Statistic

Now, we apply the *COVRATIO* statistic to detect any possible outliers in the wind direction data. The determinant of the covariance matrix of the residual for the full data set $|COV|$ is 0.0043 and the corresponding values of $\left|COVRATIO_{(-j)} - 1\right|$, $j = 1, \ldots, n$, are then calculated. Due to the large number of parameters in the model, the cut-off point for this data can be approximated from the tabulated cut-off point in Section 4.3 or can be obtained directly using the simulation program used in Section 4.3 by taking the LS estimated parameter values as the true values. The program is given in Appendix 5. Consequently, using the later approach we find the cut-off point at 5% significance level to be 0.3544. Hence, we identify observation 38 to be an outlier. It is illustrated in Figure 4.3 which shows the corresponding value for observation number

38 is different from the others. This is further supported by considering the spoke plot in Figure 3.3 where the line for observation 111 only slightly cuts the inner circle.

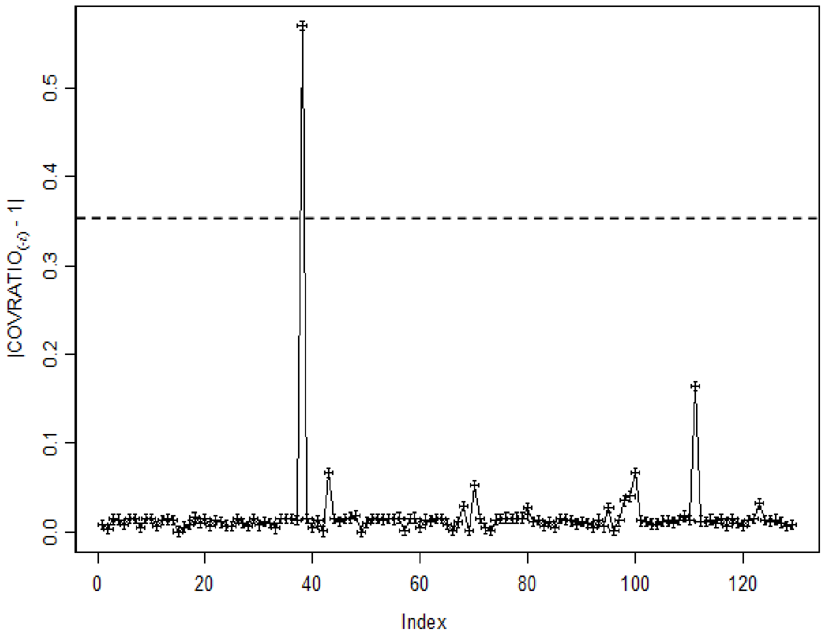
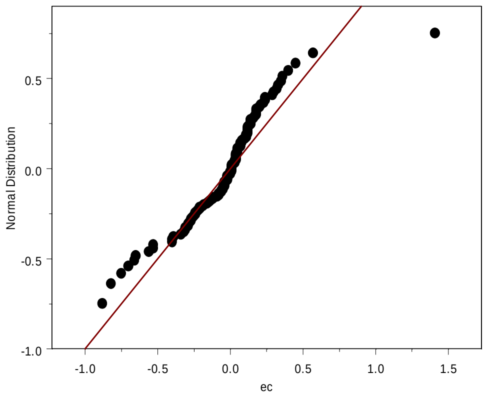


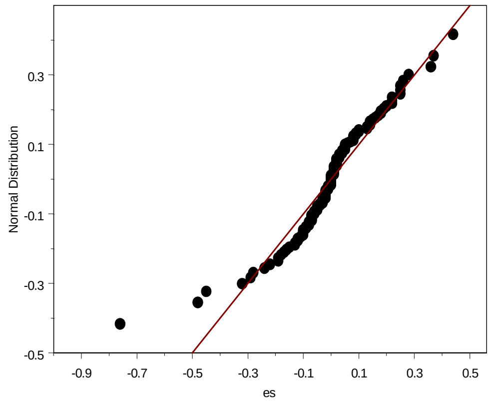
Figure 4.3: The values of the $\left| COVRATIO_{(-j)} - 1 \right|$ statistic for the wind direction data

Table 4.4: Parameter estimates for clean and contaminated data

Parameter estimates	**Contaminated data**	**Standard error**	**Clean data (case 38 deleted)**	**Standard error**
$\hat{A}_0$	0.0674	0.0361	0.0633	0.0365
$\hat{A}_1$	0.7559	0.0598	0.7609	0.0602
$\hat{B}_1$	-0.0948	0.0323	-0.0974	0.0325
$\hat{C}_0$	-0.0470	0.0291	-0.0114	0.0196
$\hat{C}_1$	0.1049	0.0483	0.062	0.0324
$\hat{D}_1$	0.9762	0.0261	0.9981	0.0175
$\hat{\sigma}_1$	0.300	0.2849	0.2800	0.2853
$\hat{\sigma}_2$	0.300	0.2300	0.1500	0.1533
$A(\hat{\kappa})$	0.9329	-	0.9479	-
$\hat{\kappa}$	7.7247	-	9.8749	-
$\hat{\rho}$	0.9322	-	0.9474	-



(a) ε_1



(b) ε_1

Figure 4.4: Q-Q plot for circular residuals without observations number 38

4.5.2 The Effect of Outliers on the Parameter Estimates

Table 4.4 summarizes the effect of excluding the outliers on the parameter estimates. The removal of observation number 38 significantly changes some of the estimated parameters of JS circular regression model. For instance, the estimated values of $\hat{C}_0$ and $\hat{D}_1$ changes the most with the estimated value $\hat{\sigma}^2$ changes by half, from 0.30 to 0.15. As expected, the estimated value of the concentration parameter $\hat{\kappa}$ also increases from 7.7247 to 9.8749, meaning that we have estimated $\hat{v}$ which are closer to observed values v of the wind direction data. The value of standard error for parameter estimates $\hat{A}_0$, $\hat{A}_1$ and $\hat{B}_1$ significantly not many changes for clean data, but the values of $\hat{C}_0$, $\hat{C}_1$ and $\hat{D}_1$ are smaller than contaminated data. On the other hand, Figure 4.4 gives the Q-Q plots of the resulting residuals corresponding to the observational regression-like models after removing observation number 38 from the wind data set. The points on the plots for ε_1 are now closer to the straight line compared to that in Figure 3.5(b). We conclude, by omitting observation 38 from the analysis, the reduced data is now better fitted using the JS circular regression model, though we note that there is a points lying a bit far from the straight line in Figure 4.4(a) which needs further investigation.

4.6 Summary

In this chapter, we use the idea of the *COVRATIO* statistic in linear case to identify influential observations in JS circular regression models. Here, the covariance considered is the covariance of the observational regression-like model of the model. The cut-off points are obtained and the powers performance examined through simulation studies for a simple model. We show that the sample size and dispersion of the residuals determine the level of cut-off points. The procedure also shows a good performance in identifying outlier in JS circular regression models.

CHAPTER FIVE

OUTLIER DETECTION IN A CIRCULAR REGRESSION MODEL USING *DMCE* STATISTICS

5.1 Introduction

It is important to study the residuals resulting from any regression modeling in order to investigate its model adequacy. In the case of multiple linear regression models, errors are assumed to be random, independent, identically and normally distributed with mean zero and constant variance. The standard definition of residuals for a linear regression model given by $e_j = y_j - \hat{y}_j$, where y_j and $\hat{y}_j$ are the observed and predicted values respectively, cannot be used directly on the circular regression models. For instance, let $y_j = 340°$ and $\hat{y}_j = 10°$. Then, the value $e_j = 340° - 10° = 330°$ is a totally different from the actual circular residual, which is $30°$.

Few definitions of circular residuals can be found in the literature. Using the definition of circular distance proposed by Rao (1969), Mardia (1972) defined the circular residual for the *j*th observation as

$$e_j^* = 1 - \cos\left(v_j - \hat{v}_j\right).$$

Here, e_j^* is linear and is bounded within the interval $[0,2]$. Thus, we are not able to use this residual to investigate the assumption of error that follows a specific circular distribution such as the von Mises distribution. Similarly, we may use the definition of circular distance as given Jammalamadaka & SenGupta (2001) to give

$$e_j^{'} = \pi - \left|\pi - \left|v_j - \hat{v}_j\right|\right|$$

where $e_j^{'} \in [0, \pi]$. However, again, this residual cannot be used directly to investigate the assumption of circular errors. For example, it is not possible to show the circular mean of such residuals to be zero. Moreover, the estimated concentration parameter also tends to increase as the residuals are distributed in the interval $[0, \pi]$ instead of the entire circumference. Thus, Abuzaid *et al*. (2008) proposed a new definition of circular residual based on circular distance given by

$$r_{A_j} = \begin{cases} (\pi - |\pi - |y_j - \hat{y}_j||), & \text{if} \quad \hat{y}_j \leq y_j,\ y_j - \hat{y}_j \leq \pi \quad \text{or} \quad \hat{y}_j > y_j,\ \hat{y}_j - y_j > \pi \\ -(\pi - |\pi - |y_j - \hat{y}_j||), & \text{if} \quad \hat{y}_j \leq y_j,\ y_j - \hat{y}_j > \pi \quad \text{or} \quad \hat{y}_j > y_j,\ \hat{y}_j - y_j \leq \pi \end{cases}.$$

which is in the range $[-\pi, \pi]$. These residuals have been shown to be useful in investigating the goodness-of-fit of simple linear regression models of Hussin *et al*. (2004). Numerical and simulation studies were carried out to show that the circular residuals r_{A_j}, j=1,2,…,n are uncorrelated and follow a von Mises distribution with circular mean *0* and concentration parameter κ.

The circular residuals above can be used to detect outliers in circular regression models. Abuzaid *et al*. (2008) looked at the possibility of identifying outliers in Hussin's circular regression model via residual analysis using a new definition of circular residuals based on circular distance. Later, the same authors proposed a statistic in terms of the circular distance for detecting outliers in the same type of circular regression model by using row deletion approach. In the next section, we look at two statistics that can be used to detect outliers in JS circular regression model.

5.2 Difference Mean Circular Error (*DMCE*) Statistics

The circular distance between two circular observations θ_i and θ_j is defined by Rao (1969) as $d_{ij} = 1 - \cos(\theta_i - \theta_j)$, where d_{ij} is a monotone increasing function of $(\theta_i - \theta_j)$ and $d_{ij} \in [0,2]$. Let the statistic be known as mean circular error $(MCEc)$ given by

$$MCEc = 1 - \frac{1}{n}\sum_{j=1}^{n} \cos(v_j - \hat{v}_j)\,, \tag{5.1}$$

where n is the sample size, $\hat{v}_j$ is the fitted values of v_j under model (3.1) and $MCEc \in [0,2]$.

The circular distance between v_j and the fitted value $\hat{v}_j$ is expected to be relatively large if an observation v_j is defined as an outliers in the data set. Thus, the existence of such observation in a data set will increase the summation of all circular distances as well as the value of $MCEc$ statistic. Subsequently, the removal of the *j*th observation denoted by $MCEc_{(-j)}$ from the data set will decrease the value of the statistic. The maximum absolute difference between the value of the statistics for full and reduced data sets is given by

$$DMCEc = \max_j \left\{ \left| MCEc - MCEc_{(-j)} \right| \right\}. \tag{5.2}$$

The *j*th observation is identified as an outliers if $DMCEc$ exceeds a pre-specified cut-off point.

On the other hand, Jammalamadaka & SenGupta (2001) gave a new defintion of circular distance between any two points as the smaller of the two arc length between the two points along the circumferences. For any two angles ϕ and θ, the circular distance is defined by

$$d_{\circ}(\phi,\theta)=\min(\phi-\theta,2\pi-(\phi-\theta))=\pi-\left|\pi-|\phi-\theta|\right|.$$

Using this particular definition of circular distance, we can develop an alternative statistic as a measure of mean circular error using the *sine* function, where *sine* is an increasing function on the interval $[0,\pi/2]$. This mean circular error is defined as

$$MCEs=\frac{1}{n}\sum_{j=1}^{n}\sin\left(\frac{d_j}{2}\right), \tag{5.3}$$

where $d_j=\pi-\left|\pi-|v_i-\hat{v}_i|\right|$ is the circular distance between v_j and $\hat{v}_j$, with sample size *n* and $MCEs\in[0,1]$.

Using similar argument used in the *MCEc* statistic, if an observation v_j is an outlier, then the half of the circular distance $\frac{d_j}{2}$ is expected to be relatively large compared to other $\frac{d_j}{2}$'s. Thus, the existence of such observation in a data set will increase the value of $MCEs$. Consequently, the removal of the *j*th observation denoted by $MCEs_{(-j)}$ from the data set will decrease the value of the statistic. The maximum absolute difference between the value of the statistics for full and reduced data sets is given by

$$DMCEs=\max_{j}\left\{\left|MCEs-MCEs_{(-j)}\right|\right\}. \tag{5.4}$$

The *j*th observation is also identified as an outlier if the corresponding value of $DMCEs$ exceeds a pre-specified cut-off point.

In the following section, simulation studies are carried out to find the cut-off points and to investigate the power of performance of *DMCEc* and $DMCEs$ for the JS circular regression models.

5.3 Percentiles Points of Test Statistics

(i) Description of Simulation Process

The percentile points are obtained by using Monte Carlo simulation method for *DMCEc* and *DMCEs* statistics for different sample sizes n and different values of standard deviation σ_1 and σ_2. Specifically, we generate sets of random errors from the bivariate Normal distribution with mean vector $\boldsymbol{0}$ for various combination of (σ_1, σ_2) in the range of [0.03,0.6] and n in the range [10,150]. Samples of von Mises distribution $VM(\pi, 2)$ with corresponding size n are generated to represent the values of $\boldsymbol{U}$ variable. Then, we generate ε_1 and ε_2 of size n from $N\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}\sigma_1 & 0\\ 0 & \sigma_2\end{pmatrix}\right)$. For a fixed a, obtain the true values of A_1, B_1, C_1 and D_1. The true values of A_0 and C_0 being zero here. We obtain the variable $\boldsymbol{V}$ using equation (3.6). We then compute the value of $MCEc$ and $MCEs$ statistics for full data set by using equation (5.1) and (5.3), respectively. Sequentially, we exclude the jth row from the generated sample, where j = 1, ..., n. We fit the reduced data using equation (3.6) and calculate the values of $MCEc_{(-j)}$ and $MCEs_{(-j)}$. Then, we obtain the value of $DMCEc$ and $DMCEs$, respectively. The process is repeated 500 times for each combination of sample size n and various combination of standard deviations (σ_1, σ_2).

Then the 1%, 5% and 10% upper percentiles of the maximum values of $DMCEc$ and $DMCEs$ are calculated and used as the cut-off points of the proposed procedure. Tables 5.1-5.6 give the 1%, 5% and 10% cut-off points for a = 2 for $DMCEc$ and $DMCEs$ respectively.

Table 5.1: The simulated 1% points of *DMCEc* statistic for a = 2

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
10	**0.03**	0.0004	0.0009	0.0021	0.0033	0.0388	0.1970
	0.05	0.0009	0.0012	0.0026	0.0035	0.0375	0.1975
	0.08	0.0024	0.0024	0.0035	0.0043	0.0388	0.1962
	0.1	0.0038	0.0037	0.0043	0.0053	0.0397	0.2004
	0.3	0.0355	0.0373	0.0402	0.0423	0.0659	0.2209
	0.6	0.1380	0.1437	0.1528	0.1525	0.1984	0.2885
20	**0.03**	0.0002	0.0005	0.0012	0.0020	0.0307	0.0964
	0.05	0.0006	0.0007	0.0013	0.0019	0.0313	0.0954
	0.08	0.0015	0.0015	0.0017	0.0023	0.0329	0.0953
	0.1	0.0022	0.0023	0.0024	0.0026	0.0344	0.0958
	0.3	0.0241	0.0240	0.0214	0.0230	0.0469	0.0973
	0.6	0.0920	0.0929	0.0926	0.0949	0.1044	0.1291
30	**0.03**	0.0002	0.0004	0.0011	0.0017	0.0261	0.1200
	0.05	0.0004	0.0006	0.0012	0.0019	0.0262	0.1207
	0.08	0.0011	0.0012	0.0016	0.0024	0.0268	0.1230
	0.1	0.0018	0.0018	0.0022	0.0029	0.0270	0.1252
	0.3	0.0233	0.0254	0.0259	0.0278	0.0652	0.1488
	0.6	0.0969	0.1080	0.1111	0.1180	0.1533	0.2048
40	**0.03**	0.0002	0.0003	0.0009	0.0014	0.0185	0.0853
	0.05	0.0004	0.0004	0.0010	0.0016	0.0188	0.0888
	0.08	0.0010	0.0011	0.0013	0.0018	0.0196	0.0922
	0.1	0.0015	0.0016	0.0019	0.0021	0.0197	0.0921
	0.3	0.0172	0.0175	0.0190	0.0195	0.0393	0.1196
	0.6	0.0808	0.0805	0.0804	0.0830	0.1111	0.1688
50	**0.03**	0.0001	0.0003	0.0009	0.0014	0.0196	0.0700
	0.05	0.0003	0.0004	0.0009	0.0014	0.0205	0.0702
	0.08	0.0009	0.0009	0.0012	0.0015	0.0207	0.0730
	0.1	0.0014	0.0015	0.0016	0.0018	0.0216	0.0751
	0.3	0.0136	0.0143	0.0156	0.0161	0.0353	0.0891
	0.6	0.0656	0.0689	0.0725	0.0747	0.0898	0.1223
60	**0.03**	0.0001	0.0002	0.0006	0.0010	0.0139	0.0573
	0.05	0.0003	0.0003	0.0006	0.0010	0.0136	0.0590
	0.08	0.0007	0.0008	0.0008	0.0011	0.0144	0.0602
	0.1	0.0011	0.0012	0.0013	0.0013	0.0146	0.0616
	0.3	0.0116	0.0120	0.0130	0.0136	0.0365	0.0722
	0.6	0.0505	0.0527	0.0563	0.0593	0.0715	0.0969
70	**0.03**	0.0001	0.0002	0.0006	0.0010	0.0209	0.0487
	0.05	0.0003	0.0003	0.0006	0.0010	0.0221	0.0507
	0.08	0.0007	0.0007	0.0009	0.0011	0.0233	0.0519
	0.1	0.0011	0.0011	0.0012	0.0014	0.0243	0.0523
	0.3	0.0127	0.0133	0.0145	0.0153	0.0371	0.0608
	0.6	0.0457	0.0475	0.0495	0.0507	0.0626	0.0872

Table 5.1, continued.

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
80	**0.03**	0.0001	0.0002	0.0005	0.0008	0.0119	0.0404
	0.05	0.0002	0.0002	0.0005	0.0008	0.0122	0.0418
	0.08	0.0005	0.0006	0.0007	0.0009	0.0124	0.0434
	0.1	0.0009	0.0009	0.0010	0.0011	0.0124	0.0446
	0.3	0.0120	0.0126	0.0128	0.0131	0.0216	0.0512
	0.6	0.0403	0.0408	0.0422	0.0437	0.0546	0.0677
90	**0.03**	0.0001	0.0002	0.0004	0.0007	0.0144	0.0378
	0.05	0.0002	0.0002	0.0004	0.0007	0.0145	0.0383
	0.08	0.0005	0.0006	0.0006	0.0008	0.0150	0.0393
	0.1	0.0008	0.0008	0.0009	0.0010	0.0152	0.0402
	0.3	0.0094	0.0112	0.0114	0.0119	0.0221	0.0460
	0.6	0.0341	0.0355	0.0377	0.0381	0.0481	0.0580
100	**0.03**	0.0001	0.0002	0.0004	0.0007	0.0021	0.0102
	0.05	0.0002	0.0002	0.0005	0.0008	0.0159	0.0340
	0.08	0.0005	0.0005	0.0006	0.0008	0.0161	0.0345
	0.1	0.0008	0.0008	0.0008	0.0009	0.0159	0.0349
	0.3	0.0111	0.0111	0.0121	0.0124	0.0223	0.0416
	0.6	0.0306	0.0319	0.0330	0.0342	0.0438	0.0489
110	**0.03**	0.0001	0.0002	0.0004	0.0099	0.0122	0.0301
	0.05	0.0002	0.0002	0.0004	0.0006	0.0120	0.0311
	0.08	0.0004	0.0005	0.0005	0.0007	0.0121	0.0313
	0.1	0.0007	0.0007	0.0008	0.0009	0.0125	0.0326
	0.3	0.0089	0.0092	0.0096	0.0096	0.0201	0.0376
	0.6	0.0279	0.0297	0.0297	0.0308	0.0402	0.0473
130	**0.03**	0.0001	0.0001	0.0004	0.0006	0.0110	0.0247
	0.05	0.0001	0.0002	0.0003	0.0006	0.0115	0.0255
	0.08	0.0004	0.0004	0.0004	0.0006	0.0122	0.0259
	0.1	0.0006	0.0006	0.0006	0.0007	0.0127	0.0270
	0.3	0.0078	0.0083	0.0085	0.0087	0.0227	0.0298
	0.6	0.0230	0.0240	0.0247	0.0015	0.0351	0.0407
150	**0.03**	0.0001	0.0001	0.0003	0.0005	0.0137	0.0212
	0.05	0.0001	0.0001	0.0003	0.0005	0.0143	0.0215
	0.08	0.0004	0.0004	0.0004	0.0005	0.0145	0.0223
	0.1	0.0005	0.0006	0.0006	0.0007	0.0145	0.0224
	0.3	0.0080	0.0086	0.0099	0.0105	0.0164	0.0263
	0.6	0.0197	0.0208	0.0220	0.0222	0.0281	0.0327

Table 5.2: The simulated 5% points of *DMCEc* statistic for $a = 2$

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
10	**0.03**	0.0003	0.0005	0.0012	0.0019	0.0220	0.1700
	0.05	0.0007	0.0008	0.0014	0.0021	0.0225	0.1694
	0.08	0.0016	0.0018	0.0021	0.0026	0.0230	0.1706
	0.1	0.0026	0.0028	0.0031	0.0033	0.0224	0.1710
	0.3	0.0224	0.0228	0.0253	0.0240	0.0467	0.1928
	0.6	0.0766	0.0785	0.0895	0.0920	0.1269	0.2080
20	**0.03**	0.0002	0.0004	0.0009	0.0013	0.0165	0.0881
	0.05	0.0004	0.0005	0.0010	0.0014	0.0168	0.0882
	0.08	0.0010	0.0011	0.0014	0.0017	0.0164	0.0884
	0.1	0.0016	0.0017	0.0019	0.0022	0.0166	0.0889
	0.3	0.0133	0.0135	0.0145	0.0149	0.0273	0.0918
	0.6	0.0674	0.0668	0.0722	0.0752	0.0905	0.1025
30	**0.03**	0.0001	0.0003	0.0008	0.0012	0.0151	0.0957
	0.05	0.0003	0.0004	0.0008	0.0013	0.0155	0.0963
	0.08	0.0008	0.0009	0.0012	0.0015	0.0160	0.1000
	0.1	0.0013	0.0014	0.0017	0.0019	0.0162	0.1025
	0.3	0.0151	0.0155	0.0167	0.0173	0.0324	0.1240
	0.6	0.0778	0.0849	0.0848	0.0851	0.1033	0.1601
40	**0.03**	0.0001	0.0003	0.0007	0.0011	0.0123	0.0750
	0.05	0.0003	0.0003	0.0007	0.0011	0.0128	0.0771
	0.08	0.0007	0.0008	0.0010	0.0013	0.0138	0.0810
	0.1	0.0011	0.0012	0.0014	0.0017	0.0140	0.0819
	0.3	0.0114	0.0123	0.0129	0.0141	0.0273	0.0968
	0.6	0.0611	0.0639	0.0652	0.0650	0.0842	0.1258
50	**0.03**	0.0001	0.0002	0.0006	0.0009	0.0119	0.0595
	0.05	0.0002	0.0003	0.0006	0.0009	0.0120	0.0622
	0.08	0.0006	0.0007	0.0008	0.0010	0.0124	0.0636
	0.1	0.0010	0.0011	0.0011	0.0013	0.0127	0.0647
	0.3	0.0101	0.0105	0.0110	0.0111	0.0208	0.0754
	0.6	0.0512	0.0541	0.0551	0.0570	0.0718	0.0967
60	**0.03**	0.0001	0.0002	0.0005	0.0008	0.0096	0.0506
	0.05	0.0002	0.0002	0.0005	0.0008	0.0099	0.0516
	0.08	0.0005	0.0005	0.0007	0.0009	0.0102	0.0532
	0.1	0.0008	0.0009	0.0010	0.0011	0.0104	0.0542
	0.3	0.0084	0.0086	0.0094	0.0096	0.0195	0.0636
	0.6	0.0437	0.0448	0.0475	0.0484	0.0605	0.0817
70	**0.03**	0.0001	0.0002	0.0004	0.0007	0.0091	0.0430
	0.05	0.0002	0.0002	0.0004	0.0007	0.0093	0.0439
	0.08	0.0005	0.0005	0.0007	0.0008	0.0101	0.0451
	0.1	0.0007	0.0008	0.0009	0.0010	0.0103	0.0458
	0.3	0.0081	0.0085	0.0089	0.0091	0.0207	0.0546
	0.6	0.0388	0.0404	0.0415	0.0430	0.0516	0.0666

Table 5.2, continued.

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
80	**0.03**	0.0001	0.0002	0.0004	0.0006	0.0079	0.0380
	0.05	0.0002	0.0002	0.0004	0.0006	0.0080	0.0390
	0.08	0.0004	0.0004	0.0005	0.0007	0.0083	0.0403
	0.1	0.0007	0.0007	0.0008	0.0009	0.0085	0.0406
	0.3	0.0069	0.0072	0.0076	0.0076	0.0161	0.0479
	0.6	0.0350	0.0365	0.0378	0.0392	0.0473	0.0579
90	**0.03**	0.0001	0.0001	0.0003	0.0005	0.0080	0.0333
	0.05	0.0001	0.0002	0.0004	0.0005	0.0081	0.0342
	0.08	0.0004	0.0004	0.0005	0.0006	0.0086	0.0354
	0.1	0.0006	0.0006	0.0007	0.0008	0.0092	0.0358
	0.3	0.0065	0.0069	0.0071	0.0073	0.0163	0.0413
	0.6	0.0306	0.0315	0.0329	0.0341	0.0423	0.0510
100	**0.03**	0.0001	0.0001	0.0003	0.0005	0.0073	0.0304
	0.05	0.0001	0.0002	0.0003	0.0005	0.0075	0.0312
	0.08	0.0004	0.0004	0.0005	0.0006	0.0078	0.0322
	0.1	0.0006	0.0006	0.0007	0.0008	0.0081	0.0325
	0.3	0.0061	0.0064	0.0067	0.0069	0.0146	0.0367
	0.6	0.0273	0.0280	0.0294	0.0304	0.0390	0.0451
110	**0.03**	0.0001	0.0001	0.0003	0.0005	0.0066	0.0272
	0.05	0.0001	0.0002	0.0003	0.0005	0.0069	0.0279
	0.08	0.0003	0.0003	0.0004	0.0006	0.0073	0.0285
	0.1	0.0005	0.0005	0.0006	0.0007	0.0074	0.0293
	0.3	0.0057	0.0059	0.0061	0.0063	0.0124	0.0339
	0.6	0.0257	0.0266	0.0280	0.0285	0.0343	0.0409
130	**0.03**	0.0000	0.0001	0.0003	0.0004	0.0070	0.0230
	0.05	0.0001	0.0001	0.0003	0.0004	0.0070	0.0234
	0.08	0.0003	0.0003	0.0003	0.0005	0.0074	0.0239
	0.1	0.0004	0.0004	0.0005	0.0006	0.0075	0.0245
	0.3	0.0052	0.0055	0.0056	0.0058	0.0141	0.0281
	0.6	0.0214	0.0223	0.0233	0.0237	0.0288	0.0347
150	**0.03**	0.0001	0.0001	0.0002	0.0003	0.0067	0.0200
	0.05	0.0001	0.0001	0.0002	0.0004	0.0069	0.0205
	0.08	0.0003	0.0003	0.0003	0.0004	0.0071	0.0209
	0.1	0.0004	0.0004	0.0004	0.0005	0.0074	0.0213
	0.3	0.0044	0.0045	0.0048	0.0050	0.0115	0.0246
	0.6	0.0189	0.0194	0.0204	0.0209	0.0246	0.0302

Table 5.3: The simulated 10% points of *DMCEc* statistic for a = 2

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
10	**0.03**	0.0002	0.0004	0.0010	0.0015	0.0166	0.1409
	0.05	0.0005	0.0007	0.0011	0.0016	0.0166	0.1383
	0.08	0.0013	0.0014	0.0018	0.0021	0.0174	0.1363
	0.1	0.0020	0.0021	0.0024	0.0028	0.0175	0.1376
	0.3	0.0172	0.0177	0.0190	0.0197	0.0364	0.1655
	0.6	0.0587	0.0589	0.0626	0.0635	0.0945	0.1902
20	**0.03**	0.0002	0.0003	0.0007	0.0011	0.0127	0.0814
	0.05	0.0004	0.0004	0.0008	0.0011	0.0126	0.0817
	0.08	0.0009	0.0009	0.0011	0.0014	0.0131	0.0823
	0.1	0.0013	0.0014	0.0016	0.0017	0.0135	0.0822
	0.3	0.0112	0.0113	0.0117	0.0122	0.0220	0.0862
	0.6	0.0528	0.0529	0.0542	0.0545	0.0794	0.0952
30	**0.03**	0.0001	0.0003	0.0006	0.0010	0.0117	0.0859
	0.05	0.0003	0.0004	0.0007	0.0010	0.0120	0.0859
	0.08	0.0007	0.0008	0.0010	0.0012	0.0130	0.0877
	0.1	0.0012	0.0012	0.0015	0.0016	0.0129	0.0905
	0.3	0.0118	0.0123	0.0128	0.0131	0.0239	0.1086
	0.6	0.0569	0.0621	0.0643	0.0660	0.0862	0.1390
40	**0.03**	0.0001	0.0002	0.0005	0.0008	0.0104	0.0700
	0.05	0.0002	0.0003	0.0006	0.0008	0.0105	0.0711
	0.08	0.0006	0.0006	0.0008	0.0011	0.0110	0.0745
	0.1	0.0009	0.0010	0.0011	0.0014	0.0113	0.0744
	0.3	0.0095	0.0095	0.0102	0.0106	0.0206	0.0859
	0.6	0.0477	0.0498	0.0505	0.0519	0.0720	0.1158
50	**0.03**	0.0001	0.0002	0.0005	0.0007	0.0090	0.0556
	0.05	0.0002	0.0002	0.0005	0.0008	0.0092	0.0564
	0.08	0.0005	0.0005	0.0007	0.0009	0.0093	0.0578
	0.1	0.0008	0.0008	0.0010	0.0011	0.0098	0.0584
	0.3	0.0078	0.0082	0.0085	0.0090	0.0170	0.0706
	0.6	0.0448	0.0460	0.0478	0.0488	0.0640	0.0892
60	**0.03**	0.0001	0.0002	0.0004	0.0006	0.0076	0.0463
	0.05	0.0002	0.0002	0.0004	0.0006	0.0079	0.0475
	0.08	0.0004	0.0005	0.0006	0.0008	0.0082	0.0484
	0.1	0.0007	0.0007	0.0008	0.0010	0.0083	0.0495
	0.3	0.0071	0.0073	0.0077	0.0081	0.0152	0.0572
	0.6	0.0386	0.0395	0.0407	0.0416	0.0540	0.0737
70	**0.03**	0.0001	0.0001	0.0003	0.0006	0.0072	0.0402
	0.05	0.0002	0.0002	0.0004	0.0006	0.0076	0.0411
	0.08	0.0004	0.0004	0.0005	0.0007	0.0078	0.0420
	0.1	0.0006	0.0007	0.0007	0.0009	0.0079	0.0431
	0.3	0.0067	0.0071	0.0075	0.0076	0.0153	0.0503
	0.6	0.0342	0.0355	0.0366	0.0378	0.0477	0.0621

Table 5.3, continued.

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
80	**0.05**	0.0001	0.0002	0.0003	0.0005	0.0066	0.0364
	0.08	0.0004	0.0004	0.0005	0.0006	0.0069	0.0375
	0.1	0.0006	0.0006	0.0007	0.0008	0.0072	0.0375
	0.3	0.0057	0.0059	0.0063	0.0064	0.0123	0.0442
	0.6	0.0322	0.0330	0.0346	0.0353	0.0432	0.0546
90	**0.03**	0.0001	0.0001	0.0003	0.0004	0.0060	0.0315
	0.05	0.0001	0.0002	0.0003	0.0005	0.0060	0.0323
	0.08	0.0003	0.0003	0.0004	0.0006	0.0064	0.0334
	0.1	0.0005	0.0005	0.0006	0.0007	0.0065	0.0340
	0.3	0.0055	0.0058	0.0061	0.0062	0.0129	0.0384
	0.6	0.0275	0.0284	0.0304	0.0322	0.0388	0.0471
100	**0.03**	0.0000	0.0001	0.0003	0.0004	0.0059	0.0288
	0.05	0.0001	0.0001	0.0003	0.0004	0.0059	0.0294
	0.08	0.0003	0.0003	0.0004	0.0005	0.0062	0.0301
	0.1	0.0005	0.0005	0.0005	0.0006	0.0063	0.0307
	0.3	0.0050	0.0052	0.0055	0.0056	0.0115	0.0349
	0.6	0.0255	0.0262	0.0270	0.0280	0.0352	0.0425
110	**0.03**	0.0000	0.0001	0.0003	0.0004	0.0052	0.0260
	0.05	0.0001	0.0001	0.0003	0.0004	0.0054	0.0265
	0.08	0.0003	0.0003	0.0004	0.0005	0.0055	0.0271
	0.1	0.0004	0.0004	0.0005	0.0006	0.0056	0.0280
	0.3	0.0044	0.0049	0.0051	0.0052	0.0105	0.0320
	0.6	0.0239	0.0254	0.0261	0.0267	0.0321	0.0391
130	**0.03**	0.0000	0.0001	0.0002	0.0003	0.0048	0.0221
	0.05	0.0001	0.0001	0.0002	0.0004	0.0049	0.0226
	0.08	0.0002	0.0002	0.0003	0.0004	0.0052	0.0229
	0.1	0.0004	0.0004	0.0004	0.0005	0.0054	0.0232
	0.3	0.0043	0.0045	0.0046	0.0047	0.0107	0.0269
	0.6	0.0205	0.0212	0.0220	0.0224	0.0274	0.0328
150	**0.03**	0.0001	0.0001	0.0002	0.0003	0.0049	0.0195
	0.05	0.0001	0.0001	0.0002	0.0003	0.0049	0.0199
	0.08	0.0002	0.0002	0.0003	0.0004	0.0050	0.0203
	0.1	0.0003	0.0004	0.0004	0.0004	0.0052	0.0205
	0.3	0.0037	0.0038	0.0040	0.0042	0.0081	0.0236
	0.6	0.0180	0.0187	0.0196	0.0199	0.0231	0.0285

Table 5.4: The simulated 1% points of *DMCEs* statistic for a = 2

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
	0.03	0.0057	0.0083	0.0138	0.0163	0.0595	0.1221
	0.05	0.0076	0.0102	0.0133	0.0183	0.0595	0.1227
10	**0.08**	0.0122	0.0130	0.0166	0.0199	0.0607	0.1199
	0.1	0.0154	0.0157	0.0197	0.0226	0.0619	0.1227
	0.3	0.0509	0.0510	0.0562	0.0561	0.0723	0.1304
	0.6	0.1057	0.1127	0.1111	0.1096	0.1176	0.1526
	0.03	0.0025	0.0038	0.0060	0.0076	0.0311	0.0547
	0.05	0.0037	0.0043	0.0061	0.0077	0.0314	0.0543
20	**0.08**	0.0061	0.0062	0.0070	0.0081	0.0293	0.0549
	0.1	0.0076	0.0076	0.0079	0.0083	0.0293	0.0549
	0.3	0.0250	0.0259	0.0265	0.0261	0.0348	0.0526
	0.6	0.0546	0.0550	0.0537	0.0545	0.0570	0.0696
	0.03	0.0022	0.0033	0.0060	0.0079	0.0290	0.0668
	0.05	0.0030	0.0039	0.0057	0.0075	0.0300	0.0673
30	**0.08**	0.0050	0.0049	0.0068	0.0081	0.0285	0.0698
	0.1	0.0061	0.0062	0.0074	0.0087	0.0288	0.0711
	0.3	0.0250	0.0272	0.0285	0.0280	0.0395	0.0887
	0.6	0.0615	0.0654	0.0685	0.0709	0.0791	0.1301
	0.03	0.0016	0.0026	0.0041	0.0053	0.0206	0.0526
	0.05	0.0024	0.0027	0.0047	0.0057	0.0224	0.0525
40	**0.08**	0.0040	0.0041	0.0048	0.0064	0.0209	0.0554
	0.1	0.0048	0.0050	0.0056	0.0068	0.0206	0.0559
	0.3	0.0189	0.0187	0.0199	0.0210	0.0331	0.0653
	0.6	0.0502	0.0517	0.0504	0.0507	0.0620	0.0883
	0.03	0.0009	0.0020	0.0031	0.0041	0.0183	0.0392
	0.05	0.0019	0.0022	0.0034	0.0043	0.0185	0.0398
50	**0.08**	0.0031	0.0033	0.0037	0.0046	0.0185	0.0406
	0.1	0.0008	0.0041	0.0046	0.0046	0.0179	0.0403
	0.3	0.0142	0.0146	0.0160	0.0165	0.0258	0.0480
	0.6	0.0389	0.0405	0.0411	0.0436	0.0510	0.0640
	0.03	0.0010	0.0016	0.0027	0.0033	0.0157	0.0307
	0.05	0.0015	0.0018	0.0027	0.0033	0.0149	0.0313
60	**0.08**	0.0024	0.0027	0.0030	0.0035	0.0139	0.0320
	0.1	0.0030	0.0034	0.0037	0.0041	0.0146	0.0327
	0.3	0.0122	0.0120	0.0129	0.0135	0.0225	0.0355
	0.6	0.0287	0.0314	0.0332	0.0356	0.0387	0.0555
	0.03	0.0009	0.0014	0.0023	0.0029	0.0130	0.0263
	0.05	0.0014	0.0016	0.0022	0.0029	0.0140	0.0261
70	**0.08**	0.0023	0.0024	0.0026	0.0031	0.0137	0.0272
	0.1	0.0030	0.0030	0.0032	0.0034	0.0144	0.0278
	0.3	0.0114	0.0125	0.0131	0.0130	0.0203	0.0317
	0.6	0.0273	0.0260	0.0267	0.0276	0.0323	0.0449

Table 5.4, continued.

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
80	**0.03**	0.0007	0.0012	0.0019	0.0024	0.0103	0.0228
	0.05	0.0012	0.0012	0.0019	0.0024	0.0104	0.0231
	0.08	0.0019	0.0019	0.0021	0.0026	0.0106	0.0242
	0.1	0.0024	0.0024	0.0026	0.0027	0.0106	0.0242
	0.3	0.0107	0.0114	0.0116	0.0116	0.0148	0.0257
	0.6	0.0234	0.0244	0.0252	0.0260	0.0310	0.0344
90	**0.03**	0.0007	0.0010	0.0016	0.0021	0.0112	0.0194
	0.05	0.0011	0.0012	0.0016	0.0020	0.0106	0.0198
	0.08	0.0016	0.0018	0.0020	0.0023	0.0107	0.0202
	0.1	0.0020	0.0022	0.0024	0.0026	0.0115	0.0204
	0.3	0.0093	0.0095	0.0101	0.0101	0.0148	0.0242
	0.6	0.0197	0.0209	0.0218	0.0225	0.0250	0.0292
100	**0.03**	0.0006	0.0010	0.0017	0.0021	0.0102	0.0177
	0.05	0.0010	0.0011	0.0016	0.0022	0.0100	0.0179
	0.08	0.0016	0.0017	0.0019	0.0022	0.0102	0.0189
	0.1	0.0020	0.0021	0.0022	0.0025	0.0101	0.0188
	0.3	0.0094	0.0094	0.0096	0.0100	0.0131	0.0201
	0.6	0.0182	0.0191	0.0194	0.0200	0.0226	0.0259
110	**0.03**	0.0005	0.0009	0.0015	0.0019	0.0087	0.0161
	0.05	0.0009	0.0010	0.0014	0.0019	0.0088	0.0166
	0.08	0.0015	0.0015	0.0016	0.0019	0.0093	0.0169
	0.1	0.0018	0.0020	0.0020	0.0021	0.0099	0.0176
	0.3	0.0068	0.0075	0.0077	0.0079	0.0121	0.0186
	0.6	0.0165	0.0176	0.0187	0.0190	0.0230	0.0239
130	**0.03**	0.0004	0.0007	0.0012	0.0016	0.0076	0.0127
	0.05	0.0008	0.0008	0.0011	0.0015	0.0077	0.0133
	0.08	0.0013	0.0013	0.0013	0.0015	0.0078	0.0134
	0.1	0.0016	0.0017	0.0016	0.0017	0.0081	0.0137
	0.3	0.0064	0.0068	0.0066	0.0068	0.0109	0.0152
	0.6	0.0139	0.0147	0.0152	0.0150	0.0172	0.0197
150	**0.03**	0.0004	0.0006	0.0010	0.0013	0.0078	0.0118
	0.05	0.0006	0.0007	0.0010	0.0013	0.0078	0.0124
	0.08	0.0010	0.0011	0.0012	0.0014	0.0075	0.0120
	0.1	0.0013	0.0013	0.0014	0.0015	0.0078	0.0119
	0.3	0.0057	0.0059	0.0060	0.0065	0.0086	0.0134
	0.6	0.0122	0.0125	0.0128	0.0125	0.0143	0.0173

Table 5.5: The simulated 5% points of *DMCEs* statistic for a = 2

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
10	**0.03**	0.0045	0.0058	0.0093	0.0119	0.0413	0.1074
	0.05	0.0061	0.0074	0.0095	0.0119	0.0413	0.1059
	0.08	0.0095	0.0104	0.0118	0.0130	0.0421	0.1039
	0.1	0.0123	0.0133	0.0139	0.0149	0.0406	0.1053
	0.3	0.0383	0.0392	0.0413	0.0421	0.0545	0.1042
	0.6	0.0791	0.0781	0.0817	0.0811	0.0901	0.1153
20	**0.03**	0.0021	0.0031	0.0048	0.0060	0.0212	0.0483
	0.05	0.0032	0.0035	0.0050	0.0062	0.0216	0.0481
	0.08	0.0050	0.0052	0.0058	0.0066	0.0216	0.0480
	0.1	0.0062	0.0064	0.0066	0.0071	0.0215	0.0482
	0.3	0.0189	0.0190	0.0191	0.0191	0.0258	0.0456
	0.6	0.0459	0.0438	0.0443	0.0445	0.0478	0.0547
30	**0.03**	0.0016	0.0026	0.0043	0.0056	0.0217	0.0550
	0.05	0.0025	0.0029	0.0057	0.0055	0.0215	0.0556
	0.08	0.0040	0.0043	0.0051	0.0063	0.0225	0.0568
	0.1	0.0051	0.0053	0.0060	0.0066	0.0225	0.0578
	0.3	0.0198	0.0203	0.0209	0.0218	0.0313	0.0659
	0.6	0.0481	0.0514	0.0520	0.0535	0.0628	0.0882
40	**0.03**	0.0012	0.0019	0.0033	0.0042	0.0166	0.0412
	0.05	0.0019	0.0022	0.0034	0.0046	0.0170	0.0421
	0.08	0.0031	0.0032	0.0040	0.0048	0.0168	0.0440
	0.1	0.0040	0.0040	0.0046	0.0051	0.0170	0.0453
	0.3	0.0147	0.0151	0.0156	0.0162	0.0242	0.0498
	0.6	0.0373	0.0390	0.0398	0.0397	0.0495	0.0664
50	**0.03**	0.0010	0.0016	0.0026	0.0034	0.0129	0.0342
	0.05	0.0015	0.0017	0.0026	0.0033	0.0133	0.0345
	0.08	0.0025	0.0026	0.0030	0.0034	0.0132	0.0349
	0.1	0.0040	0.0033	0.0035	0.0039	0.0139	0.0346
	0.3	0.0111	0.0120	0.0126	0.0123	0.0180	0.0391
	0.6	0.0308	0.0325	0.0340	0.0352	0.0382	0.0497
60	**0.03**	0.0009	0.0013	0.0021	0.0028	0.0110	0.0279
	0.05	0.0013	0.0015	0.0022	0.0028	0.0112	0.0284
	0.08	0.0021	0.0022	0.0026	0.0030	0.0114	0.0281
	0.1	0.0026	0.0027	0.0030	0.0033	0.0116	0.0288
	0.3	0.0093	0.0096	0.0103	0.0107	0.0161	0.0322
	0.6	0.0251	0.0260	0.0264	0.0282	0.0324	0.0418
70	**0.03**	0.0008	0.0011	0.0018	0.0023	0.0100	0.0228
	0.05	0.0011	0.0013	0.0019	0.0024	0.0102	0.0234
	0.08	0.0018	0.0019	0.0021	0.0025	0.0103	0.0240
	0.1	0.0023	0.0024	0.0026	0.0028	0.0101	0.0241
	0.3	0.0089	0.0089	0.0093	0.0094	0.0145	0.0271
	0.6	0.0217	0.0226	0.0228	0.0230	0.0271	0.0338

Table 5.5, continued.

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
80	**0.03**	0.0006	0.0010	0.0016	0.0021	0.0081	0.0198
	0.05	0.0010	0.0011	0.0017	0.0021	0.0080	0.0203
	0.08	0.0016	0.0016	0.0018	0.0022	0.0084	0.0208
	0.1	0.0020	0.0021	0.0022	0.0024	0.0086	0.0210
	0.3	0.0074	0.0074	0.0078	0.0077	0.0117	0.0224
	0.6	0.0200	0.0209	0.0214	0.0218	0.0238	0.0290
90	**0.03**	0.0006	0.0009	0.0014	0.0018	0.0077	0.0178
	0.05	0.0009	0.0010	0.0014	0.0018	0.0084	0.0181
	0.08	0.0014	0.0015	0.0017	0.0019	0.0083	0.0184
	0.1	0.0018	0.0019	0.0020	0.0022	0.0085	0.0187
	0.3	0.0066	0.0069	0.0074	0.0076	0.0116	0.0197
	0.6	0.0166	0.0179	0.0187	0.0192	0.0214	0.0247
100	**0.03**	0.0005	0.0008	0.0013	0.0015	0.0069	0.0159
	0.05	0.0008	0.0009	0.0013	0.0016	0.0071	0.0160
	0.08	0.0013	0.0013	0.0016	0.0018	0.0072	0.0165
	0.1	0.0016	0.0017	0.0018	0.0020	0.0074	0.0170
	0.3	0.0059	0.0062	0.0066	0.0067	0.0101	0.0176
	0.6	0.0152	0.0157	0.0166	0.0170	0.0194	0.0218
110	**0.03**	0.0005	0.0007	0.0012	0.0015	0.0063	0.0146
	0.05	0.0007	0.0008	0.0012	0.0015	0.0063	0.0149
	0.08	0.0012	0.0012	0.0014	0.0016	0.0064	0.0154
	0.1	0.0015	0.0015	0.0016	0.0018	0.0065	0.0154
	0.3	0.0056	0.0059	0.0061	0.0061	0.0093	0.0161
	0.6	0.0142	0.0154	0.0159	0.0162	0.0181	0.0202
130	**0.03**	0.0004	0.0006	0.0010	0.0013	0.0058	0.0120
	0.05	0.0006	0.0007	0.0010	0.0013	0.0060	0.0122
	0.08	0.0010	0.0010	0.0012	0.0013	0.0062	0.0125
	0.1	0.0012	0.0013	0.0014	0.0015	0.0063	0.0127
	0.3	0.0050	0.0051	0.0051	0.0054	0.0088	0.0134
	0.6	0.0121	0.0124	0.0129	0.0132	0.0148	0.0171
150	**0.03**	0.0003	0.0005	0.0009	0.0011	0.0050	0.0105
	0.05	0.0005	0.0006	0.0009	0.0011	0.0051	0.0108
	0.08	0.0009	0.0009	0.0010	0.0012	0.0051	0.0110
	0.1	0.0011	0.0011	0.0012	0.0013	0.0053	0.0110
	0.3	0.0045	0.0046	0.0046	0.0047	0.0070	0.0114
	0.6	0.0104	0.0106	0.0114	0.0116	0.0127	0.0148

Table 5.6: The simulated 10% points of $DMCEs$ statistic for $a = 2$

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
10	**0.03**	0.0037	0.0052	0.0079	0.0097	0.0330	0.0945
	0.05	0.0052	0.0064	0.0084	0.0098	0.0334	0.0943
	0.08	0.0083	0.0090	0.0101	0.0114	0.0339	0.0937
	0.1	0.0106	0.0109	0.0118	0.0124	0.0343	0.0941
	0.3	0.0318	0.0336	0.0340	0.0343	0.0447	0.0928
	0.6	0.0654	0.0651	0.0641	0.0652	0.0793	0.1022
20	**0.03**	0.0019	0.0027	0.0043	0.0054	0.0185	0.0450
	0.05	0.0028	0.0032	0.0043	0.0053	0.0184	0.0445
	0.08	0.0045	0.0045	0.0051	0.0058	0.0185	0.0435
	0.1	0.0056	0.0055	0.0059	0.0064	0.0180	0.0432
	0.3	0.0168	0.0166	0.0169	0.0171	0.0224	0.0429
	0.6	0.0386	0.0384	0.0390	0.0402	0.0417	0.0475
30	**0.03**	0.0015	0.0022	0.0036	0.0045	0.0186	0.0481
	0.05	0.0022	0.0026	0.0037	0.0046	0.0184	0.0498
	0.08	0.0036	0.0038	0.0046	0.0051	0.0192	0.0514
	0.1	0.0046	0.0047	0.0055	0.0057	0.0195	0.0519
	0.3	0.0169	0.0177	0.0182	0.0189	0.0265	0.0597
	0.6	0.0407	0.0421	0.0444	0.0452	0.0536	0.0780
40	**0.03**	0.0011	0.0017	0.0028	0.0036	0.0144	0.0382
	0.05	0.0017	0.0020	0.0028	0.0036	0.0142	0.0385
	0.08	0.0027	0.0028	0.0033	0.0040	0.0149	0.0398
	0.1	0.0034	0.0036	0.0039	0.0045	0.0148	0.0400
	0.3	0.0122	0.0125	0.0130	0.0135	0.0207	0.0439
	0.6	0.0315	0.0330	0.0342	0.0350	0.0424	0.0570
50	**0.03**	0.0009	0.0013	0.0023	0.0028	0.0116	0.0310
	0.05	0.0013	0.0016	0.0023	0.0030	0.0119	0.0313
	0.08	0.0021	0.0023	0.0027	0.0031	0.0119	0.0318
	0.1	0.0031	0.0028	0.0031	0.0035	0.0119	0.0320
	0.3	0.0096	0.0097	0.0102	0.0108	0.0163	0.0354
	0.6	0.0274	0.0282	0.0287	0.0293	0.0346	0.0456
60	**0.03**	0.0008	0.0012	0.0019	0.0025	0.0097	0.0252
	0.05	0.0011	0.0013	0.0019	0.0025	0.0099	0.0256
	0.08	0.0018	0.0019	0.0023	0.0027	0.0100	0.0259
	0.1	0.0023	0.0024	0.0027	0.0030	0.0101	0.0263
	0.3	0.0083	0.0086	0.0090	0.0091	0.0138	0.0285
	0.6	0.0219	0.0226	0.0234	0.0243	0.0290	0.0362
70	**0.03**	0.0007	0.0010	0.0016	0.0021	0.0085	0.0211
	0.05	0.0010	0.0011	0.0017	0.0021	0.0085	0.0215
	0.08	0.0016	0.0017	0.0020	0.0023	0.0086	0.0219
	0.1	0.0021	0.0022	0.0023	0.0025	0.0086	0.0221
	0.3	0.0076	0.0079	0.0083	0.0086	0.0125	0.0248
	0.6	0.0195	0.0200	0.0209	0.0211	0.0249	0.0305

Table 5.6, continued.

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
80	**0.03**	0.0006	0.0009	0.0014	0.0019	0.0071	0.0187
	0.05	0.0008	0.0010	0.0014	0.0019	0.0073	0.0190
	0.08	0.0014	0.0015	0.0017	0.0020	0.0075	0.0192
	0.1	0.0018	0.0019	0.0020	0.0022	0.0077	0.0195
	0.3	0.0065	0.0066	0.0068	0.0070	0.0102	0.0206
	0.6	0.0182	0.0193	0.0197	0.0198	0.0217	0.0255
90	**0.03**	0.0005	0.0008	0.0013	0.0017	0.0069	0.0168
	0.05	0.0008	0.0009	0.0013	0.0017	0.0069	0.0170
	0.08	0.0013	0.0014	0.0015	0.0018	0.0071	0.0173
	0.1	0.0016	0.0017	0.0018	0.0020	0.0071	0.0175
	0.3	0.0059	0.0062	0.0064	0.0066	0.0101	0.0185
	0.6	0.0153	0.0156	0.0163	0.0167	0.0194	0.0221
100	**0.03**	0.0005	0.0007	0.0011	0.0015	0.0062	0.0150
	0.05	0.0007	0.0008	0.0012	0.0015	0.0062	0.0152
	0.08	0.0012	0.0012	0.0014	0.0016	0.0062	0.0155
	0.1	0.0015	0.0016	0.0017	0.0018	0.0063	0.0157
	0.3	0.0054	0.0055	0.0057	0.0057	0.0086	0.0167
	0.6	0.0142	0.0146	0.0150	0.0154	0.0181	0.0202
110	**0.03**	0.0004	0.0006	0.0011	0.0013	0.0055	0.0137
	0.05	0.0006	0.0007	0.0011	0.0014	0.0056	0.0140
	0.08	0.0010	0.0011	0.0013	0.0015	0.0058	0.0140
	0.1	0.0013	0.0014	0.0015	0.0016	0.0059	0.0143
	0.3	0.0047	0.0050	0.0053	0.0054	0.0080	0.0152
	0.6	0.0132	0.0139	0.0144	0.0147	0.0161	0.0187
130	**0.03**	0.0003	0.0005	0.0009	0.0012	0.0049	0.0114
	0.05	0.0005	0.0006	0.0009	0.0012	0.0050	0.0116
	0.08	0.0009	0.0009	0.0011	0.0012	0.0053	0.0119
	0.1	0.0012	0.0012	0.0013	0.0014	0.0050	0.0119
	0.3	0.0044	0.0045	0.0046	0.0047	0.0074	0.0127
	0.6	0.0111	0.0113	0.0119	0.0121	0.0136	0.0158
150	**0.03**	0.0003	0.0005	0.0008	0.0010	0.0044	0.0100
	0.05	0.0005	0.0005	0.0008	0.0010	0.0045	0.0102
	0.08	0.0008	0.0008	0.0009	0.0011	0.0046	0.0104
	0.1	0.0010	0.0010	0.0011	0.0012	0.0047	0.0106
	0.3	0.0038	0.0039	0.0040	0.0040	0.0060	0.0109
	0.6	0.0098	0.0102	0.0105	0.0105	0.0116	0.0133

(ii) Discussion

Tables 5.1-5.6 present the cut-off points for *DMCEc* and *DMCEs* statistics at different level of significance. The result shows that, for fixed σ_1 and $\sigma_2 \geq \sigma_1$, the cut-off points for *DMCEc* and *DMCEs* statistics show an increasing trend as σ_2 gets larger. The same trend is seen when σ_2 is fixed and $\sigma_1 \geq \sigma_2$. On the other hand, the cut-off points are a decreasing function of the sample size *n*. Similar results are observed for other values of *a*. For all cases considered, the cut-off point of *DMCEc* statistics are generally smaller than that of *DMCEs*.

5.4 The Power of Performance of *DMCE* Statistics

Monte Carlo simulation method is used to investigate the power of performance of the *DMCEc* and *DMCEs* statistics. Six different sample sizes are considered: *n* = 20, 30, 50, 70, 100 and 130. Similar procedure employed in Section 5.3 is used here to generate the data set. Then, the observation at position *d*, say v_d, is contaminated as follows;

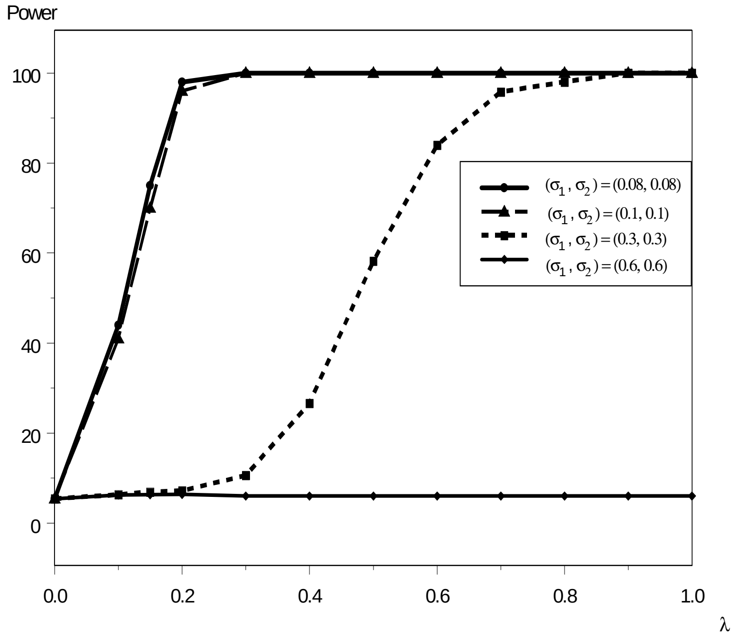
$$v_d^* = v_d + \lambda\pi \quad \mod(2\pi),$$

where v_d^* is the value after contamination and λ is the degree of contamination in the range $0 \leq \lambda \leq 1$. The generated data of ***U*** and ***V*** are then fitted by the JS circular regression model to give the parameter estimate $\hat{A}_0, \hat{A}_1, \hat{B}_1, \hat{C}_0, \hat{C}_1$ and $\hat{D}_1$. Then, the values of *MCEc* and *MCEs* are calculated using equation (5.1) and (5.3) respectively. Next, the values of *DMCEc* and *DMCEs* statistics are calculated for each generated data set. Consequently, we exclude the *j*th row from sample, for *j* = 1, …, *n* and fit the remaining data using equation (3.6). The $MCEc_{(-j)}$ and $MCEs_{(-j)}$ are then calculated by using equation (5.1) and (5.3), respectively. Finally, we specify the maximum value of *DMCEc* and *DMCEs* statistics. The process is repeated for 500 times. The power of performance of *DMCEc* and *DMCEs* statistics are examined by computing the percentage of correct detection of the contaminated observation at position *d*.

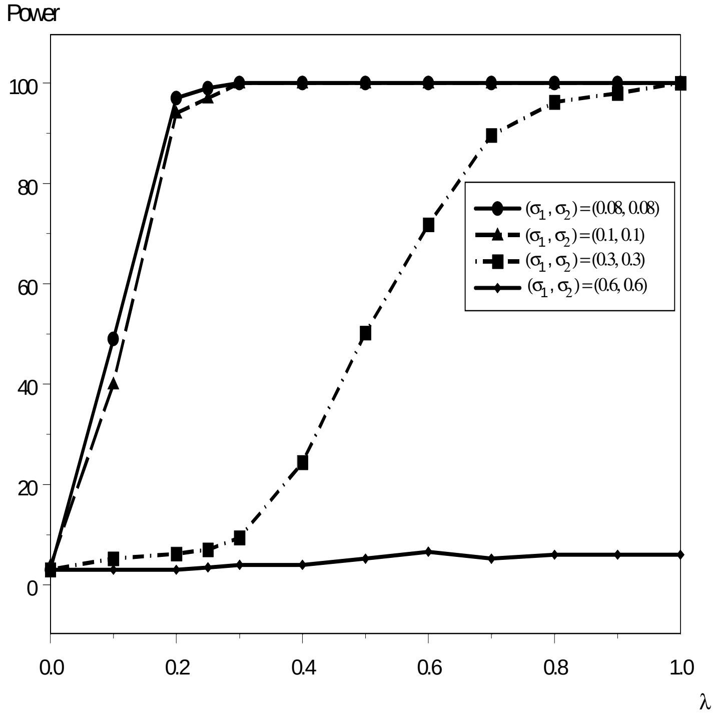
A part of simulation results are displayed in the Figures 5.1-5.3 and the other simulation results are given in Appendix 3. Figure 5.1 shows the performance of *DMCEc* and *DMCEs* statistics for *n*=70 and different values of (σ_1, σ_2). It is obvious that the performance of both statistics highly depend on the values of (σ_1, σ_2), where the power of performances is a decreasing function of the (σ_1, σ_2). In other word, the performance is increasing as σ_1 and σ_2 get smaller. This is expected as V_{1j} and V_{2j} in equation (3.6) will fluctuate closer to the horizontal axis when σ_1 and σ_2 are closer to zero, and hence, better chance to detect the outlier even when λ is small.

Figure 5.2 shows the performance of *DMCEc* and *DMCEs* statistics with fixed $\sigma_1 = \sigma_2 = 0.1$ and different sample size *n*. For both statistics the power of performances does not differ much when we vary sample size *n*, though we note that the performance only increase slightly when *n* is larger. On the other hand, Figure 5.3 shows that the performance of both tests does not differ much when (σ_1, σ_2) are small as can be seen from Figure 5.3(a). However, the difference in performance is clearer for larger (σ_1, σ_2) as shown in Figure 5.3(b). These results agree with that observed in Figure 5.1, that is when the value of (σ_1, σ_2) get smaller, the performance gets better.

Meanwhile, Figure 5.4 shows the performance of the *COVRATIO*, *DMCEc* and *DMCEs* statistics. The power of performance of *DMCEc* and *DMCEs* statistics does not differ much when (σ_1, σ_2) are small compared to *COVRATIO* statistic as can be seen from Figure 5.4(a). However, the difference in performance is clearer for larger (σ_1, σ_2) as shown in Figure 5.4 (b). From these graph, the *DMCEc* statistics is the best measures among the three because the performance of *DMCEc* is superior compared to the others.

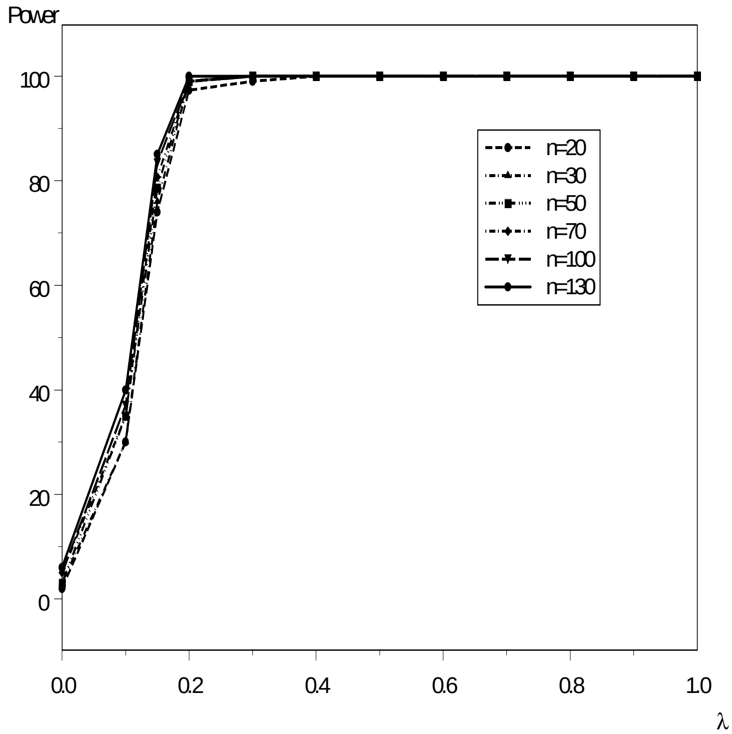


(a) *DMCEc*

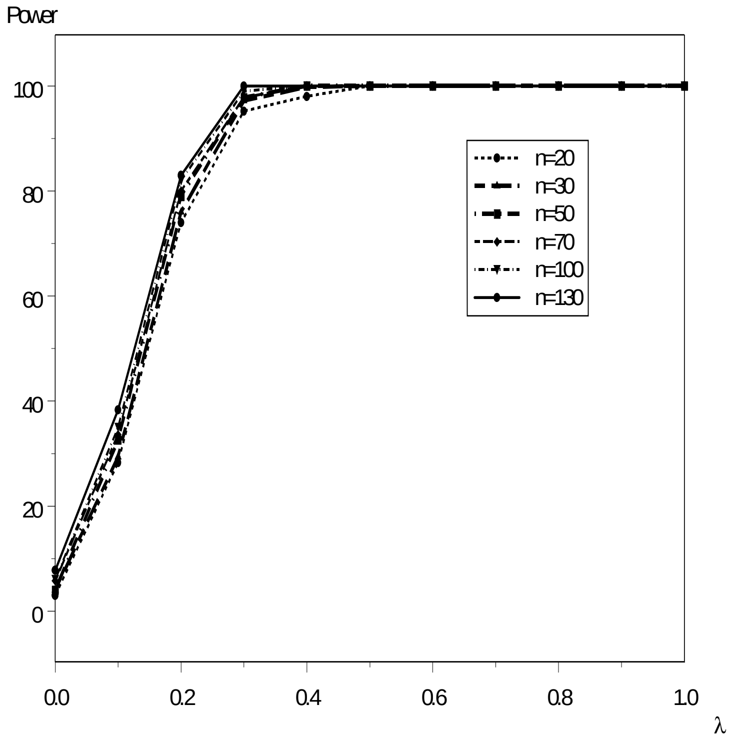


(b) *DMCEs*

Figure 5.1: Graph of power performance for *DMCEc* and *DMCEs* statistics, for *n*=70

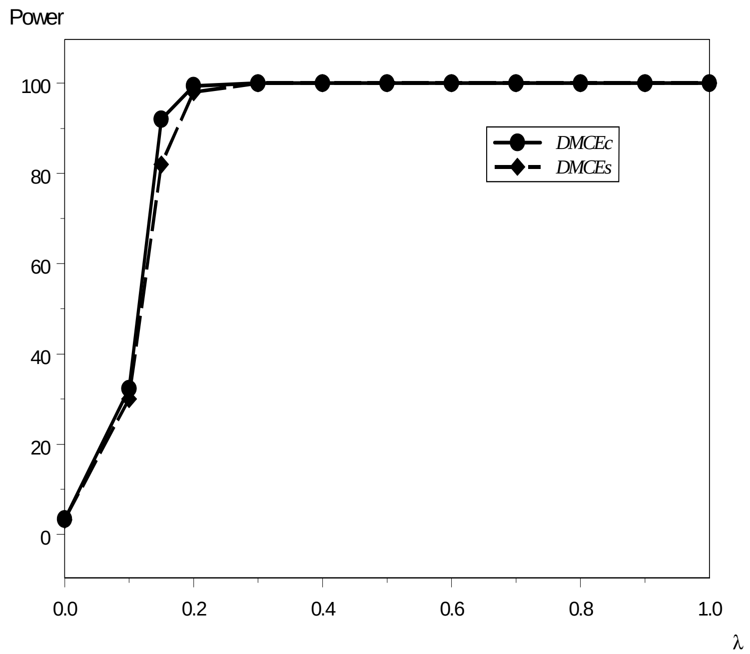


(a) *DMCEc*

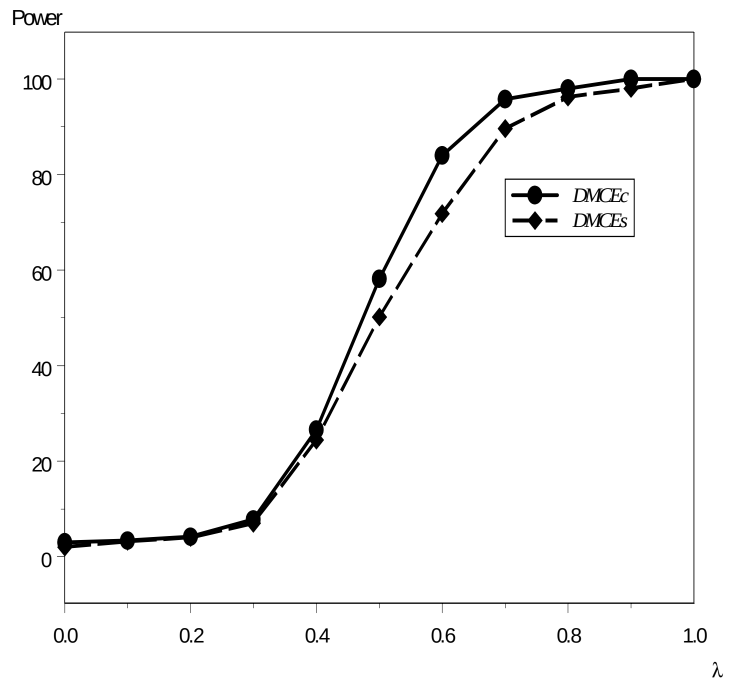


(b) *DMCEs*

Figure 5.2: Graph of power performance for *DMCEc* and *DMCEs* statistics, for $\sigma_1 = \sigma_2 = 0.1$

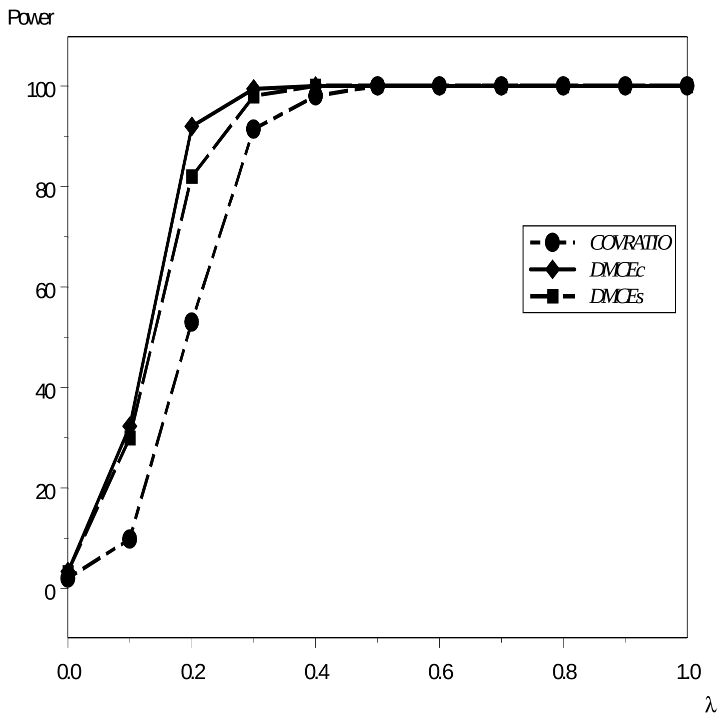


(a) n=100, $(\sigma_1, \sigma_2) = (0.1, 0.1)$

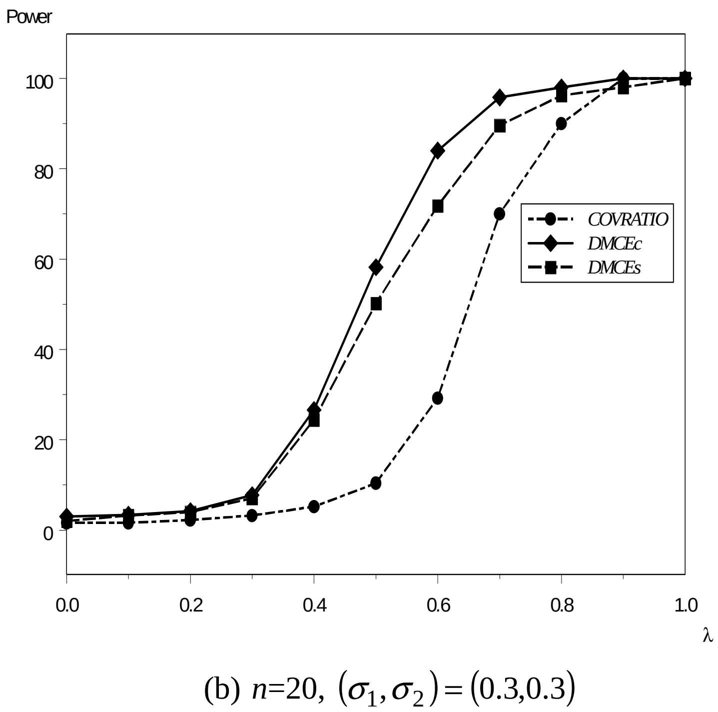


(b) n=70, $(\sigma_1, \sigma_2) = (0.3, 0.3)$

Figure 5.3: Power of performance of *DMCEc* and *DMCEs* statistics



(a) n=100, $(\sigma_1, \sigma_2) = (0.1, 0.1)$



(b) n=20, $(\sigma_1, \sigma_2) = (0.3, 0.3)$

Figure 5.4: Power of performance of *COVRATIO*, *DMCEc* and *DMCEs* statistics

5.5 Practical Example: Eye data

We consider the eye data which is described in Section 3.5.1. The least squares parameter estimates of the JS circular regression model are $\hat{A}_0 = 1.0822$, $\hat{A}_1 = -0.1497$, $\hat{B}_1 = -0.3837$, $\hat{C}_0 = 0.0986$, $\hat{C}_1 = 0.2534$, $\hat{D}_1 = 0.5935$, $\hat{\kappa} = 22.51$, $\hat{\sigma}_1 = 0.16$ and $\hat{\sigma}_2 = 0.16$ with the fitted observational regression-like model with respect to $\hat{g}_1(u)$ and $\hat{g}_2(u)$ are as follows:

$$\hat{g}_1(u) = 1.0822 - 0.1497 \cos u - 0.3837 \sin u$$

$$\hat{g}_2(u) = 0.0986 + 0.2534 \cos u + 0.5935 \sin u .$$

The diagnostic plots of the two set of resulting residuals from the fitted model suggest the possibility of occurrence of outliers in the data set. We now apply the outlier detection procedures based on *DMCEc* and *DMCEs* on the data.

5.5.1 *DMCE* Statistics

We now apply the *DMCEc* and *DMCEc* statistics on the eye data in order to detect any possible outliers in the data set. The mean circular errors for full data set based on the cosine and sine functions are $MCEc = 0.0225$ and $MCEs = 0.0859$ respectively. Then, the values of $|MCEc - MCEc_{(-j)}|$ and $|MCEs - MCEs_{(-j)}|$ for $j = 1, \ldots, n$ are calculated. The values are plotted in Figures 5.5-5.6.

The sample size of the eye data is 23 with the estimated values of σ_1 and σ_2 are 0.16 and 0.16 respectively. Then, by running the program as given in Appendix 6, we use the LS estimates for the eye data to find the appropriate the cut-off point at 0.05 level of significance which is found to be 0.0024 for *DMCEc* and 0.0082 for *DMCEs* .

As a result, it can be seen in Figure 5.5 that the *DMCEc* statistic values for observation number 2 and 15 exceed the cut-off point as shown by the dashed line. However, for *DMCEs* statistic, only observation number 15 exceeds the cut-off point as shown by the dashed line in Figure 5.5, though the corresponding value for observation number 2 is closer to the dashed line compared to the others. It is interesting to investigate the reason behind these results. Note that the sample size of the eye data is rather small, which only involve 23 eye patients. We have commented in Section 5.4 that the performance of the *DMCEc* statistic is better when *n* is small, and is almost similar for large sample size. Hence, the results on the eye data further confirmed the finding in the simulation study that the procedure based on *DMCEc* are superior than the procedure based on the *DMCEs*.

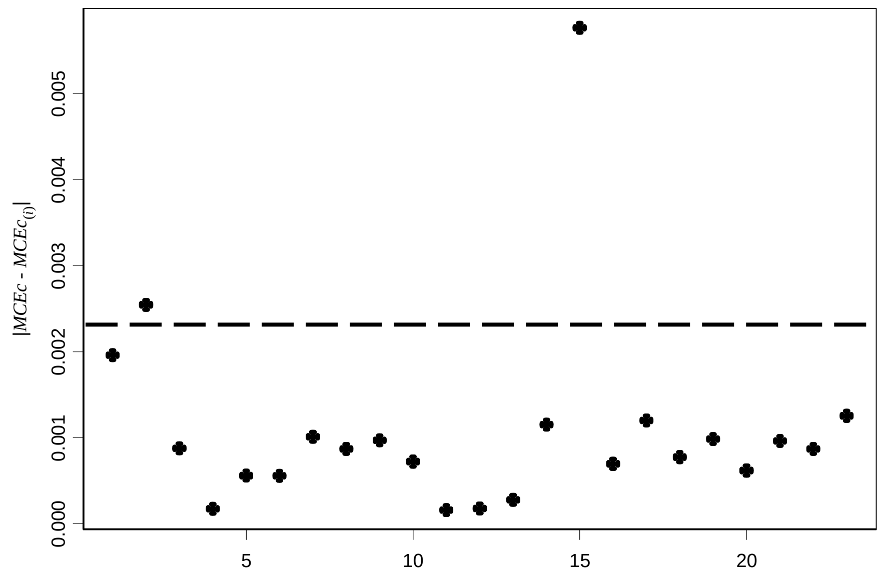


Figure 5.5: The values of the *DMCEc* statistic for eye data

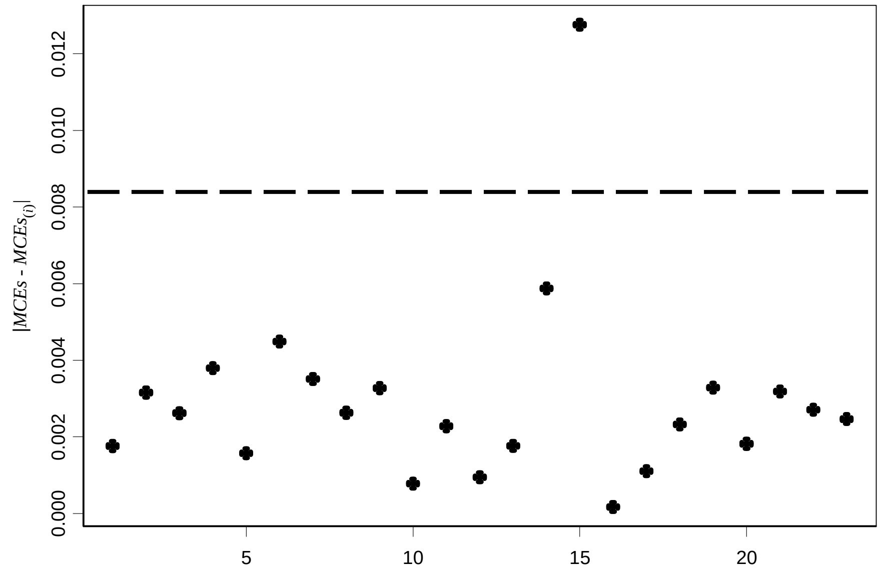
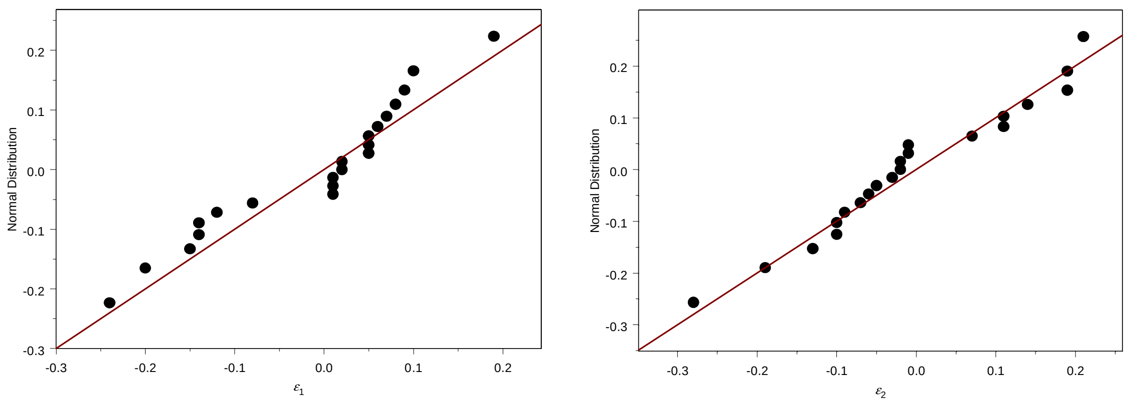


Figure 5.6: The values of the *DMCEs* statistic for eye data

Table 5.7: Parameter estimates for clean and contaminated data

Parameter estimates	**Contaminated data**	**Standard error**	**Clean data (case 2 and 15 deleted)**	**Standard error**
$\hat{A}_0$	1.0822	0.2664	1.0592	0.2002
$\hat{A}_1$	-0.1497	0.1026	-0.1857	0.2793
$\hat{B}_1$	-0.3836	0.2873	-0.3499	0.2173
$\hat{C}_0$	0.0986	0.2776	0.0855	0.2396
$\hat{C}_1$	0.2534	0.1070	0.3059	0.0950
$\hat{D}_1$	0.5935	0.2994	0.6106	0.2601
$\hat{\sigma}_1$	0.16	0.1198	0.12	0.1168
$\hat{\sigma}_2$	0.16	0.1635	0.14	0.1397
$A(\hat{\kappa})$	0.9775	-	0.9857	-
$\hat{\kappa}$	22.5056	-	35.3085	-
$\hat{\rho}$	0.9774	-	0.9857	-



(a) ε_1 (b) ε_2

Figure 5.7: Q-Q plot of residual for circular residuals without observations number 2 and 15

5.5.2 The Effect of Outliers on the Parameter Estimates

Table 5.7 summarizes the effect of excluding the outliers on the parameter estimates. The removal of observation numbers 2 and 15 significantly change the value of $\hat{A}_0$, $\hat{A}_1$, $\hat{B}_1$, $\hat{C}_0$, $\hat{C}_1$, $\hat{D}_1$, $\hat{\sigma}_1$ and $\hat{\sigma}_2$. Furthermore, the values of the standard errors for all the parameter estimates in clean data are smaller than contaminated data. Meanwhile, the estimated concentration parameter has increased from 0.9774 to 0.9857 and $A(\hat{\kappa})$ is increased from 0.9775 to 0.9857, as well as $\hat{\kappa}$ increased from 22.5056 to 35.3085. Thus, the estimation is more accurate and we may have better model fitting for the data when observation 2 and 15 are excluded from the data set. The Q-Q plots of circular residuals after removing observations number 2 and 15 from the eye data are shown in Figure 5.7. The points are now much closer to the straight line in the plot indicating a better fit for the data.

5.6 Summary

In this chapter, we have considered *DMCEc* and *DMCEs* statistics to identify possible outliers in the JS circular regression model. The cut-off points of the statistics are obtained and the power of performance is examined through extensive simulation study. The *DMCEc* statistic shows better performance than *DMCEs* for small sample size but becomes closer for larger sample size. The outlier detection procedures based on the *DMCEc* and *DMCEs* statistics are utilized to identify outliers in JS circular regression. Using *DMCEc* statistic, observations number 2 and 15 are identified as outliers when applied on the eye data. The exclusion of these two observations from the original data set improves the fitted JS circular regression model on the data.

CHAPTER SIX

GENERALIZED JS CIRCULAR REGRESSION MODEL

6.1 Introduction

The problem of regressing a circular response variable on more than one circular explanatory variable has not been explored. Among the circular regression models available in the literature, we foresee that the JS circular model has the flexibility to be extended to include more circular explanatory variables. In this chapter, we present the formulation of the generalized JS circular regression model and the estimation of the regression parameters using the least squares method. We then investigate the problem of multicollinearity in the model and provided the solution using ridge regression approach.

6.2 The Model

Let $(v, u_1, u_2, ..., u_p)$ be a set of random variables which are measured with reference to the same zero direction and the same sense of rotation. The angles can be treated as unit vectors in the plane in terms of their sine and cosine components. To predict v for a given value of $u_1, u_2, ..., u_p$, the vector corresponding to v is predicted by the conditional expectation (or regression) of e^{iv} given $\boldsymbol{u} = (u_1, u_2 ..., u_p)$, namely

$$E\left(e^{iv} \mid \boldsymbol{u}\right) = \rho(\boldsymbol{u}) e^{i\mu(\boldsymbol{u})}. \tag{6.1}$$

By letting $e^{iv} = \cos v + i \sin v$, we have

$$\begin{aligned} E(\cos v + i \sin v \mid \boldsymbol{u}) &= \rho(\boldsymbol{u}) \cos \mu(\boldsymbol{u}) + i\rho(\boldsymbol{u}) \sin \mu(\boldsymbol{u}) \\ &= g_1(\boldsymbol{u}) + i\, g_2(\boldsymbol{u}) \end{aligned} \tag{6.2}$$

or equivalently, we may write

$$E(\cos v \mid \boldsymbol{u}) = g_1(\boldsymbol{u})$$

$$E(\sin v \mid \boldsymbol{u}) = g_2(\boldsymbol{u}). \tag{6.3}$$

Then, we may estimate the parameters $\mu(\boldsymbol{u})$ and $\rho(\boldsymbol{u})$ such that

$$\mu(\boldsymbol{u}) = \hat{\boldsymbol{v}} = \begin{cases} \arctan \dfrac{g_2(\boldsymbol{u})}{g_1(\boldsymbol{u})} & \text{if} \quad g_1(\boldsymbol{u}) \geq 0 \\ \pi + \arctan \dfrac{g_2(\boldsymbol{u})}{g_1(\boldsymbol{u})} & \text{if} \quad g_1(\boldsymbol{u}) \leq 0 \\ \text{undefined} & \text{if} \quad g_1(\boldsymbol{u}) = g_2(\boldsymbol{u}) = 0, \end{cases} \tag{6.4}$$

where $\mu(\boldsymbol{u})$ represents the conditional mean direction of v given $\boldsymbol{u}$ and $\rho(\boldsymbol{u})$ is the conditional concentration parameter, $0 \leq \rho(\boldsymbol{u}) \leq 1$.

Motivated by the approximation made for JS model with one independent variable, we may approximate $g_i(\boldsymbol{u})$, i=1,2 using trigonometric polynomial of functions of p variables $u_1, u_2, \ldots, u_p$. Kufner & Kadlec (1971) presented the functions for 2 variables in detail, noting that the theory can be extended to include more variables. Now, the trigonometric polynomials will have cross-products terms involving the cosine and sine functions. In general, the number of cross-product terms is given by the formula $(m+1)\times 2^{p}$. In our present case with p=2, there are $4(m+1)$ terms forming the following system of functions:

$$\cos ku_1 \cos \ell u_2 \,; \quad \cos ku_1 \sin \ell u_2 \,;$$

$$\sin ku_1 \cos \ell u_2 \,; \quad \sin ku_1 \sin \ell u_2 \,; \quad k, \ell = 0,1,2,\ldots m\,.$$

Hence, the trigonometric polynomials of function of two variables for a suitable degree m to approximate (6.3) are given by

$$g_1(u) \approx \sum_{k,\ell=0}^{m} \vartheta_{k\ell} \begin{bmatrix} A_{k\ell} \cos ku_1 \cos \ell u_2 + B_{k\ell} \cos ku_1 \sin \ell u_2 \\ + C_{k\ell} \sin ku_1 \cos \ell u_2 + D_{k\ell} \sin ku_1 \sin \ell u_2 \end{bmatrix}$$

$$g_2(u) \approx \sum_{k,\ell=0}^{m} \vartheta_{k\ell} \begin{bmatrix} E_{k\ell} \cos ku_1 \cos \ell u_2 + F_{k\ell} \cos ku_1 \sin \ell u_2 \\ + G_{k\ell} \sin ku_1 \cos \ell u_2 + H_{k\ell} \sin ku_1 \sin \ell u_2 \end{bmatrix}, \tag{6.5}$$

where

$$A_{k\ell} = \frac{1}{\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} f(u_1, u_2) \cos ku_1 \cos \ell u_2 \, du_1 \, du_2;$$

$$B_{k\ell} = \frac{1}{\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} f(u_1, u_2) \cos ku_1 \sin \ell u_2 \, du_1 \, du_2;$$

$$C_{k\ell} = \frac{1}{\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} f(u_1, u_2) \sin ku_1 \cos \ell u_2 \, du_1 \, du_2;$$

$$D_{k\ell} = \frac{1}{\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} f(u_1, u_2) \sin ku_1 \sin \ell u_2 \, du_1 \, du_2$$

(k, ℓ = 0,1,2,..., m) and

$$\vartheta_{k\ell} = \begin{cases} \dfrac{1}{4} & \text{for} \quad k = \ell = 0 \\ \dfrac{1}{2} & \text{for} \quad k > 0, \ell = 0 \quad \text{and for} \quad k = 0,\ \ell > 0 \\ 1 & \text{for} \quad k > 0, \ell > 0 \end{cases} .$$

Hence, we have the following two observational regression-like models:

$$\boldsymbol{V}_{1j} = \cos v_j = \sum_{k,l=0}^{m} \begin{pmatrix} A_{k\ell} \cos ku_1 \cos \ell u_2 + B_{k\ell} \cos ku_1 \sin \ell u_2 \\ + C_{k\ell} \sin ku_1 \cos \ell u_2 + D_{k\ell} \sin ku_1 \sin \ell u_2 \end{pmatrix} + \varepsilon_{1j}$$

$$\boldsymbol{V}_{2j} = \sin v_j = \sum_{k,l=0}^{m} \begin{pmatrix} E_{k\ell} \cos ku_1 \cos \ell u_2 + F_{k\ell} \cos ku_1 \sin \ell u_2 \\ + G_{k\ell} \sin ku_1 \cos \ell u_2 + H_{k\ell} \sin ku_1 \sin \ell u_2 \end{pmatrix} + \varepsilon_{2j} \tag{6.6}$$

for j = 1, … , n and $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2)$ is the random error vector following normal distribution with mean $\boldsymbol{0}$ and dispersion matrix $\boldsymbol{\Sigma}$, unknown.

For the case with three independent variables u_1, u_2, u_3, we have $8(m+1)$ cross-product terms in $g_i(\boldsymbol{u})$ for each level of order m giving the following system:

$\cos ku_1 \cos \ell u_2 \cos tu_3$; $\cos ku_1 \cos \ell u_2 \sin tu_3$; $\cos ku_1 \sin \ell u_2 \cos tu_3$;
$\cos ku_1 \sin \ell u_2 \sin tu_3$; $\sin ku_1 \cos \ell u_2 \sin tu_3$; $\sin ku_1 \cos \ell u_2 \cos tu_3$;
$\sin ku_1 \sin \ell u_2 \cos tu_3$; $\sin ku_1 \sin \ell u_2 \sin tu_3$.

Hence, the trigonometric polynomials of function of three variables are

$$g_1(\boldsymbol{u}) \approx \sum_{k,\ell,t=0}^{m} \begin{pmatrix} A_{k\ell t} \cos ku_{1j} \cos \ell u_{2j} \cos tu_{3j} + B_{k\ell t} \cos ku_{1j} \cos \ell u_{2j} \sin tu_{3j} \\ + C_{k\ell t} \cos ku_{1j} \sin \ell u_{2j} \cos tu_{3j} + D_{k\ell t} \cos ku_{1j} \sin \ell u_{2j} \sin tu_{3j} \\ + E_{k\ell t} \sin ku_{1j} \cos \ell u_{2j} \sin tu_{3j} + F_{k\ell t} \sin ku_{1j} \cos \ell u_{2j} \cos tu_{3j} \\ + G_{k\ell t} \sin ku_{1j} \sin \ell u_{2j} \cos tu_{3j} + H_{k\ell t} \sin ku_{1j} \sin \ell u_{2j} \sin tu_{3j} \end{pmatrix}$$

$$g_2(\boldsymbol{u}) \approx \sum_{k,\ell,t=0}^{m} \begin{pmatrix} I_{k\ell t} \cos ku_{1j} \cos \ell u_{2j} \cos tu_{3j} + J_{k\ell t} \cos ku_{1j} \cos \ell u_{2j} \sin tu_{3j} \\ + K_{k\ell t} \cos ku_{1j} \sin \ell u_{2j} \cos tu_{3j} + L_{k\ell t} \cos ku_{1j} \sin \ell u_{2j} \sin tu_{3j} \\ + M_{k\ell t} \sin ku_{1j} \cos \ell u_{2j} \sin tu_{3j} + N_{k\ell t} \sin ku_{1j} \cos \ell u_{2j} \cos tu_{3j} \\ + P_{k\ell t} \sin ku_{1j} \sin \ell u_{2j} \cos tu_{3j} + Q_{k\ell t} \sin ku_{1j} \sin \ell u_{2j} \sin tu_{3j} \end{pmatrix}. \quad (6.7)$$

Hence, we have the following observational regression-like models:

$$\boldsymbol{V}_{1j} = \cos v_j = \sum_{k,\ell,t=0}^{m} \begin{pmatrix} A_{k\ell t} \cos ku_{1j} \cos \ell u_{2j} \cos tu_{3j} + B_{k\ell t} \cos ku_{1j} \cos \ell u_{2j} \sin tu_{3j} \\ + C_{k\ell t} \cos ku_{1j} \sin \ell u_{2j} \cos tu_{3j} + D_{k\ell t} \cos ku_{1j} \sin \ell u_{2j} \sin tu_{3j} \\ + E_{k\ell t} \sin ku_{1j} \cos \ell u_{2j} \sin tu_{3j} + F_{k\ell t} \sin ku_{1j} \cos \ell u_{2j} \cos tu_{3j} \\ + G_{k\ell t} \sin ku_{1j} \sin \ell u_{2j} \cos tu_{3j} + H_{k\ell t} \sin ku_{1j} \sin \ell u_{2j} \sin tu_{3j} \end{pmatrix} + \varepsilon_{1j}$$

$$\boldsymbol{V}_{2j} = \sin v_j = \sum_{k,\ell,t=0}^{m} \begin{pmatrix} I_{k\ell t} \cos ku_{1j} \cos \ell u_{2j} \cos tu_{3j} + J_{k\ell t} \cos ku_{1j} \cos \ell u_{2j} \sin tu_{3j} \\ + K_{k\ell t} \cos ku_{1j} \sin \ell u_{2j} \cos tu_{3j} + L_{k\ell t} \cos ku_{1j} \sin \ell u_{2j} \sin tu_{3j} \\ + M_{k\ell t} \sin ku_{1j} \cos \ell u_{2j} \sin tu_{3j} + N_{k\ell t} \sin ku_{1j} \cos \ell u_{2j} \cos tu_{3j} \\ + P_{k\ell t} \sin ku_{1j} \sin \ell u_{2j} \cos tu_{3j} + Q_{k\ell t} \sin ku_{1j} \sin \ell u_{2j} \sin tu_{3j} \end{pmatrix} + \varepsilon_{2j}$$

(6.8)

for j = 1, … , n. Overall, we have 2(8m+1) regression coefficients to estimate after assuming some coefficients to equal zero for identifiability purposes. The idea described above can be extended for more than 3 independents variables.

6.3 Least Squares Estimation Method

We now consider the least squares estimation method for estimating the parameters of the generalized JS circular regression model. By following the notation in Section 3.3.1, the observational regression-like model for any number of p independent variables, equation (6.6) for p=2 and equation (6.8) for p=3, can be summarized as

$$\boldsymbol{V}^{(1)} = \boldsymbol{U}\boldsymbol{\lambda}^{(1)} + \boldsymbol{\varepsilon}^{(1)}$$

$$\boldsymbol{V}^{(2)} = \boldsymbol{U}\boldsymbol{\lambda}^{(2)} + \boldsymbol{\varepsilon}^{(2)}, \tag{6.9}$$

where

$$\begin{aligned} \boldsymbol{V}^{(1)} &= \left(V_{11},\dots,V_{1n}\right)' \\ \boldsymbol{V}^{(2)} &= \left(V_{21},\dots,V_{2n}\right)' \\ \boldsymbol{\varepsilon}^{(1)} &= \left(\varepsilon_{11},\dots,\varepsilon_{1n}\right)' \\ \boldsymbol{\varepsilon}^{(2)} &= \left(\varepsilon_{21},\dots,\varepsilon_{2n}\right)' \end{aligned} \tag{6.10}$$

with the matrix $\boldsymbol{U}$ is the combination of cosine and sine functions

$$\underset{n\times(2^p\times m+1)}{\boldsymbol{U}} = \begin{bmatrix} \underset{n\times 1}{1} & \underset{n\times m}{\boldsymbol{CC}} & \underset{n\times m}{\boldsymbol{CS}} & \underset{n\times m}{\boldsymbol{SC}} & \underset{n\times m}{\boldsymbol{SS}} \end{bmatrix} \tag{6.11}$$

such that

$$\boldsymbol{CC} = \begin{bmatrix} \cos u_{11}\cos u_{21} & \cdots & \cos mu_{11}\cos mu_{21} \\ \vdots & & \vdots \\ \cos u_{1n}\cos u_{2n} & \cdots & \cos mu_{1n}\cos mu_{2n} \end{bmatrix},$$

$$\boldsymbol{CS} = \begin{bmatrix} \cos u_{11}\sin u_{21} & \cdots & \cos mu_{11}\sin mu_{21} \\ \vdots & & \vdots \\ \cos u_{1n}\sin u_{2n} & \cdots & \cos mu_{1n}\sin mu_{2n} \end{bmatrix},$$

$$\boldsymbol{SC} = \begin{bmatrix} \sin u_{11}\cos u_{21} & \cdots & \sin mu_{11}\cos mu_{21} \\ \vdots & & \vdots \\ \sin u_{1n}\cos u_{2n} & \cdots & \sin mu_{1n}\cos mu_{2n} \end{bmatrix},$$

and

$$\boldsymbol{SS} = \begin{bmatrix} \sin u_{11}\sin u_{21} & \cdots & \sin mu_{11}\sin mu_{21} \\ \vdots & & \vdots \\ \sin u_{1n}\sin u_{2n} & \cdots & \sin mu_{1n}\sin mu_{2n} \end{bmatrix}.$$

The parameters to be estimated are

$$\boldsymbol{\lambda}^{(1)} = \left(A_0, A_1,\dots,A_m, B_1,\dots,B_m, C_1,\dots,C_m, D_1,\dots,D_m\right)'$$

$$\boldsymbol{\lambda}^{(2)} = \left(E_0, E_1,\dots,E_m, F_1,\dots,F_m, G_1,\dots,G_m, H_1,\dots,H_m\right)'. \tag{6.12}$$

The least squares estimates turn out to be

$$\hat{\boldsymbol{\lambda}}^{(1)} = (\boldsymbol{U}'\boldsymbol{U})^{-1}\boldsymbol{U}'\boldsymbol{V}^{(1)}$$

$$\hat{\boldsymbol{\lambda}}^{(2)} = (\boldsymbol{U}'\boldsymbol{U})^{-1}\boldsymbol{U}'\boldsymbol{V}^{(2)}. \tag{6.13}$$

Then, the covariance matrix of the residuals, $\boldsymbol{\Sigma}$, takes the same form as described in Chapter 3. That is, by letting

$$R_0(p,q) = \boldsymbol{V}^{(p)\prime}\boldsymbol{V}^{(q)} - \boldsymbol{V}^{(p)\prime}\boldsymbol{U}(\boldsymbol{U}'\boldsymbol{U})^{-1}\boldsymbol{U}'\boldsymbol{V}^{(q)}$$

$$= \boldsymbol{V}^{(p)\prime}(\mathrm{I}-\mathrm{M})\boldsymbol{V}^{(q)},$$

where $\mathrm{M} = \boldsymbol{U}(\boldsymbol{U}'\boldsymbol{U})^{-1}\boldsymbol{U}'$ and $R_0 = (R_0(p,q))_{p,q=1,2}$, we have

$$\hat{\boldsymbol{\Sigma}} = \left[n - 2(2^{\boldsymbol{p}}m + 1)\right]^{-1} R_0 \tag{6.14}$$

which is an unbiased estimate of $\boldsymbol{\Sigma}$. Henceforth, the standard errors of the estimators can then be obtained as well as the $\rho(\boldsymbol{u})$ which is simpler to be expressed as follows:

$$\rho(\boldsymbol{u}) = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left[\boldsymbol{V}_{1j}^{2}(\boldsymbol{u}) + \boldsymbol{V}_{2j}^{2}(\boldsymbol{u})\right]} \tag{6.15}$$

since $0 \le \rho(\boldsymbol{u}) \le 1$, for any given $\boldsymbol{u}$.

6.4 Performance of the LS Method

A simulation study was carried out to investigate the performance of the LS estimation method for m=1 and p=2. Thus, the coefficients to be estimated are

$$\boldsymbol{\lambda} = \left(\boldsymbol{\lambda}^{(1)}, \boldsymbol{\lambda}^{(2)}\right) = (\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6, \lambda_7, \lambda_8, \lambda_9, \lambda_{10})$$

$$= (A_0, A_1, B_1, C_1, D_1, E_0, E_1, F_1, G_1, H_1). \tag{6.16}$$

We consider the set of uncorrelated random errors $(\varepsilon_1, \varepsilon_2)$ from the bivariate Normal distribution with mean vector $(\boldsymbol{0},\boldsymbol{0})$ and variances $(\sigma_1, \sigma_2) = (0.3, 0.3)$, respectively. For simplicity, we set the true values of A_0 and E_0 of the generalized JS multiple circular regression models of order m=1 to be zero, while $A_1, B_1, C_1, D_1, E_1, F_1, G_1$ and H_1 are obtained by using the standard additive trigonometric polynomial equations $\cos(a+u_1+u_2)$ and $\sin(a+u_1+u_2)$. Then, we expand these functions using standard additive trigonometric function as employed in the previous two chapters. For example, when a = 2, we have the true values of $A_1, B_1, C_1, D_1, E_1, F_1, G_1$ and H_1 to be 0.4161, -0.9093, -0.9093, 0.4161, 0.9093, -0.4161, -0.4161 and -0.9093 respectively. Similarly, we can also get different sets of true values by choosing different values of a. The full steps to investigate the performance of the LS method are as follows:

(i) Generate fixed $\boldsymbol{U_1}$ and $\boldsymbol{U_2}$ variables of size n from $VM(\pi,4)$ and $VM(\pi,2)$ respectively.

(ii) Generate $\boldsymbol{\varepsilon}$ of size n from $\boldsymbol{N}_2\left(\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}0.3 & 0\\0 & 0.3\end{pmatrix}\right)$. For a fixed a, obtain the true values of $A_0, A_1, B_1, C_1, D_1, E_0, E_1, F_1, G_1$ and H_1. We let the true values of A_0 and C_0 to be zero. Then, we calculate V_{1j} and V_{2j}, $j=1,\ldots,n$ using equation (6.6) .

(iii) Obtain the circular variable $v_j = \arctan\left(\frac{V_{2j}}{V_{1j}}\right)$, $j=1,\ldots,n$.

(iv) Fit the generated data using the generalized JS circular regression model to give the vector of parameter estimates $\hat{\boldsymbol{\lambda}}$ and concentration parameter estimate $\hat{\rho}(\boldsymbol{u})$.

(v) Finally, steps (i) – (iv) are repeated for *simu*=1000 times. For each parameter λ_i, i = 1, 2, …, 10, we calculate the mean, bias, standard error (*SE*) and root mean squared error (*RMSE*) using the same formula as in Section 3.4.1.

The results are tabulated in Table 6.1- Table 6.5 for each values of a = -3, -2, 2, 3 and 6. The following trends are observed:

1. The estimated mean for all parameter estimates are consistently close to the true values.
2. The bias is consistently small for all parameter estimates.
3. The *SE* for all parameter estimates are generally small.
4. The values for *RMSE* of each parameter estimates are small.

By looking the above results, the least squares estimation method performs well in estimating the parameters of the generalized JS circular regression models.

Table 6.1: Parameter estimates for a = -3

Estimates	true values	mean	*SE*	bias	*RMSE*
$\hat{A}_0$	0.0000	-0.0044	0.0689	-0.0044	0.1385
$\hat{A}_1$	-0.9900	-0.9227	0.0863	0.0673	2.1269
$\hat{B}_1$	0.1411	0.1301	0.0672	-0.0110	0.3478
$\hat{C}_1$	0.1411	0.1304	0.0817	-0.0108	0.3405
$\hat{D}_1$	0.9900	0.9189	0.1235	-0.0711	2.2488
$\hat{E}_0$	0.0000	0.0007	0.0576	0.0007	0.0230
$\hat{E}_1$	-0.1411	-0.1360	0.0918	0.0052	0.1633
$\hat{F}_1$	-0.9900	-0.9377	0.0565	0.0523	1.6548
$\hat{G}_1$	-0.9900	-0.9412	0.0853	0.0488	1.5418
$\hat{H}_1$	0.1411	0.1420	0.1211	0.0009	0.0276

Table 6.2: Parameter estimates for a = -2

Estimates	true values	mean	*SE*	bias	*RMSE*
$\hat{A}_0$	0.0000	0.0087	0.0612	0.0087	0.2749
$\hat{A}_1$	-0.4161	-0.4036	0.0904	0.0125	0.3956
$\hat{B}_1$	0.9093	0.8543	0.0601	-0.0550	1.7397
$\hat{C}_1$	0.9093	0.8658	0.0909	-0.0435	1.3741
$\hat{D}_1$	0.4161	0.4110	0.1222	-0.0052	0.1631
$\hat{E}_0$	0.0000	-0.0024	0.0669	-0.0024	0.0759
$\hat{E}_1$	-0.9093	-0.8511	0.0861	0.0582	1.8407
$\hat{F}_1$	-0.4161	-0.3828	0.0663	0.0334	1.0548
$\hat{G}_1$	-0.4161	-0.3831	0.0836	0.0330	1.0438
$\hat{H}_1$	0.9093	0.8542	0.1252	-0.0551	1.7423

Table 6.3: Parameter estimates for a = 2

Estimates	true values	mean	*SE*	bias	*RMSE*
$\hat{A}_0$	0.0000	0.0072	0.0603	0.0072	0.2277
$\hat{A}_1$	-0.4161	-0.4043	0.0922	0.0119	0.3759
$\hat{B}_1$	-0.9093	-0.8604	0.0619	0.0489	1.5453
$\hat{C}_1$	-0.9093	-0.8556	0.0910	0.0537	1.6967
$\hat{D}_1$	0.4161	0.4075	0.1194	-0.0086	0.2727
$\hat{E}_0$	0.0000	0.0034	0.0677	0.0034	0.1068
$\hat{E}_1$	0.9093	0.8512	0.0863	-0.0581	1.8370
$\hat{F}_1$	-0.4161	-0.3834	0.0678	0.0328	1.0364
$\hat{G}_1$	-0.4161	-0.3764	0.0859	0.0397	1.2555
$\hat{H}_1$	-0.9093	-0.8526	0.1226	0.0567	1.7922

Table 6.4: Parameter estimates for a = 3

Estimates	true values	mean	*SE*	bias	*RMSE*
$\hat{A}_0$	0.0000	-0.0055	0.0670	-0.0055	0.1727
$\hat{A}_1$	-0.9900	-0.9221	0.0843	0.0679	2.1474
$\hat{B}_1$	-0.1411	-0.1303	0.0673	0.0108	0.3418
$\hat{C}_1$	-0.1411	-0.1324	0.0835	0.0087	0.2761
$\hat{D}_1$	0.9900	0.9185	0.1271	-0.0715	2.2601
$\hat{E}_0$	0.0000	-0.0044	0.0580	-0.0044	0.1385
$\hat{E}_1$	0.1411	0.1392	0.0922	-0.0019	0.0609
$\hat{F}_1$	-0.9900	-0.9378	0.0561	0.0521	1.6490
$\hat{G}_1$	-0.9900	-0.9405	0.0864	0.0495	1.5647
$\hat{H}_1$	-0.1411	-0.1333	0.1212	0.0078	0.2473

Table 6.5: Parameter estimates for a = 6

Estimates	true values	mean	*SE*	bias	*RMSE*
$\hat{A}_0$	0.0000	0.0057	0.0661	0.0057	0.1798
$\hat{A}_1$	0.9602	0.8945	0.0828	-0.0656	2.0756
$\hat{B}_1$	0.2794	0.2572	0.0632	-0.0223	0.7041
$\hat{C}_1$	0.2794	0.2543	0.0878	-0.0251	0.7945
$\hat{D}_1$	-0.9602	-0.8941	0.1224	0.0661	2.0895
$\hat{E}_0$	0.0000	0.0048	0.0606	0.0048	0.1523
$\hat{E}_1$	-0.2794	-0.2707	0.0961	0.0088	0.2772
$\hat{F}_1$	0.9602	0.9115	0.0610	-0.0487	1.5386
$\hat{G}_1$	0.9602	0.9049	0.0896	-0.0553	1.7477
$\hat{H}_1$	0.2794	0.2736	0.1219	-0.0058	0.1842

6.5 Problem of Multicollinearity

In Sections 2.6 and 2.7, we have discussed the multicollinearity problem in multiple linear regressions. One of the methods available to handle the problem in multiple linear regressions is by considering the ridge regression modelling as an alternative procedure to the LS method, see for example, Hoerl & Kennard (1970). This is carried out by adding a constants k to the diagonal of $(\boldsymbol{X}'\boldsymbol{X})$ matrix before computing the $\hat{\boldsymbol{\beta}}$ using $\hat{\boldsymbol{\beta}}^{*} = (\boldsymbol{X}'\boldsymbol{X} + \boldsymbol{kI})^{-1}\boldsymbol{X}'\boldsymbol{Y}$ as given by equation (2.12). Note that if k = 0, the ridge regression estimator reduces to the LS estimator. In the presence of multicollinearity, the ridge regression estimator is much more stable with smaller variance than that from LS estimator.

Now, we look at the problem of multicollinearity in the generalized JS circular regression model which, as we have defined earlier, involves the study on the relationship between one dependent circular variable v and few explanatory circular variables u_i, i = 1, …, p. The problem happens when at least two of the circular variables u_i are correlated. We expect the problem will give adverse effects on estimated coefficients in generalized JS circular regression model and will be illustrated later by simulated examples. Hence, proper examination on the data set should be carried out to identify the existence of multicollinearity problem in the data.

As described in Section 2.6.2, for multiple linear regression model, the *VIF* is calculated using the formula

$$VIF = \frac{1}{1 - R_j^2}$$

where R_j^2 is the coefficient of determination obtained by regressing an independent linear variable, say X_j, on other independent linear variables. Here, we will also be using the variance inflation factor to identify the problem of multicollinearity in the generelized JS circular regression model, but with modified procedure. Note that from Section 6.2, by regressing a circular variable U_j on other variables $\boldsymbol{U}_{(j)} = (U_1, \ldots, U_{j-1}, U_{j+1}, \ldots, U_p)$ such that

$$E\left(e^{i\boldsymbol{u}_j} \mid \boldsymbol{u}_{(j)}\right) = \rho(\boldsymbol{u}(j))e^{i\mu(\boldsymbol{u}_{(j)})}$$

results in two observational regression-like models $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$ of the form given by equations (6.6) and (6.8) for the case $p = 2$ and $p = 3$ respectively. As such, we will have two *VIF* values corresponding to each of these observational regression-like models. That is, for the case $p = 2$, we have a total of four *VIF* values corresponding to $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$, while, for $p = 3$, we have six *VIF* values. In general, the procedure to calculate the *VIF* in our present case can be described below:

(1) For each U_j, $j = 1, \ldots, p$, regress circular variable U_j on other variables $\boldsymbol{U}_{(j)} = (U_1, \ldots, U_{j-1}, U_{j+1}, \ldots, U_p)$. From the observational regression-like models $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$, calculate the coefficient of determination $R^2_{V_1, j}$ and $R^2_{V_2, j}$ respectively.

(2) Calculate the corresponding *VIF* using the formula

$$VIF_{V_i, j} = \frac{1}{1 - R^2_{V_i, j}}, \quad i = 1,2 \text{ and } j = 1, \ldots, p. \tag{6.17}$$

Using similar rule as in the linear case, if at least one of the $VIF_{V_i, j}$ is greater than or equal to 5 or 10, we have multicollinearity problem to be taken care of.

One way to overcome the problem is by employing the ridge regression technique on the observational regression-like models. With a suitable choice of k, we should get better estimates for the generalized JS circular regression model than the LS estimates by controlling the general instability associated with the LS estimates. We shall describe the ridge regression method in the next section.

6.5.1 Circular Ridge Regression for Generalized JS Circular Regression Model

Here, we use the same notation as equation (2.12) to estimate the parameters of the generalized JS circular regression model in the presence of multicollinearity. The least square estimation of the parameter based on the properties of the best linear unbiased estimation has been discussed in Section 6.3 and is given by

$$\hat{\boldsymbol{\lambda}}^{(j)} = \left(\boldsymbol{U}'\boldsymbol{U}\right)^{-1}\boldsymbol{U}'\boldsymbol{V}^{(j)} \tag{6.18}$$

with the minimum sum of squares of the residuals is given by

$$\phi\left(\hat{\boldsymbol{\lambda}}^{(j)}\right) = \left(\boldsymbol{V}^{(j)} - \boldsymbol{U}\,\hat{\boldsymbol{\lambda}}^{(j)}\right)'\left(\boldsymbol{V}^{(j)} - \boldsymbol{U}\,\hat{\boldsymbol{\lambda}}^{(j)}\right) \tag{6.19}$$

for j=1, 2. We have presented the background of the "ridge circular regression" for multiple linear regression models in Section 2.7. Using similar notation and motivation, the ridge circular regression model attempts to control the instability of the LS estimates of the generalized JS circular regression model by introducing a constant k in equation (6.18) giving

$$\begin{aligned}\hat{\boldsymbol{\lambda}}^{(j)^*} &= \left(\boldsymbol{U}'\boldsymbol{U} + k\boldsymbol{I}\right)^{-1}\boldsymbol{U}'\boldsymbol{V}^{(j)};\ k \ge 0 \\ &= \boldsymbol{W}\boldsymbol{U}'\boldsymbol{V}^{(j)}\end{aligned} \tag{6.20}$$

where $\boldsymbol{W} = \left(\boldsymbol{U}'\boldsymbol{U} + k\boldsymbol{I}\right)^{-1}$ for j=1, 2. The relationship of a ridge circular estimate to an ordinary LS estimate is given by the alternative form

$$\hat{\boldsymbol{\lambda}}^{(j)^*} = \left[\boldsymbol{I}_{\boldsymbol{p}} + k\left(\boldsymbol{U}'\boldsymbol{U}\right)^{-1}\right]^{-1} \hat{\boldsymbol{\lambda}}^{(j)}$$

$$= \boldsymbol{Z}\,\hat{\boldsymbol{\lambda}}^{(j)}, \qquad (6.21)$$

where $\boldsymbol{Z} = \left[\boldsymbol{I}_{\boldsymbol{p}} + k\left(\boldsymbol{U}'\boldsymbol{U}\right)^{-1}\right]^{-1}$. For an estimate $\hat{\boldsymbol{\lambda}}^{(j)}$, the residual sum of squares is

$$\phi^*(k) = \left(\boldsymbol{V}^{(j)} - \boldsymbol{U}\hat{\boldsymbol{\lambda}}^{(j)}\right)'\left(\boldsymbol{V}^{(j)} - \boldsymbol{U}\hat{\boldsymbol{\lambda}}^{(j)}\right) \qquad (6.22)$$

which can be written in the form

$$\phi^*(k) = \boldsymbol{V}^{(j)'}\boldsymbol{V}^{(j)} - \left(\hat{\boldsymbol{\lambda}}^{(j)^*}\right)'\boldsymbol{U}'\boldsymbol{V}^{(j)} - k\left(\hat{\boldsymbol{\lambda}}^{(j)^*}\right)'\left(\hat{\boldsymbol{\lambda}}^{(j)^*}\right). \qquad (6.23)$$

Further, by following Hoerl & Kennard (1970), we will obtain the ridge traces for the two observational regression-like models $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$. Consequently, we use the ridge traces to determine a value of k to be used in obtaining the ridge estimates of the parameters of the generalized JS circular regression model that gives better estimates of $\boldsymbol{\lambda}^{(j)}$, j=1, 2. Note that, when k = 0, the ridge estimates given by equation (6.20) reduces to the LS estimates given by equation (6.18).

6.5.2 Example using Simulated Data

A simulation study is conducted to investigate the effect of multicollinearity problem and to show the application of the proposed procedure of dealing with the problem in generalized JS circular regression model. The correlated circular variables are simulated by modifying the algorithm employed in Lawrence & Arthur (1990) for generating correlated variables in linear case to the circular case. Let us consider two circular explanatory variables U_1 and U_2. These variables are obtained using the following:

$$\boldsymbol{U}_{ij} = \left(1 - \rho_C{}^2\right)\boldsymbol{T}_{ij} + \rho_C\,\boldsymbol{T}_{i4}, \qquad i = 1, 2, \ldots, n \quad j = 1, 2, \qquad (6.24)$$

where ρ_C is the circular correlation between U_1 and U_2 , while $\boldsymbol{T}$ are independent wrapped normal variables with $\mu = \pi/2$ and ρ_C. From Jammalamadaka & Sengupta (2001), $U_1 + U_2$ are wrapped normal variates with parameter $\mu = \pi$ and concentration ρ_c^2. We let two values of $\rho_C = 0.9$ and 0.98 and investigate the effect for different size of sample n=10, 30 and 50 and $k \geq 0$. For simplicity, we consider the generalized JS circular regression model for $p = 2$ and, hence, we have 10 parameters to be estimated, that is, $\hat{\lambda}^{(1)} = (A_0, A_1, B_1, C_1, D_1)$ and $\hat{\lambda}^{(2)} = (E_0, E_1, F_1, G_1, H_1)$. The *VIF* values for cases considered above are presented in Table 6.6. It can be seen that the *VIF* values for $\rho_C = 0$ produced smaller *VIF* values compared to the others. When $\rho_C = 0.9$ and $\rho_C = 0.98$, at least one of the $VIF_{V_i,j}$ are greater than 10 indicating the existence of multicollinearity problem in the data. On the other hand, the resulting ridge traces corresponding to the observational regression-like models $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$ are plotted in Figures 6.1–6.3 for $\rho_C = 0.9$ and Figures 6.4–6.6 for $\rho_C = 0.98$. The following results are observed:

(i) For n=10 and $\rho_C = 0.9$, it can be seen from Figure 6.1 (i)- (ii), the traces for all parameters except E_0 and G_1 increase or decrease when k moves from 0 to 1. Some of the parameter estimates do stabilize as k approaches 1, for example, A_0 and D_1 in Figure 6.1(i) and all parameters except E_1 in Figure 6.1(ii). However, when n is increased to 30 and 50, the effect of multicollinearity on the parameter estimates is very minimal as illustrated in Figures 6.2 and 6.3 respectively.

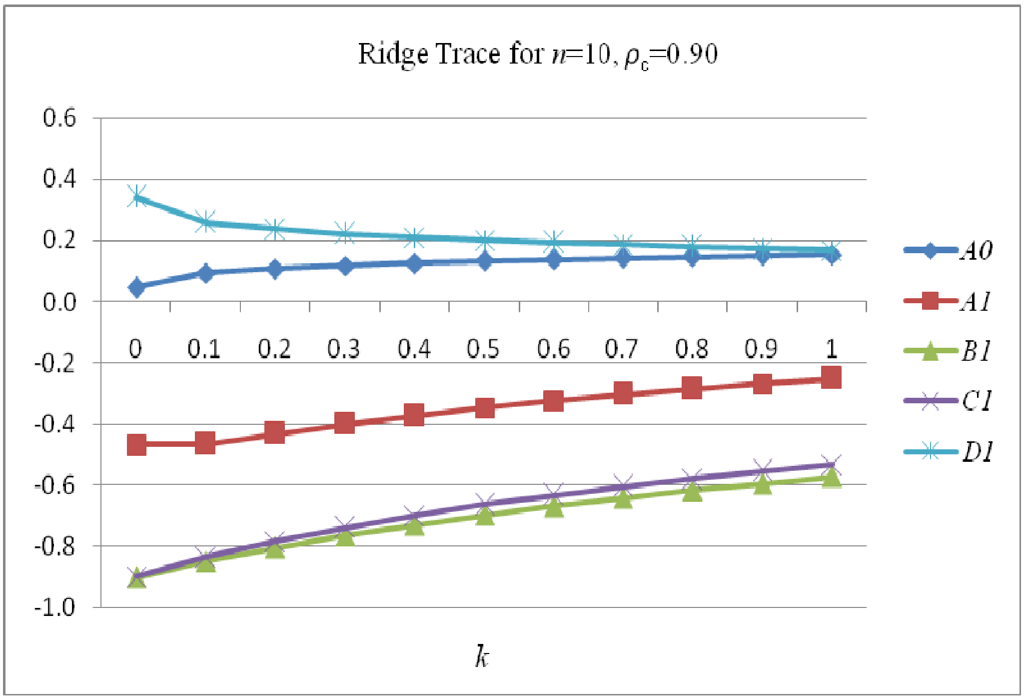
(ii) When we increase the concentration value ρ_C to 0.98, we can see that the effect of multicollinearity on the parameter estimates of the generalized JS circular regression model for the three cases considered, that is, n=10, 30 and 50 as shown in Figures 6.4.-6.6. Most LS estimates (when k=0) and the ridge estimates (when k>0) in the cases considered are changed. For example, the LS estimate $\hat{A}_{1,LS} = -1.4$ in Figure 6.4(i) is changed to $\hat{A}_{1,ridge} = -0.4$ when k=1.

(iii) The choice of k is made by looking at the point in the ridge trace whereby the traces are stabilizing. For $\rho_C = 0.9$, it is a bit difficult to find a stabilized point for n=10. However, for $\rho_C = 0.98$, most parameters stabilize at some point k.

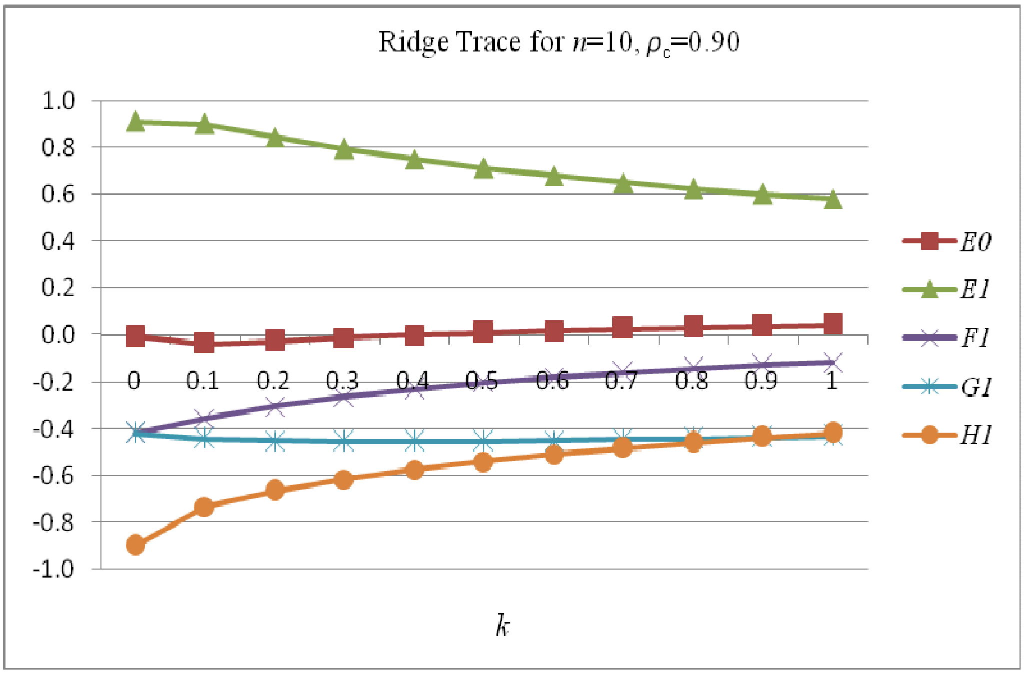
From the three simulated examples presented, we can see that the existence of multicollinearity in the data set affect the parameter estimates of the generalized JS circular regression model. In addition, this impact is also determined by the strength of the circular autocorrelation in terms of ρ_C of any two circular variables in the model. Ridge regression is then considered whereby the trend of ridge traces is examined and the 'best' value of k is chosen to give the ridge estimates of the model. For example, we may choose the value k=0.15 for the case n=50 and $\rho_C = 0.98$ giving the LS and Ridge estimates as tabulated in columns two and three of Table 6.7 respectively. Significant changes are observed for estimate of A_0, D_1, E_0, E_1 and H_1. It will be of our interest to see if the ridge regression model can overcome the problem by checking on the assumptions made on the generalized JS circular regression model, in particular, on normality assumptions of the residuals resulting from the observational regression-like models.

Table 6.6: VIF values for $\rho_C = 0$, $\rho_C = 0.9$ and $\rho_C = 0.98$

	$\rho_C = 0$		$\rho_C = 0.9$		$\rho_C = 0.98$	
n	$VIF_{V_1,j}$	$VIF_{V_2,j}$	$VIF_{V_1,j}$	$VIF_{V_2,j}$	$VIF_{V_1,j}$	$VIF_{V_2,j}$
10	1.0358	1.4489	22.5921	21.3665	44.5545	35.0007
	1.2169	1.5226	22.6006	21.2306	44.5711	33.0483
30	1.0445	1.1536	16.8884	10.4501	39.9190	22.2203
	1.1546	1.0404	17.1228	10.4783	37.0563	25.3020
50	1.0311	1.0211	11.3582	6.3343	19.9845	10.9741
	1.0250	1.0237	11.4683	6.0205	19.6421	10.6392

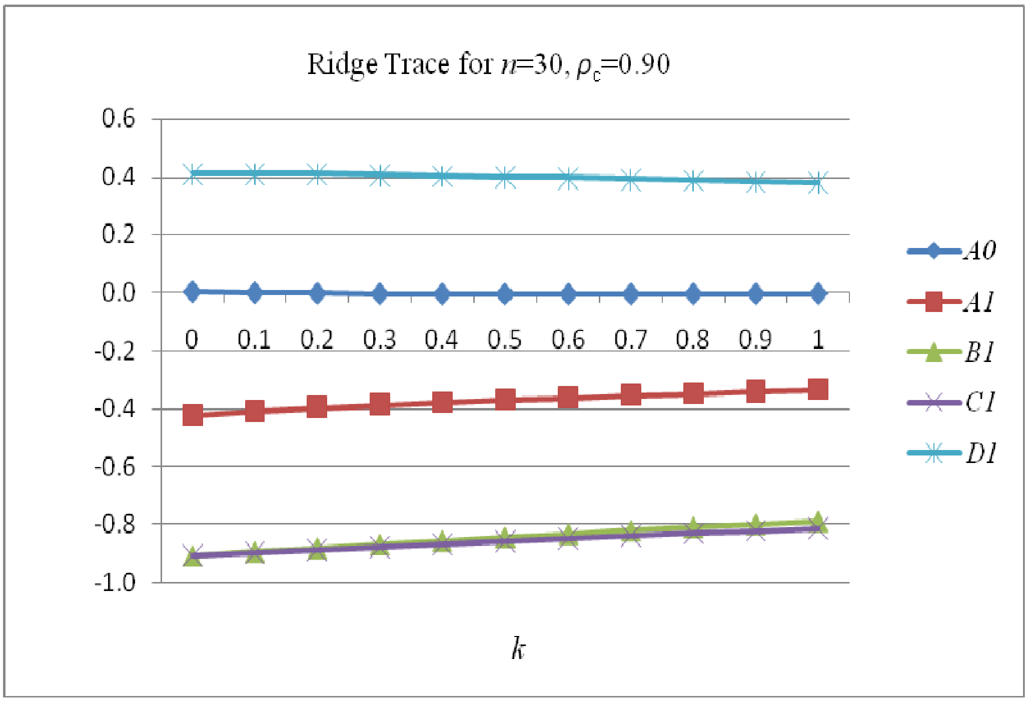


(i) $\hat{\lambda}^{(1)}$

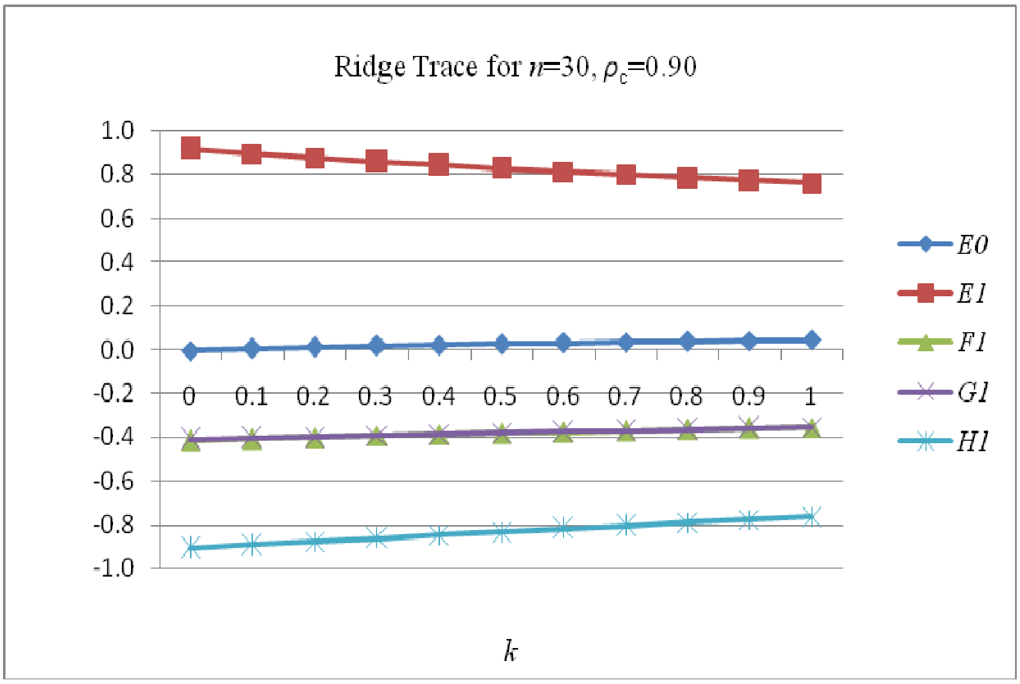


(ii) $\hat{\lambda}^{(2)}$

Figure 6.1: Ridge Trace for simulated data, n=10, $\rho_c = 0.9$

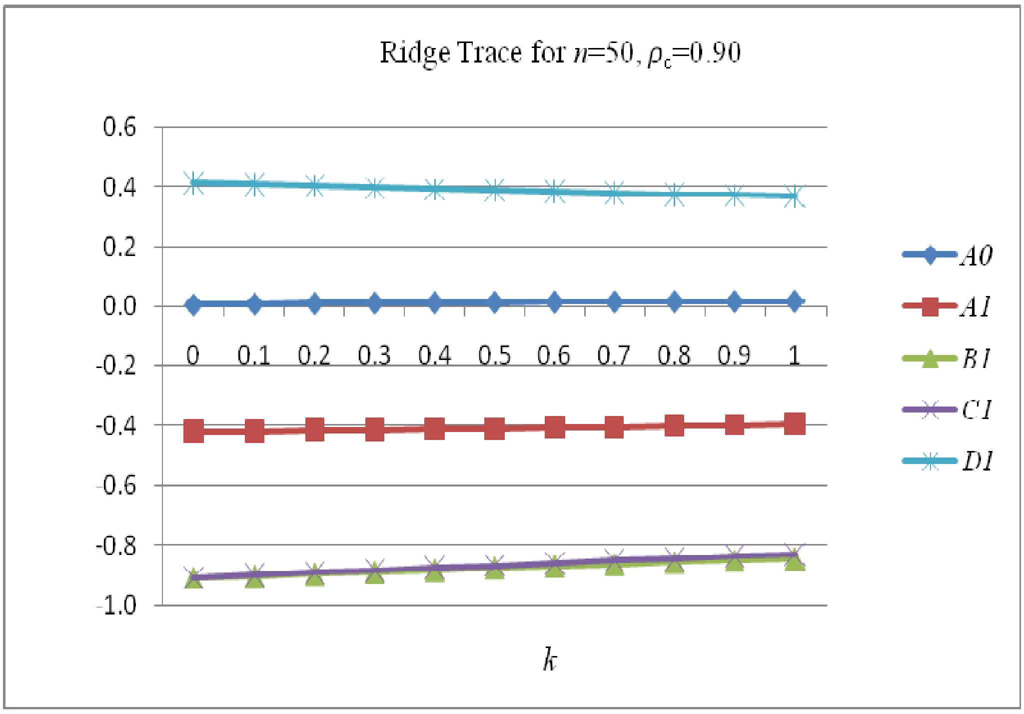


(i) $\hat{\lambda}^{(1)}$

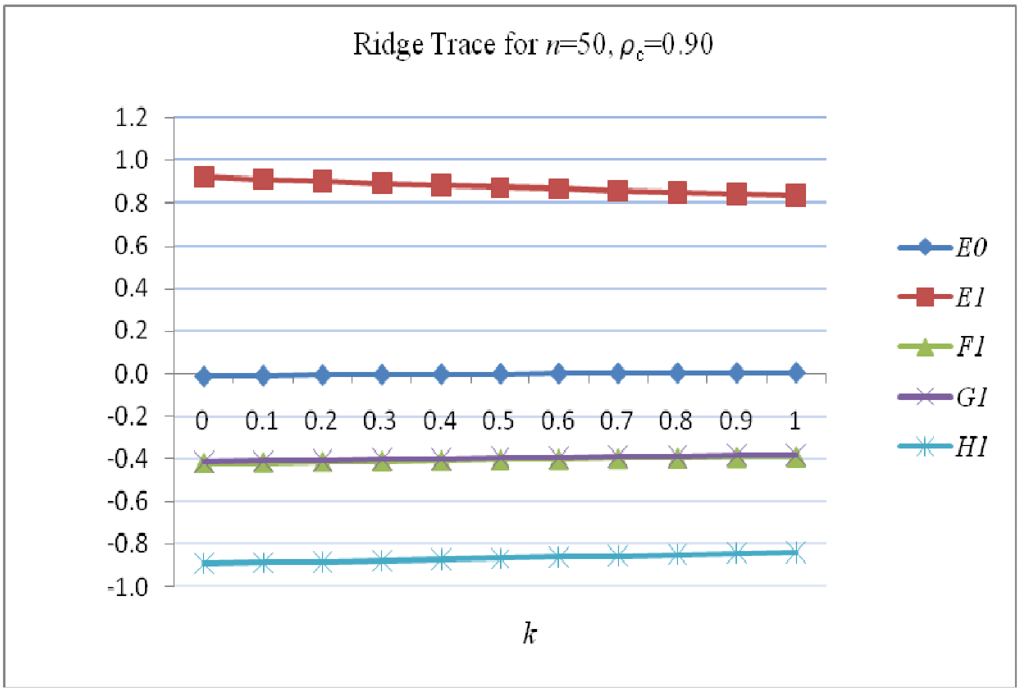


(ii) $\hat{\lambda}^{(2)}$

Figure 6.2: Ridge Trace for simulated data, n=30, $\rho_C = 0.9$

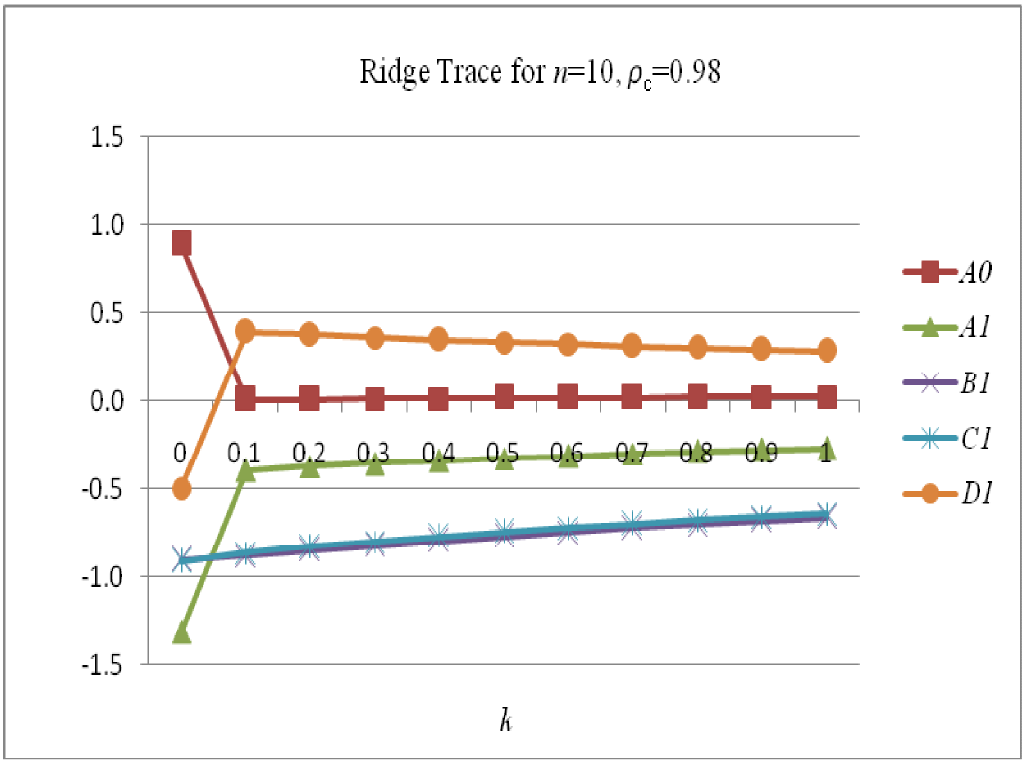


(i) $\hat{\lambda}^{(1)}$

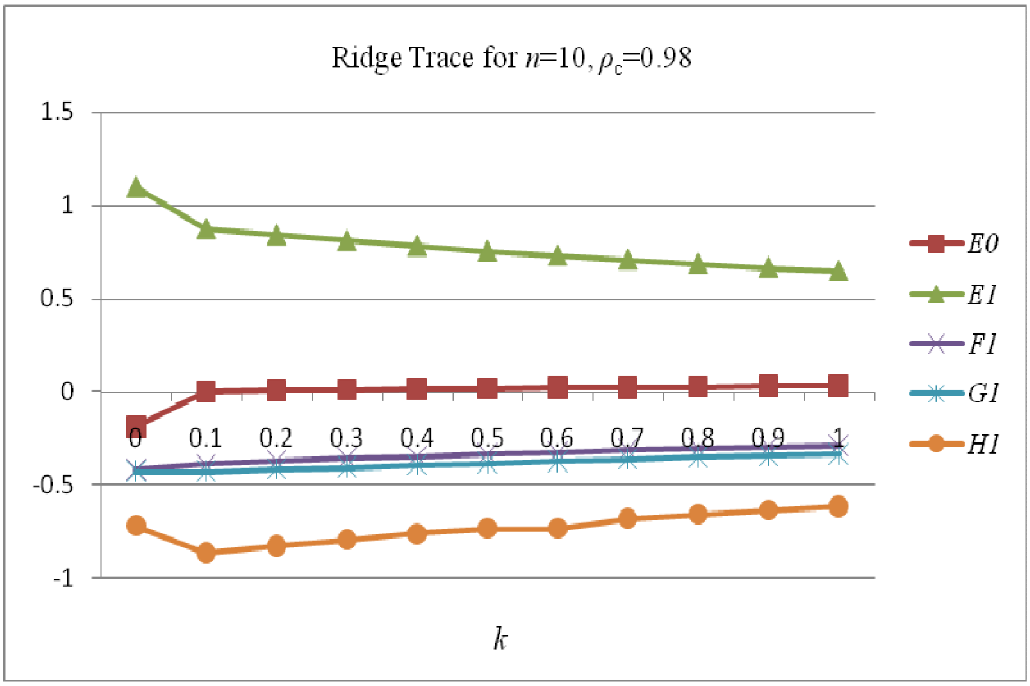


(ii) $\hat{\lambda}^{(2)}$

Figure 6.3: Ridge Trace for simulated data, n=50, $\rho_c = 0.9$

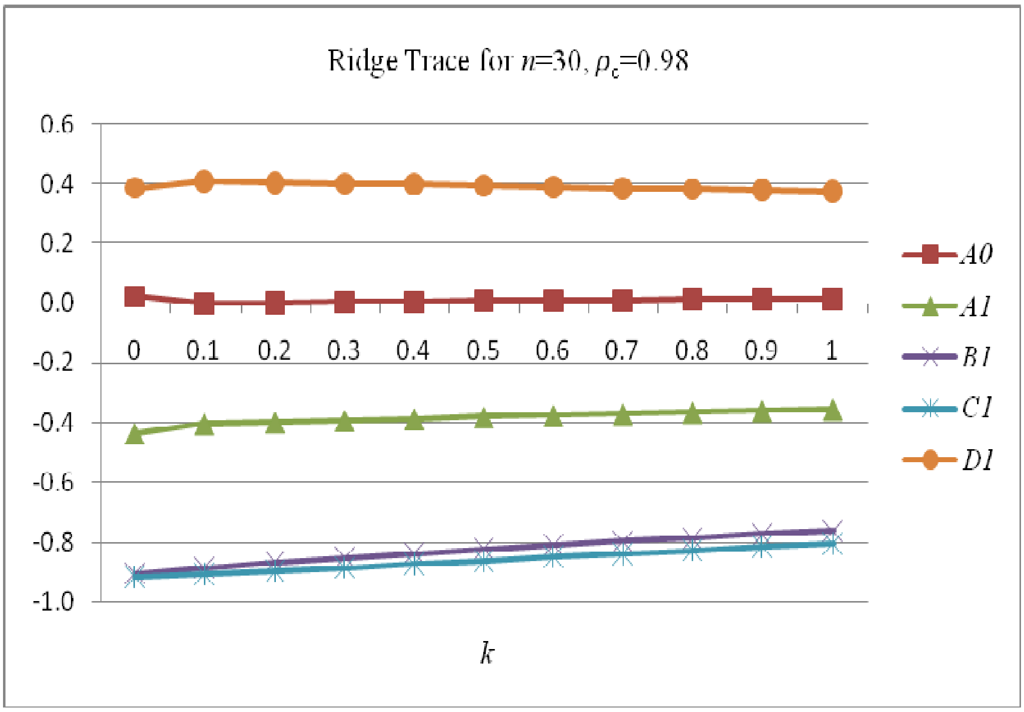


(i) $\hat{\lambda}^{(1)}$

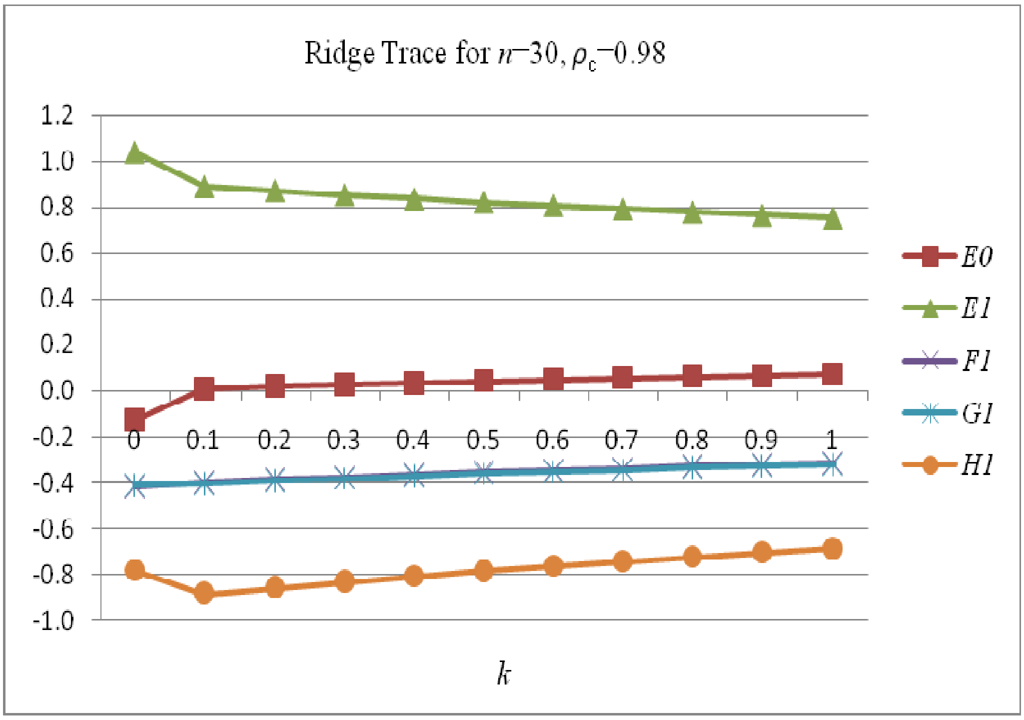


(ii) $\hat{\lambda}^{(2)}$

Figure 6.4: Ridge Trace for simulated data, n=10, $\rho_c = 0.98$

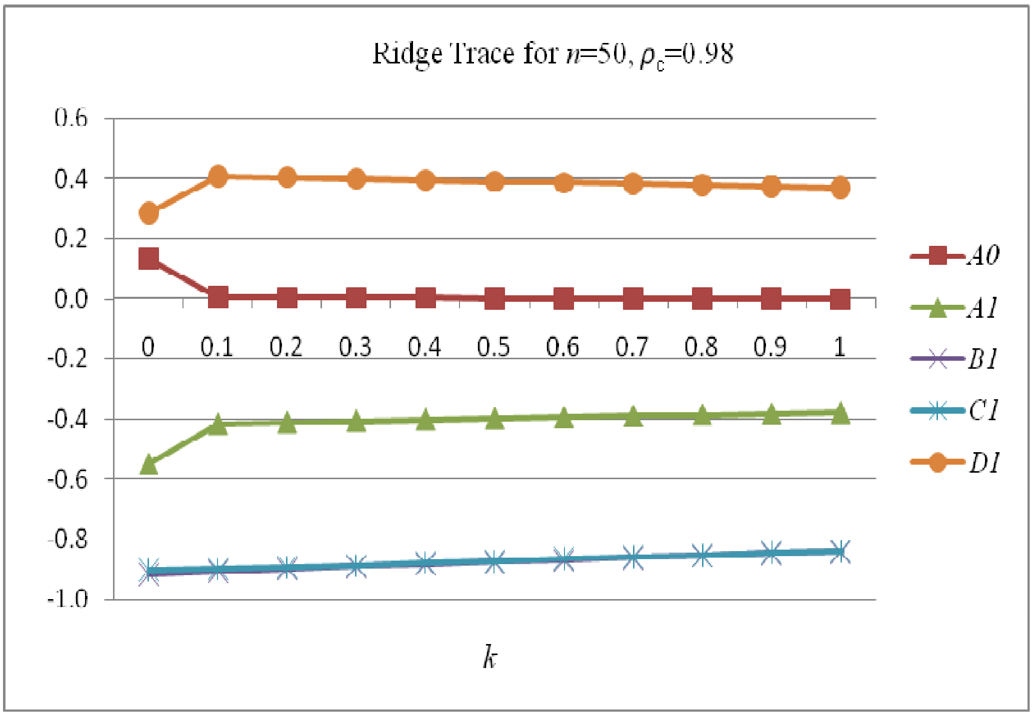


(i) $\hat{\lambda}^{(1)}$

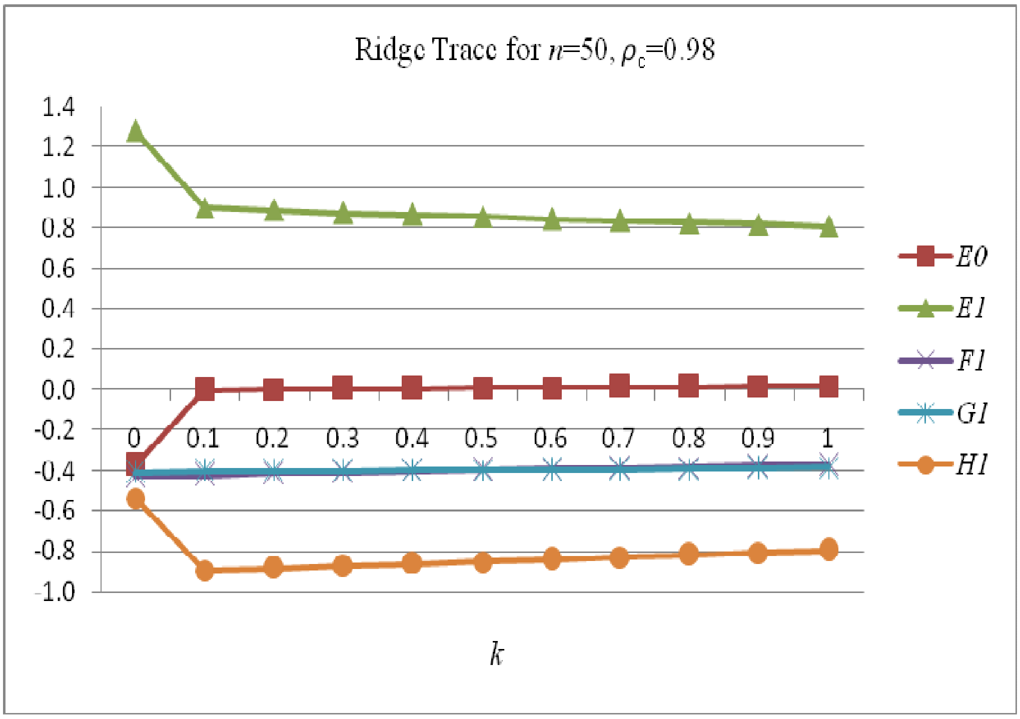


(ii) $\hat{\lambda}^{(2)}$

Figure 6.5: Ridge Trace for simulated data, n=30, $\rho_c = 0.98$



(i) $\hat{\lambda}^{(1)}$



(ii) $\hat{\lambda}^{(2)}$

Figure 6.6: Ridge Trace for simulated data, n=50, $\rho_c = 0.98$

Table 6.7: The LS and Ridge regression estimates for n = 50 at $\rho_c = 0.98$

Parameter Estimates	k = 0 (LS)	k = 0.15 (Ridge regression)
$\hat{A}_0^*$	0.0133	0.0031
$\hat{A}_1^*$	-0.4549	-0.4122
$\hat{B}_1^*$	-0.9160	-0.9010
$\hat{C}_1^*$	-0.9013	-0.8952
$\hat{D}_1^*$	0.3285	0.4069
$\hat{E}_0^*$	-0.3719	0.0016
$\hat{E}_1^*$	1.2851	0.8949
$\hat{F}_1^*$	-0.4272	-0.4185
$\hat{G}_1^*$	-0.4078	-0.4016
$\hat{H}_1^*$	-0.5376	-0.8907

6.5.3 The Performance of Least Squares and Circular Ridge Regression

(i) Simulation Procedure

A simulation study is carried out to investigate the performance of Least Squares method and Circular Ridge regression in estimating the parameters of generalized JS circular regression model. Three different values of $\rho_c = 0$, 0.9 and 0.98 when k=0, 0.5 and 0.9 and different sample size n=10, 30 and 50 are considered. The same procedure employed in Section 6.5.2 is used here to generate the data set. We consider the generalized JS circular regression model for p = 2 and, hence, we have 10 parameters to be estimated, that is $\hat{\lambda}^{(1)} = (A_0, A_1, B_1, C_1, D_1)$ and $\hat{\lambda}^{(2)} = (E_0, E_1, F_1, G_1, H_1)$. When at least one of the $VIF_{V_i,j}$ are greater than 10 indicating the existence of multicollinearity problem in the data, then the generated circular data (u_j, v_j) is fitted using ridge circular regression by adding a constant k as given in

Equation (6.20). If $VIF_{V_i,j}$ values are less than 10, the generated circular data (u_j, v_j) is fitted to LS method. Finally, the steps above are repeated for 1000 times. For each parameter estimates $\hat{\lambda}^{(1)} = (A_0, A_1, B_1, C_1, D_1)$ and $\hat{\lambda}^{(2)} = (E_0, E_1, F_1, G_1, H_1)$, the estimated mean, bias, *SE* and *RMSE* are calculated as described in Section 3.4.1.

(ii) Discussion

The results for mean, bias, *SE* and *RMSE* of each parameter estimates are tabulated in Tables 6.8 – 6.10 for different values of ρ_C and *k*. Several results are observed as follow:

1. For LS estimates (when *k*=0), the estimated mean for all parameter estimates are consistently close to the true values. The estimation further improves when the sample size *n* increase. When the concentration value ρ_C increases, the bias is generally larger than the ρ_C=0. This is because the existence of multicollinearity in the data set affects the parameter estimates of the generalized JS circular regression.
2. When the value of ρ_C increase from 0 to 0.98, the value of the bias increases. Meanwhile, the value of bias decrease when the sample size *n* increases.
3. The *SE* for all parameter estimates are generally small for ρ_C=0, but gets larger as the value of ρ_C increase. The value of *RMSE* of each parameter estimates increase when the value of ρ_C increase.
4. For ridge estimates, when the value of *k* increase to 0.9, the value of bias, *SE* and *RMSE* gets smaller for all the parameter estimates compared to LS estimates.
5. When the value of ρ_C in ridge estimates increases, the value of bias, *SE* and *RMSE* gets larger for all the parameter estimates.

6. The value of bias, *SE* and *RMSE* increase when the sample size *n* increases.

By looking at the results, the ridge circular regression performs well in estimating the parameters of generalized circular regression model when the multicollinearity exists in the data set. So, the ridge circular regression has improved the multicollinearity problem since the simulation results showed that the value of *SE* and *RMSE* for ridge circular regression is smaller compared to the value of *SE* and *RMSE* of LS estimates.

6.6 Practical Example: Multivariate Eye Data

We consider a multivariate eye data set comprising one dependent circular variable and three explanatory variables which are suspected to be dependent giving a problem of multicollinearity in the data set. We fit the generalized JS circular regression model on the data and deal with the multicollinearity problem, if exists, in the data.

6.6.1 Description of the Data

There are two types of grading of angle of eye based on Schaeffer grading giving different levels – IV. The grades I and II are the angles for closure glaucoma while the grades III and IV are for the open-angle glaucoma. The primary angle-closure glaucoma (PACG) is a significant cause of blindness in East Asia and South Asia, see Foster *et al*. (1996), Foster *et al*. (2000), Dandona *et al*. (2000) and Jacob *et al*. (1998). Optical coherence tomography (OCT) technology originally is used in ophthalmology to obtain the images of the posterior segment and the anterior segment structures such as the cornea. The angles of the anterior segment optical coherence tomography are obtained with AS-OCT.

Table 6.8: Parameter estimates for different values of ρ_C when k=0

$\rho_C = 0$

Estimates	true values	10				30				50			
		mean	*SE*	bias	*RMSE*	mean	*SE*	bias	*RMSE*	mean	*SE*	bias	*RMSE*
$\hat{A}_0$	0.0000	0.0571	1.0721	0.0571	1.8071	0.0203	0.6277	0.0203	0.6431	0.0038	0.3593	0.0038	0.1194
$\hat{A}_1$	-0.4161	-0.4535	1.1475	-0.0374	1.1824	-0.4160	0.6891	0.0002	0.0055	-0.3959	0.4044	0.0203	0.6410
$\hat{B}_1$	-0.9093	-0.8635	0.3300	0.0458	1.4487	-0.8540	0.1979	0.0553	1.7497	-0.8552	0.1361	0.0541	1.7121
$\hat{C}_1$	-0.9093	-0.8659	0.3375	0.0434	1.3716	-0.8551	0.1985	0.0542	1.7131	-0.8550	0.1294	0.0543	1.7184
$\hat{D}_1$	0.4161	0.3368	1.1548	-0.0794	2.5106	0.3716	0.6810	-0.0446	1.4088	0.3879	0.4072	-0.0283	0.8936
$\hat{E}_0$	0.0000	-0.0049	1.0992	-0.0049	0.1551	-0.0230	0.6610	-0.0230	0.7268	-0.0161	0.3611	-0.0161	0.5085
$\hat{E}_1$	0.9093	0.8777	1.1645	-0.0316	0.9980	0.8793	0.7138	-0.0300	0.9483	0.8720	0.4007	-0.0373	1.1800
$\hat{F}_1$	-0.4161	-0.4154	0.3362	0.0007	0.0234	-0.3972	0.2140	0.0189	0.5978	-0.3971	0.1356	0.0191	0.6036
$\hat{G}_1$	-0.4161	-0.3735	0.3434	0.0426	1.3479	-0.3899	0.2043	0.0262	0.8285	-0.3944	0.1340	0.0217	0.6865
$\hat{H}_1$	-0.9093	-0.8598	1.1660	0.0495	1.5652	-0.8279	0.7122	0.0814	2.5740	-0.8385	0.4025	0.0708	2.2380

$\rho_C = 0.9$

Estimates	true values	10				30				50			
		mean	*SE*	bias	*RMSE*	mean	*SE*	bias	*RMSE*	mean	*SE*	bias	*RMSE*
$\hat{A}_0$	0.0000	1.1113	21.3160	1.1113	35.1418	-0.1098	10.0777	-0.1099	3.4738	0.1995	5.8544	0.1995	6.3087
$\hat{A}_1$	-0.4161	-1.5103	21.4768	-1.0941	34.5987	-0.3571	10.1771	0.0590	1.8667	-0.6470	5.8769	-0.2308	7.2989
$\hat{B}_1$	-0.9093	-0.8615	1.2502	0.0478	1.5102	-0.8458	0.7196	0.0635	2.0087	-0.8643	0.4257	0.0450	1.4226
$\hat{C}_1$	-0.9093	-0.8915	1.2418	0.0178	0.5633	-0.8679	0.7355	0.0414	1.3079	-0.8561	0.4256	0.0532	1.6808
$\hat{D}_1$	0.4161	-0.7228	21.3667	-1.1389	36.0160	0.5178	10.1249	0.1016	3.2135	0.2013	5.8934	-0.2148	6.7938
$\hat{E}_0$	0.0000	-0.0815	18.9063	-0.0815	2.5777	-0.0943	9.2567	-0.0943	2.9805	-0.1554	6.0965	-0.1554	4.9133
$\hat{E}_1$	0.9093	0.9407	19.0221	0.0314	0.9945	0.9608	9.3344	0.0515	1.6285	1.0525	6.1365	0.1432	4.5293
$\hat{F}_1$	-0.4161	-0.3677	1.2208	0.0484	1.5305	-0.3505	0.6837	0.0657	2.0763	-0.3833	0.4317	0.0329	1.0401
$\hat{G}_1$	-0.4161	-0.4294	1.2034	-0.0132	0.4185	-0.3918	0.6840	0.0244	0.7702	-0.3617	0.4358	0.0545	1.7224
$\hat{H}_1$	-0.9093	-0.7771	18.9577	0.1321	4.1779	-0.7594	9.2967	0.1499	4.7392	-0.7001	6.1292	0.2092	6.6149

$\rho_C = 0.98$

Estimates	true values	10				30				50			
		mean	*SE*	bias	*RMSE*	mean	*SE*	bias	*RMSE*	mean	*SE*	bias	*RMSE*
$\hat{A}_0$	0.0000	3.6742	96.7721	3.6742	82.1575	-1.4098	49.3947	-1.4098	31.5240	0.9258	22.3134	0.9258	20.7014
$\hat{A}_1$	-0.4161	-4.2664	97.4482	-3.8502	86.0941	0.6361	49.5845	1.0522	23.5279	-1.4933	22.3060	-1.0771	24.0852
$\hat{B}_1$	-0.9093	-0.9540	2.8420	-0.0447	0.9988	-0.7726	1.5819	0.1367	3.0562	-0.8424	0.7455	0.0669	1.4952
$\hat{C}_1$	-0.9093	-0.8720	2.7552	0.0373	0.8349	-0.8972	1.5636	0.0121	0.2697	-0.8078	0.7912	0.1015	2.2702
$\hat{D}_1$	0.4161	-3.2859	96.8442	-3.7020	82.7793	1.8402	49.4465	1.4240	31.8421	-0.5024	22.3427	-0.9186	20.5397
$\hat{E}_0$	0.0000	-0.4229	78.9227	-0.4229	9.4574	-0.3164	42.5512	-0.3164	7.0747	0.8007	20.1894	0.8007	17.9038
$\hat{E}_1$	0.9093	1.1806	79.3632	0.2713	6.0664	0.9621	42.8126	0.0528	1.1810	0.0328	20.1240	-0.8765	19.6001
$\hat{F}_1$	-0.4161	-0.5010	2.5058	-0.0848	1.8964	-0.3306	1.4588	0.0856	1.9140	-0.3674	0.7143	0.0488	1.0906
$\hat{G}_1$	-0.4161	-0.3399	2.6204	0.0762	1.7045	-0.3353	1.4423	0.0809	1.8080	-0.2978	0.7253	0.1183	2.6461
$\hat{H}_1$	-0.9093	-0.4447	78.9721	0.4646	10.3892	-0.5198	42.5889	0.3895	8.7093	-1.6450	20.2172	-0.7357	16.4504

Table 6.9: Parameter estimates for different values of ρ_C when k=0.5

$\rho_C = 0$

Estimates	true values	10				30				50			
		mean	*SE*	bias	*RMSE*	mean	*SE*	bias	*RMSE*	mean	*SE*	bias	*RMSE*
$\hat{A}_0$	0.0000	0.0000	0.0100	0.1202	0.0100	0.2243	0.0106	0.1542	0.0106	0.2369	0.0143	0.1576	0.0143
$\hat{A}_1$	-0.4161	-0.4161	-0.3492	0.1786	0.0670	1.4976	-0.3809	0.1542	0.0353	0.7889	-0.3940	0.1853	0.0222
$\hat{B}_1$	-0.9093	-0.9093	-0.7896	0.1864	0.1197	2.6765	-0.8130	0.1589	0.0963	2.1540	-0.8341	0.1223	0.0752
$\hat{C}_1$	-0.9093	-0.9093	-0.7866	0.1794	0.1227	2.7428	-0.8267	0.1608	0.0826	1.8475	-0.8352	0.1180	0.0741
$\hat{D}_1$	0.4161	0.4161	0.3528	0.1817	-0.0634	1.4170	0.3656	0.1871	-0.0505	1.1298	0.3696	0.1914	-0.0465
$\hat{E}_0$	0.0000	0.0000	-0.0110	0.1288	-0.0110	0.2451	-0.0007	0.1597	-0.0007	0.0154	-0.0126	0.1634	-0.0126
$\hat{E}_1$	0.9093	0.9093	0.7987	0.1680	-0.1106	2.4723	0.8152	0.1891	-0.0941	2.1033	0.8467	0.1903	-0.0626
$\hat{F}_1$	-0.4161	-0.4161	-0.3758	0.1810	0.0403	0.9011	-0.3786	0.1576	0.0376	0.8405	-0.3884	0.1239	0.0277
$\hat{G}_1$	-0.4161	-0.4161	-0.3584	0.1985	0.0577	1.2906	-0.3739	0.1536	0.0423	0.9450	-0.3923	0.1182	0.0238
$\hat{H}_1$	-0.9093	-0.9093	-0.7863	0.1616	0.1230	2.7506	-0.8172	0.1825	0.0921	2.0597	-0.8198	0.1883	0.0895

$\rho_C = 0.9$

Estimates	true values	10				30				50			
		mean	*SE*	bias	*RMSE*	mean	*SE*	bias	*RMSE*	mean	*SE*	bias	*RMSE*
$\hat{A}_0$	0.0000	0.0000	0.0961	0.1138	0.0961	2.1483	0.0437	0.0929	0.0437	0.9771	0.0150	0.0727	0.0150
$\hat{A}_1$	-0.4161	-0.4161	-0.1874	0.2287	0.2287	5.1143	-0.3110	0.1860	0.1052	2.3516	-0.3657	0.1466	0.0505
$\hat{B}_1$	-0.9093	-0.9093	-0.7360	0.2291	0.1733	3.8754	-0.7856	0.2463	0.1237	2.7655	-0.8166	0.2237	0.0927
$\hat{C}_1$	-0.9093	-0.9093	-0.7516	0.2286	0.1577	3.5264	-0.7811	0.2415	0.1282	2.8676	-0.8276	0.2252	0.0816
$\hat{D}_1$	0.4161	0.4161	0.2807	0.1382	-0.1354	3.0286	0.3517	0.1119	-0.0645	1.4419	0.3784	0.0972	-0.0378
$\hat{E}_0$	0.0000	0.0000	-0.1072	0.1217	-0.1072	2.3971	-0.0540	0.1014	-0.0541	1.2093	-0.0242	0.0711	-0.0242
$\hat{E}_1$	0.9093	0.9093	0.5857	0.2651	-0.3236	7.2361	0.7154	0.2168	-0.1939	4.3357	0.7873	0.1589	-0.1220
$\hat{F}_1$	-0.4161	-0.4161	-0.3954	0.2070	0.0208	0.4645	-0.3658	0.2127	0.0504	1.1267	-0.3816	0.2393	0.0345
$\hat{G}_1$	-0.4161	-0.4161	-0.3998	0.2226	0.0163	0.3652	-0.3788	0.2245	0.0373	0.8347	-0.3730	0.2425	0.0432
$\hat{H}_1$	-0.9093	0.0000	0.0961	0.1138	0.0961	2.1483	0.0437	0.0929	0.0437	0.9771	0.0150	0.0727	0.0150

$\rho_C = 0.98$

Estimates	true values	10				30				50			
		mean	*SE*	bias	*RMSE*	mean	*SE*	bias	*RMSE*	mean	*SE*	bias	*RMSE*
$\hat{A}_0$	0.0000	0.0000	0.2650	0.0874	0.2650	5.9251	0.2353	0.1117	0.2353	5.2611	0.1975	0.1334	0.1975
$\hat{A}_1$	-0.4161	-0.4161	0.0965	0.1451	0.5126	11.4627	0.0204	0.1878	0.4366	9.7622	-0.0433	0.2344	0.3728
$\hat{B}_1$	-0.9093	-0.9093	-0.5773	0.2318	0.3320	7.4238	-0.5998	0.2452	0.3095	6.9198	-0.6416	0.2557	0.2677
$\hat{C}_1$	-0.9093	-0.9093	-0.5795	0.2250	0.3298	7.3744	-0.6006	0.2359	0.3087	6.9025	-0.6600	0.2481	0.2493
$\hat{D}_1$	0.4161	0.4161	0.1679	0.0954	-0.2482	5.5504	0.2138	0.0944	-0.2023	4.5246	0.2395	0.1133	-0.1766
$\hat{E}_0$	0.0000	0.0000	-0.3124	0.0828	-0.3124	6.9854	-0.2872	0.1117	-0.2872	6.4230	-0.2565	0.1274	-0.2565
$\hat{E}_1$	0.9093	0.9093	0.1826	0.1818	-0.7267	16.2497	0.2360	0.2392	-0.6733	15.0548	0.3073	0.2680	-0.6020
$\hat{F}_1$	-0.4161	-0.4161	-0.3844	0.1825	0.0318	0.7103	-0.3738	0.1882	0.0423	0.9460	-0.3845	0.2076	0.0316
$\hat{G}_1$	-0.4161	-0.4161	-0.3792	0.1818	0.0369	0.8256	-0.3656	0.1885	0.0505	1.1295	-0.3844	0.2023	0.0317
$\hat{H}_1$	-0.9093	-0.9093	-0.4945	0.1083	0.4148	9.2742	-0.5229	0.1322	0.3864	8.6412	-0.5641	0.1433	0.3452

Table 6.10: Parameter estimates for different values of ρ_C when k=0.9

$\rho_C = 0$

Estimates	true values	10				30				50			
		mean	SE	bias	RMSE	mean	SE	bias	RMSE	mean	SE	bias	RMSE
$\hat{A}_0$	0.0000	0.0228	0.0796	0.0228	0.7213	0.0129	0.0629	0.0129	0.4069	0.0084	0.0577	0.0084	0.2650
$\hat{A}_1$	-0.4161	-0.2445	0.1182	0.1717	5.4295	-0.3020	0.1040	0.1141	3.6088	-0.3355	0.0878	0.0806	2.5493
$\hat{B}_1$	-0.9093	-0.5795	0.1481	0.3298	10.4292	-0.6872	0.1225	0.2221	7.0238	-0.7511	0.1026	0.1582	5.0027
$\hat{C}_1$	-0.9093	-0.5873	0.1476	0.3220	10.1828	-0.6846	0.1236	0.2247	7.1048	-0.7509	0.0971	0.1584	5.0077
$\hat{D}_1$	0.4161	0.2636	0.1146	-0.1525	4.8234	0.3059	0.0965	-0.1103	3.4871	0.3386	0.0899	-0.0775	2.4513
$\hat{E}_0$	0.0000	-0.0126	0.0866	-0.0126	0.3986	-0.0077	0.0635	-0.0077	0.2442	-0.0082	0.0570	-0.0082	0.2605
$\hat{E}_1$	0.9093	0.5808	0.1130	-0.3285	10.3865	0.6811	0.0949	-0.2282	7.2173	0.7523	0.0778	-0.1570	4.9634
$\hat{F}_1$	-0.4161	-0.2822	0.1430	0.1340	4.2364	-0.3257	0.1270	0.0904	2.8596	-0.3532	0.1043	0.0630	1.9916
$\hat{G}_1$	-0.4161	-0.2740	0.1489	0.1421	4.4948	-0.3269	0.1243	0.0892	2.8210	-0.3531	0.1044	0.0630	1.9933
$\hat{H}_1$	-0.9093	-0.5863	0.1014	0.3230	10.2131	-0.6856	0.0895	0.2237	7.0739	-0.7487	0.0747	0.1606	5.0790

$\rho_C = 0.9$

Estimates	true values	10				30				50			
		mean	SE	bias	RMSE	mean	SE	bias	RMSE	mean	SE	bias	RMSE
$\hat{A}_0$	0.0000	0.1745	0.0872	0.1745	5.5189	0.1323	0.0755	0.1323	4.1829	0.0924	0.0571	0.0924	2.9212
$\hat{A}_1$	-0.4161	-0.0216	0.1329	0.3946	12.4777	-0.1198	0.1275	0.2963	9.3713	-0.1973	0.1114	0.2189	6.9211
$\hat{B}_1$	-0.9093	-0.5047	0.1303	0.4046	12.7949	-0.6051	0.1272	0.3042	9.6188	-0.6908	0.1077	0.2185	6.9098
$\hat{C}_1$	-0.9093	-0.5081	0.1327	0.4012	12.6859	-0.6052	0.1261	0.3041	9.6154	-0.6917	0.1099	0.2175	6.8795
$\hat{D}_1$	0.4161	0.1939	0.0860	-0.2222	7.0276	0.2503	0.0776	-0.1658	5.2431	0.2879	0.0716	-0.1282	4.0551
$\hat{E}_0$	0.0000	-0.1993	0.0921	-0.1993	6.3024	-0.1553	0.0814	-0.1553	4.9107	-0.1072	0.0629	-0.1072	3.3911
$\hat{E}_1$	0.9093	0.2847	0.1681	-0.6246	19.7530	0.4206	0.1657	-0.4887	15.4531	0.5547	0.1315	-0.3546	11.2131
$\hat{F}_1$	-0.4161	-0.2946	0.1131	0.1216	3.8440	-0.3255	0.1094	0.0907	2.8679	-0.3546	0.1007	0.0615	1.9451
$\hat{G}_1$	-0.4161	-0.2972	0.1115	0.1190	3.7621	-0.3280	0.1089	0.0881	2.7873	-0.3519	0.1064	0.0642	2.0301
$\hat{H}_1$	-0.9093	-0.4823	0.0935	0.4270	13.5034	-0.5753	0.0936	0.3340	10.5611	-0.6612	0.0759	0.2481	7.8468

$\rho_C = 0.98$

Estimates	true values	10				30				50			
		mean	SE	bias	RMSE	mean	SE	bias	RMSE	mean	SE	bias	RMSE
$\hat{A}_0$	0.0000	0.2852	0.0480	0.2852	9.0182	0.2749	0.0493	0.2749	8.6926	0.2852	0.0480	0.2852	9.0182
$\hat{A}_1$	-0.4161	0.0708	0.0509	0.4870	15.3998	0.0714	0.0639	0.4875	15.4177	0.0708	0.0509	0.4870	15.3998
$\hat{B}_1$	-0.9093	-0.3561	0.1218	0.5532	17.4935	-0.4460	0.1228	0.4633	14.6517	-0.3561	0.1218	0.5532	17.4935
$\hat{C}_1$	-0.9093	-0.3564	0.1197	0.5529	17.4853	-0.4465	0.1207	0.4628	14.6364	-0.3564	0.1197	0.5529	17.4853
$\hat{D}_1$	0.4161	0.2138	0.0458	-0.2023	6.3977	0.2030	0.0418	-0.2132	6.7419	0.2138	0.0458	-0.2023	6.3977
$\hat{E}_0$	0.0000	-0.3229	0.0442	-0.3229	10.2120	-0.3156	0.0467	-0.3156	9.9797	-0.3229	0.0442	-0.3229	10.2120
$\hat{E}_1$	0.9093	0.0928	0.0803	-0.8165	25.8185	0.1355	0.0937	-0.7738	24.4690	0.0928	0.0803	-0.8165	25.8185
$\hat{F}_1$	-0.4161	-0.2112	0.1002	0.2049	6.4796	-0.2752	0.0979	0.1410	4.4574	-0.2112	0.1002	0.2049	6.4796
$\hat{G}_1$	-0.4161	-0.2103	0.0996	0.2058	6.5087	-0.2759	0.0972	0.1403	4.4365	-0.2103	0.0996	0.2058	6.5087
$\hat{H}_1$	-0.9093	-0.4152	0.0500	0.4941	15.6249	-0.4507	0.0537	0.4586	14.5030	-0.4152	0.0500	0.4941	15.6249

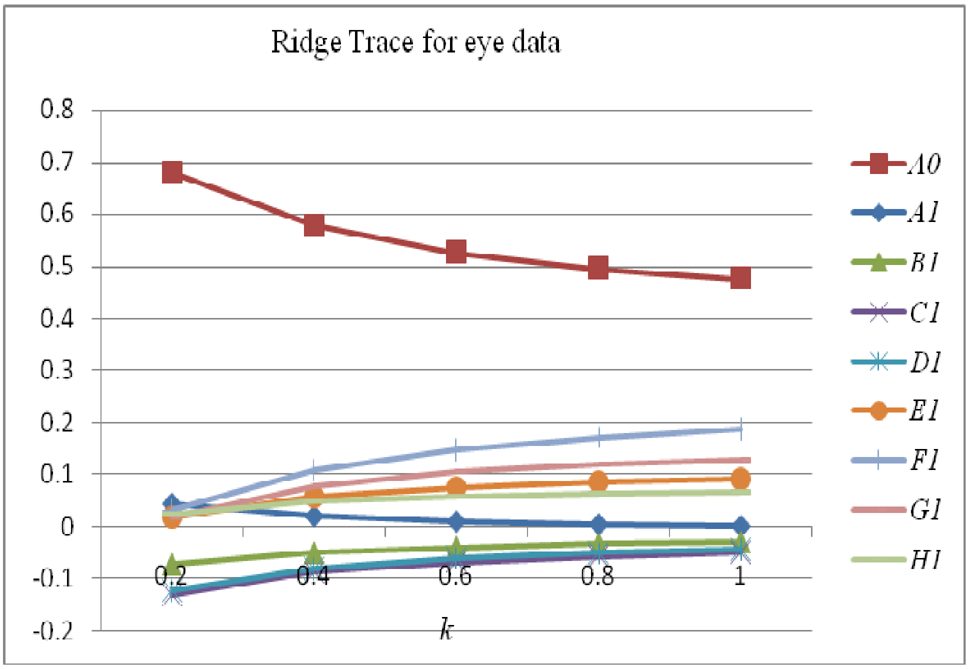
We consider the eye data consisting of 23 observations of glaucoma patients (unit in radians) recorded using Optical coherence tomography (OCT) at the University Malaya Medical Centre (UMMC). The variable measured are the angle of the eye between posterior corneal curvature to iris (v), the angle of the posterior corneal curvature (u_1), the angle of the posterior corneal curvature when length of the perpendicular is fixed to 2 mm (u_2) and the angle of the posterior corneal curvature when length of the perpendicular is fixed to 1.5 mm (u_3). It is suspected that the explanatory variables are correlated due to the way they are measured.

6.6.2 Detecting Multicollinearity

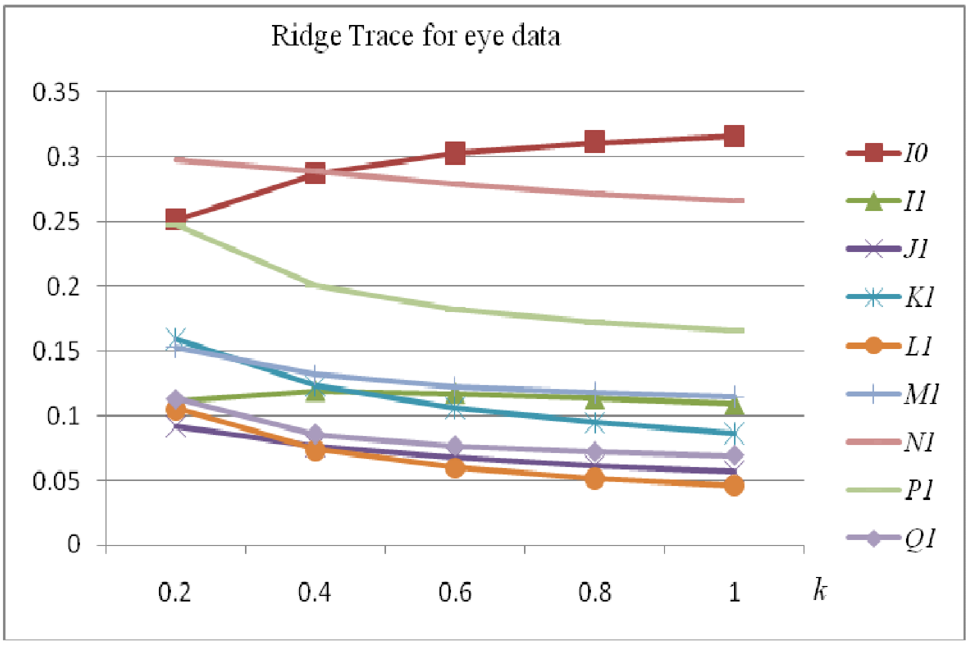
We now investigate the existence of multicollinearity problem in the multivariate eye data. The corresponding values of $VIF_{V_i,j}$ when the data is fitted using the generalized JS circular regression model is given in Table 6.11. It can be seen that at least one of the $VIF_{V_i,j}$ values are greater than 10, and hence, we conclude that there is multicollinearity in the data set. Hence, we will now analyze the data using the ridge regression approach.

Table 6.11: *VIF* values for the multivariate eye data

	$VIF_{V_1,j}$	$VIF_{V_2,j}$
VIF for u_1 on u_2 and u_3	42.20933	7.194354
VIF for u_2 on u_1 and u_3	10.13128	9.083798
VIF for u_3 on u_1 and u_2	21.91018	21.57776



(i) $\hat{\lambda}^{(1)}$



(ii) $\hat{\lambda}^{(2)}$

Figure 6.7: Ridge Trace for eye data

6.6.3 Ridge Trace

We now use the ridge regression approach on the multivariate eye data for the generalized JS circular regression model. Note that the circular regression model considered contains three explanatory variables, and hence, 18 parameters will be estimated giving the estimated parameters

$$\hat{\boldsymbol{\lambda}}^{(1)} = \left(\hat{A}_0, \hat{A}_1, \hat{B}_1, \hat{C}_1, \hat{D}_1, \hat{E}_1, \hat{F}_1, \hat{G}_1, \hat{H}_1\right)$$

$$\hat{\boldsymbol{\lambda}}^{(2)} = \left(\hat{I}_0, \hat{I}_1, \hat{J}_1, \hat{K}_1, \hat{L}_1, \hat{M}_1, \hat{N}_1, \hat{P}_1, \hat{Q}_1\right).$$

The ridge trace plots of $\hat{\boldsymbol{\lambda}}^{(1)}$ and $\hat{\boldsymbol{\lambda}}^{(2)}$ against k are given in Figure 6.7. The following results are observed:

(i) From Figure 6.7 (i) and (ii), the traces of constants A_0 and I_0 are consistently decreasing and increasing respectively throughout the interval $k \in [0,1]$ but the rate is also decreasing. Otherwise, the traces of other parameters in $\hat{\boldsymbol{\lambda}}^{(1)}$ and $\hat{\boldsymbol{\lambda}}^{(2)}$ stabilize rather quickly.

(ii) The traces of A_1, B_1, C_1 and D_1 ends at values close to zero, otherwise the other estimated parameters are different from zero.

(iii) Most of the coefficients including A_0 and I_0 seem to stabilize at k in interval (0.8, 1) and we may choose a value of k in this range. Hence, we use k = 0.9 to give the ridge estimation for the generalized JS circular regression data.

The LS (when k=0) and ridge (when k=0.9) estimates for the multivariate eye data are given in columns two and four of Table 6.12 respectively. As expected, we find that the LS and ridge estimates of the generalized JS circular regression model are different indicating the effect of multicollinearity problem in the data set. The changes of the value of $\hat{\kappa}$ has increased from 11.9041 to 19.6371 and $\hat{\rho}$ increased from 0.9570 to 0.9892. Furthermore, the standard error of the parameter estimates of ridge is smaller than LS. Therefore, fitted observational regression-like model of the generalized JS circular regression model with the ridge estimates is given by

$$\begin{aligned}\hat{g}_1(\boldsymbol{u}) = 0.4845 &+ 0.0043\cos u_1 \cos u_2 \cos u_3 - 0.0307\cos u_1 \cos u_2 \sin u_3 \\ &- 0.0519\cos u_1 \sin u_2 \cos u_3 - 0.0449\cos u_1 \sin u_2 \sin u_3 \\ &+ 0.0887\sin u_1 \cos u_2 \sin u_3 + 0.1810\sin u_1 \cos u_2 \cos u_3 \\ &+ 0.1234\sin u_1 \sin u_2 \cos u_3 + 0.0646\sin u_1 \sin u_2 \sin u_3 .\end{aligned}$$

$$\begin{aligned}\hat{g}_2(\boldsymbol{u}) = 0.3137 &+ 0.1113\cos u_1 \cos u_2 \cos u_3 + 0.0590\cos u_1 \cos u_2 \sin u_3 \\ &+ 0.0899\cos u_1 \sin u_2 \cos u_3 + 0.0484\cos u_1 \sin u_2 \sin u_3 \\ &+ 0.1159\sin u_1 \cos u_2 \sin u_3 + 0.2686\sin u_1 \cos u_2 \cos u_3 \\ &+ 0.1690\sin u_1 \sin u_2 \cos u_3 + 0.0704\sin u_1 \sin u_2 \sin u_3\end{aligned}$$

with the estimated concentration parameter $\hat{\rho}(\boldsymbol{u}) = 0.9892$.

Table 6.12: The LS and ridge circular regression estimates for multivariate eye data

Estimates	***k* = 0 (LS)**	**Standard error**	***k* = 0.9 (ridge)**	**Standard error**
$\hat{A}_0^*$	0.9549	1.7264	0.4845	0.2799
$\hat{A}_1^*$	0.3863	4.7124	0.0043	0.3248
$\hat{B}_1^*$	-0.2561	7.9341	-0.0307	0.4221
$\hat{C}_1^*$	-0.6669	6.9685	-0.0519	0.4016
$\hat{D}_1^*$	-0.6312	7.8745	-0.0449	0.4334
$\hat{E}_1^*$	-0.4672	6.2998	0.0887	0.4277
$\hat{F}_1^*$	0.4219	3.1321	0.1810	0.3591
$\hat{G}_1^*$	-0.6298	4.3270	0.1234	0.4131
$\hat{H}_1^*$	-0.5736	1.6366	0.0646	0.4361
I_0^*	0.4666	4.4675	0.3137	0.2635
I_1^*	0.3084	6.6063	0.1113	0.3059
J_1^*	0.2393	7.4652	0.0590	0.3974
K_1^*	0.8487	5.9723	0.0899	0.3782
L_1^*	0.7319	2.9693	0.0484	0.4081
M_1^*	0.6106	2.6601	0.1159	0.4026
N_1^*	0.4996	1.4250	0.2686	0.3380
P_1^*	1.0873	1.8338	0.1690	0.3889
Q_1^*	0.8702	0.9528	0.0704	0.4106
$\hat{\kappa}$	11.9041	-	19.6371	-
$\hat{\rho}$	0.9570	-	0.9892	-

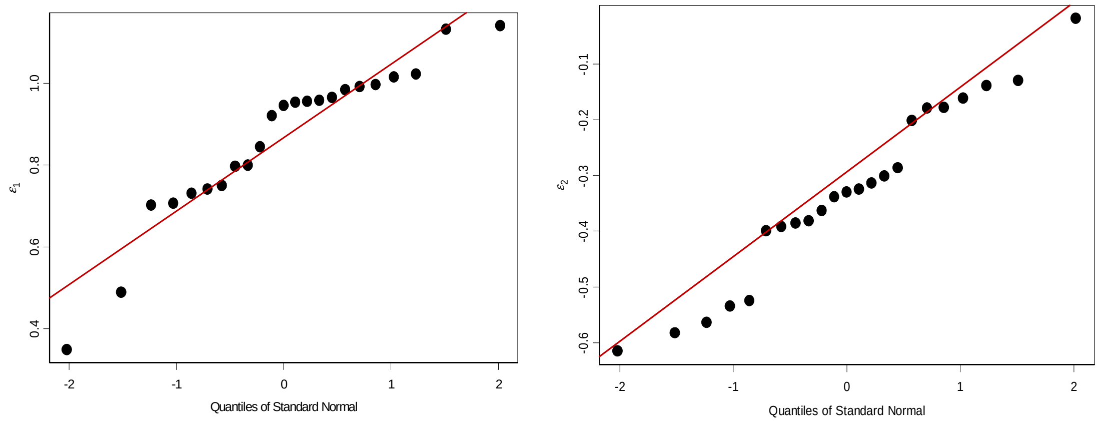
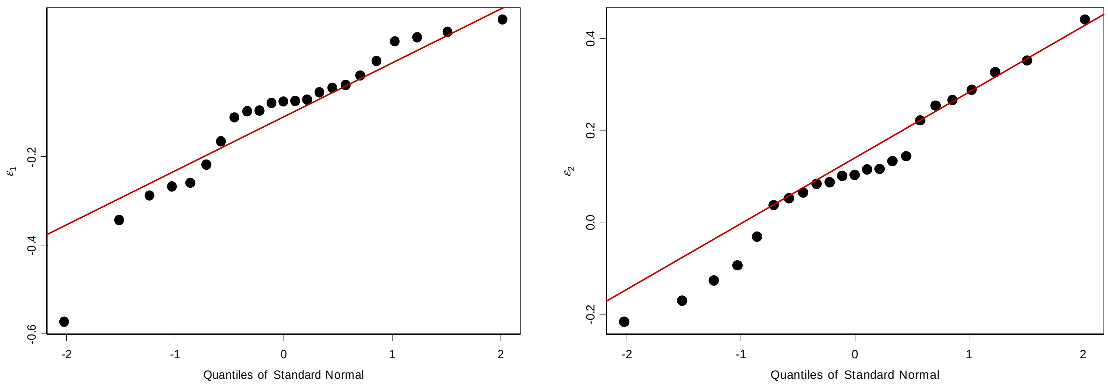


Figure 6.8: Q-Q plot of the residuals with LS estimates



(a) ε_1 (a) ε_2

Figure 6.9: Q-Q plot of the residuals with ridge estimates

The Q-Q plots for the residuals resulting from the fitted observational regression-like models with LS and ridge estimates are given in Figure 6.8 and Figure 6.9 respectively. We can see that the Q-Q plots associated to ridge regression have the points lying closer to the straight line. In addition, only one point seems to be far from to others in Figure 6.9(a) compared to two in Figure 6.8(a). This outlying point warrants further investigation on the possibility of the occurrence of outlier. We also find the *AICC* values for both cases; the value is -61.42 for LS and -80.76 for ridge, indicating the data is better fitted using the ridge regression approach.

6.7 Summary

We have proposed the formulation of the generalized JS circular regression model for more than one circular explanatory variables and the estimation of the regression parameters using the least squares method. We then investigate the problem of multicollinearity in the model and provide the solution using ridge circular regression approach. Since the effect of multicollinearity is very serious, which provides larger standard error (*SE*) of parameter estimates, the proposed ridge circular regression type approach used to remedy the problems of multicollinearity since this method has showed smaller SE of estimates compared to LS estimates. The practical example is given for illustration purpose.

CHAPTER SEVEN

FUNCTIONAL RELATIONSHIP MODEL FOR JS CIRCULAR REGRESSION

7.1 Introduction

In this chapter, we extend the linear functional relationship model to the JS circular regression models, when both variables *U* and *V* are subject to errors. We assume that the errors are independently distributed from bivariate complex Gaussian distribution. The theoretical background of the model is presented and the maximum likelihood estimators of the parameters are derived assuming the ratio of the error variances is known. We then present the application of the model on eye data.

7.2 The JS Circular Functional Relationship Model

Suppose u_j and v_j are the observed values of the circular variables *U* and *V* respectively, thus $0 < u_j, v_j \le 2\pi$, for $j = 1,...,n$. The circular variables *U* and *V* are true values corresponding to u_j and v_j respectively and there are linear relationship between these two variables. For any fixed U_j, we assume that the observations u_j and v_j have been measured with errors δ_j and ε_j respectively. Thus, the unreplicated JS circular functional relationship model can be written as

$$\left(\cos u_j + i\sin u_j\right) = \left(\cos U_j + i\sin U_j\right) + \left(\delta_{jR} + i\,\delta_{jI}\right)$$

$$\left(\cos v_j + i\sin v_j\right) = \left(\cos V_j + i\sin V_j\right) + \left(\varepsilon_{jR} + i\,\varepsilon_{jI}\right) \qquad (7.1)$$

or

$$\cos u_j = \cos U_j + \delta_{jR}$$

$$\sin u_j = \sin U_j + \delta_{jI}$$

$$\cos v_j = \cos V_j + \varepsilon_{jR}$$

$$\sin v_j = \sin V_j + \varepsilon_{jI} \,,$$

where

$$\begin{aligned} \cos V_j &= A_0 + A_1 \cos U_j + B_1 \sin U_j \\ \sin V_j &= C_0 + C_1 \cos U_j + D_1 \sin U_j \end{aligned} \tag{7.2}$$

for $j = 1,2,...,n$, $\boldsymbol{\delta}_j = \begin{pmatrix} \delta_{jR} \\ \delta_{jI} \end{pmatrix}$ are independently distributed from the bivariate complex Gaussian distribution with mean $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and covariance matrix $\boldsymbol{\Sigma}_u = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_1^2 \end{pmatrix}$ and $\boldsymbol{\varepsilon}_j = \begin{pmatrix} \varepsilon_{jR} \\ \varepsilon_{jI} \end{pmatrix}$ are also independently distributed from the bivariate complex Gaussian distribution, with mean $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and covariance matrix $\boldsymbol{\Sigma}_v = \begin{pmatrix} \sigma_2^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}$. Therefore, $\boldsymbol{\delta}_j$ and $\boldsymbol{\varepsilon}_j$ can be written as follows:

$$\boldsymbol{\delta}_j = \left(\cos u_j - \cos U_j\right) + i\left(\sin u_j - \sin U_j\right)$$

$$\begin{aligned} \boldsymbol{\varepsilon}_j &= \left(\cos v_j - \cos V_j\right) + i\left(\sin v_j - \sin V_j\right) \\ &= \left(\cos v_j - A_0 - A_1 \cos U_j - B_1 \sin U_j\right) \\ &\quad + i\left(\sin v_j - C_0 - C_1 \cos U_j - D_1 \sin U_j\right). \end{aligned} \tag{7.3}$$

We will estimate the parameters using the maximum likelihood estimation (MLE) method to be described in the following section.

7.3 Estimation of the JS Circular Functional Relationship Model

The parameters $A_0, A_1, B_1, C_0, C_1, D_1, \sigma_1^2$ of the model are to be estimated using the maximum likelihood estimation (MLE) method as described in Section 3.3.2 assuming that the ratio of two error variances is known with $\frac{\sigma_2^2}{\sigma_1^2} = \gamma$, γ is a constant. This constraint is important in order to overcome the problem of unbounded log likelihood function. Subsequently, using the estimates, the circular variable $\boldsymbol{U}$ is obtained. Now, the log likelihood function is given by

$$\begin{aligned}
&\log L\left(A_0, A_1, B_1, C_0, C_1, D_1, \sigma_1^2,\ U_1,\ldots,U_n\ ;\gamma, u_1,\ldots,u_n, v_1,\ldots,v_n\right)\\
&= -2n\log(\pi) - n\log\left(\gamma\sigma_1^2\right) - \frac{1}{\sigma_1^2}\sum_j \left|\delta_j\right|^2 - \frac{1}{\gamma\sigma_1^2}\sum_j \left|\varepsilon_j\right|^2\\
&= -2n\log(\pi) - n\log\left(\gamma\sigma_1^2\right) - \frac{1}{\sigma_1^2}\sum_j \left[\left(\cos u_j - \cos U_j\right)^2 + \left(\sin u_j - \sin U_j\right)^2\right]\\
&\quad - \frac{1}{\gamma\sigma_1^2}\sum_j \left[\left(\cos v_j - A_0 - A_1\cos U_j - B_1 \sin U_j\right)^2\right.\\
&\quad\quad \left. + \left(\sin v_j - C_0 - C_1\cos U_j - D_1 \sin U_j\right)^2\right]. \qquad (7.4)
\end{aligned}$$

Differentiating $\log L$ with respect to parameters $A_0, A_1, B_1, C_0, C_1, D_1, \sigma_1^2$ and U_j, we obtain the MLE estimators $\hat{A}_0, \hat{A}_1, \hat{B}_1, \hat{C}_0, \hat{C}_1, \hat{D}_1, \hat{\sigma}_1^2$ and $\hat{U}_j$ for the parameters given by

$$\hat{A}_0 = \frac{1}{n}\sum_j \left(\cos v_j - \hat{A}_1 \cos\ \hat{U}_j - \hat{B}_1 \sin \hat{U}_j\right)$$

$$\hat{A}_1 = \frac{\sum_j \left(\cos v_j - \hat{A}_0 - \hat{B}_1 \sin \hat{U}_j\right)}{\sum_j \left(\cos \hat{U}_j\right)}$$

$$\hat{B}_1 = \frac{\sum_j \left(\cos v_j - \hat{A}_0 - \hat{A}_1 \cos\ \hat{U}_j\right)}{\sum_j \left(\sin\ \hat{U}_j\right)}$$

$$\hat{C}_0 = \frac{1}{n}\sum_j \left(\sin v_j - \hat{C}_1 \cos \hat{U}_j - \hat{D}_1 \sin \hat{U}_j\right)$$

$$\hat{C}_1 = \frac{\sum_j \left(\sin v_j - \hat{C}_0 - \hat{D}_1 \sin \hat{U}_j\right)}{\sum_j \left(\cos \hat{U}_j\right)}$$

$$\hat{D}_1 = \frac{\sum_j \left(\sin v_j - \hat{C}_0 - \hat{C}_1 \cos \hat{U}_j\right)}{\sum_j \left(\sin \hat{U}_j\right)}$$

and

$$\hat{\sigma}_1^2 = \frac{1}{n}\left[\begin{array}{l} \sum_j \left[\left(\cos u_j - \cos \hat{U}_j\right)^2 + \left(\sin u_j - \sin \hat{U}_j\right)^2\right] \\ \qquad + \frac{1}{\gamma}\sum_j \left(\cos v_j - \hat{A}_0 - \hat{A}_1 \cos \hat{U}_j - B_1 \sin \hat{U}_j\right)^2 \\ \qquad\qquad + \frac{1}{\gamma}\sum_j \left(\sin v_j - \hat{C}_0 - \hat{C}_1 \cos \hat{U}_j - \hat{D}_1 \sin \hat{U}_j\right)^2 \end{array}\right].$$

while

$$\hat{U}_j = \tan^{-1}\left(\frac{\sin \hat{U}_j}{\cos \hat{U}_j}\right), \qquad (7.5)$$

where

$$\cos \hat{U}_j = \sum_j \frac{\cos u_j \sin \hat{U}_j}{\sin u_j} + \frac{W_1}{2\hat{A}_1\hat{B}_1} + \frac{W_2}{2\hat{C}_1\hat{D}_1}$$

$$\sin \hat{U}_j = \sum_j \frac{\sin u_j \cos \hat{U}_j}{\cos u_j} + \frac{W_3}{2\hat{A}_1\hat{B}_1} + \frac{W_4}{2\hat{C}_1\hat{D}_1}$$

$$W_1 = \sum_j \left(\begin{array}{l} \sqrt{\begin{array}{l} \left(\hat{B}_1^4 + 6\hat{A}_1^2 + \hat{B}_1^2 + \hat{A}_1^4\right)\sin \hat{U}_j^2 \\ \quad + \left[\left(6\hat{A}_0 - 6\sin v_j\right)\hat{A}_1\hat{B}_1^{\,2} + \left(2\hat{A}_0 - 2\sin v_j\right)\hat{A}_1^{\,3}\right]\sin \hat{U}_j \\ + \left(\hat{A}_0^2 - 2\sin v_j \hat{A}_0 + \sin v_j^{\,2}\right)\hat{A}_1^2 \end{array}} \\ + \left(\hat{B}_1^2 + \hat{A}_1^2\right)\sin \hat{U}_j + \left(\hat{A}_0 - \sin v_j\right)\hat{A}_1 \end{array}\right)$$

$$W_2 = \sum_j \left(\begin{array}{l} \sqrt{\begin{array}{l} \left(\hat{D}_1^4 + 6\hat{C}_1^2 + \hat{D}_1^2 + \hat{C}_1^4\right)\sin\hat{U}_j^2 \\ + \left[\left(6\hat{C}_0 - 6\sin v_j\right)\hat{C}_1\hat{D}_1^2 + \left(2\hat{C}_0 - 2\sin v_j\right)\hat{C}_1^3\right]\sin\hat{U}_j \\ + \left(\hat{C}_0^2 - 2\sin v_j\hat{C}_0 + \sin v_j^2\right)\hat{C}_1^2 \end{array}} \\ + \left(\hat{D}_1^2 + \hat{C}_1^2\right)\sin\hat{U}_j + \left(\hat{C}_0 - \sin v_j\right)\hat{C}_1 \end{array} \right)$$

$$W_3 = \sum_j \left(\begin{array}{l} \sqrt{\begin{array}{l} \left(\hat{B}_1^4 + 6\hat{A}_1^2 + \hat{B}_1^2 + \hat{A}_1^4\right)\cos\hat{U}_j^2 \\ + \left[\left(6\hat{A}_0 - 6\cos v_j\right)\hat{A}_1\hat{B}_1^2 + \left(2\hat{A}_0 - 2\cos v_j\right)\hat{A}_1^3\right]\cos\hat{U}_j \\ + \left(\hat{A}_0^2 - 2\cos v_j\hat{A}_0 + \cos v_j^2\right)\hat{A}_1^2 \end{array}} \\ + \left(\hat{B}_1^2 + \hat{A}_1^2\right)\cos\hat{U}_j + \left(\hat{A}_0 - \cos v_j\right)\hat{A}_1 \end{array} \right)$$

$$W_4 = \sum_j \left(\begin{array}{l} \sqrt{\begin{array}{l} \left(\hat{D}_1^4 + 6\hat{C}_1^2 + \hat{D}_1^2 + \hat{C}_1^4\right)\cos\hat{U}_j^2 \\ + \left[\left(6\hat{C}_0 - 6\cos v_j\right)\hat{C}_1\hat{D}_1^2 + \left(2\hat{C}_0 - 2\cos v_j\right)\hat{C}_1^3\right]\cos\hat{U}_j \\ + \left(\hat{C}_0^2 - 2\cos v_j\hat{C}_0 + \cos v_j^2\right)\hat{C}_1^2 \end{array}} \\ + \left(\hat{D}_1^2 + \hat{C}_1^2\right)\cos\hat{U}_j + \left(\hat{C}_0 - \cos v_j\right)\hat{C}_1 \end{array} \right).$$

Kendall & Stuart (1979) suggested to multiply $\hat{\sigma}_1^2$ with $\frac{2n}{n-2}$ to give consistent estimate of the parameter σ_1^2. The ML estimates of the model can now be obtained iteratively. The initial values of the iterative procedure are determined by searching the set of parameters values $\hat{\lambda}_{ini} = \left(\hat{A}_{0,ini}, \hat{A}_{1,ini}, \hat{B}_{1,ini}, \hat{C}_{0,ini}, \hat{C}_{1,ini}, \hat{D}_{1,ini}\right)$ which give a maximum likelihood function in the regions $-1 \le \hat{\lambda}_{i,ini} \le 1$, $i = 1,\ldots,6$, and $0 < \hat{\sigma}_1^2 < 1$. On the other hand, due to the complexity of the estimators, we obtained the standard errors of the parameters using the bootstrap method. This will be covered in Section 7.4.

7.4 Finding Standard Errors using Bootstrapping method

Bootstrapping method was first introduced by Efron (1979) and it is a very useful alternative method to estimate the parameters and variance-covariance matrix of the parameters. It is carried out through the process of resampling method with replacement from the data set. We will apply the approach in finding the standard errors of parameters of the JS circular functional relationship model.

Assume that we have a data set with observed values of two variables u and v of size n. The estimated values $\hat{A}_0, \hat{A}_1, \hat{B}_1, \hat{C}_0, \hat{C}_1, \hat{D}_1$ and $\hat{\sigma}_1^2$ of model parameters $(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6, \lambda_7) = (A_0, A_1, B_1, C_0, C_1, D_1, \sigma_1^2)$ are obtained using the MLE method described in Section 7.3 assuming $\sigma_1^2 = \sigma_2^2$. Hence, we can find the residuals $\hat{\delta}_j$ and $\hat{\varepsilon}_j$, $j = 1,2,...,n$. The steps of the bootstrapping method of finding the standard errors of the MLE estimates are described below:

(i) Sampling with replacement is carried out from the residuals $\hat{\delta}_j$ and $\hat{\varepsilon}_j$, $j = 1,2,...,n$ giving a bootstrap sample of size n, $D^{*(1)} = \left(\left(\hat{\delta}_1^*, \hat{\varepsilon}_1^*\right), \ldots, \left(\hat{\delta}_n^*, \hat{\varepsilon}_n^*\right)\right)$.

This is repeated B times giving B sets of bootstrap samples $D^{*(1)}$, …, $D^{*(\mathrm{B})}$.

(ii) For the jth bootstrap sample, j=1,…,B, we calculate the MLE estimates $(\lambda_1^{*(j)}, \lambda_2^{*(j)}, \ldots, \lambda_7^{*(j)}) = (A_0^{*(j)}, A_1^{*(j)}, B_1^{*(j)}, C_0^{*(j)}, C_1^{*(j)}, D_1^{*(j)}, \sigma_1^{2*(j)})$.

(iii) Then, calculate the sample standard deviation of $\hat{\lambda}_i, i = 1, \ldots, 7$, as given by

$$\hat{\sigma}^2_{\lambda_i,B} = \sqrt{\frac{\sum_{j=1}^{B}\left(\lambda_i^{*(j)} - \overline{\lambda}_{i,B}\right)^2}{B-1}}\,, \tag{7.6}$$

where $\overline{\lambda}_{i,B} = \frac{\sum_{\mathrm{j}=1}^{B} \lambda_i^{*(j)}}{B}$.

Hence, we obtain the bootstrap estimates for the model parameters using equation (7.6).

7.5 Performance of the Functional vs Non-functional Models.

In this section, we first investigate the performance of the proposed JS circular functional relationship model (functional) in estimating the parameters of the JS circular functional relationship model described in Section 7.3 for various value of $\left(\sigma_1^2, \sigma_2^2\right)$. Second, we compare the performance of the functional model with the JS circular regression model (non-functional) as described in Section 3.3.

7.5.1 Simulation Study

A simulation studies was carried out in order to assess the performance of the JS circular functional relationship model and JS circular regression models. Sample sizes of *n* were generated. We consider two factors in the simulation; the first is the sample size *n*=20, 30, 70 100 and 130 and the second is the values of $\left(\sigma_1^2, \sigma_2^2\right)$, $\sigma_1^2 = \sigma_2^2$, where σ_1^2 are taken as 0.01, 0.09, and 0.36. The parameters of both JS circular functional relationship models and JS circular regression models are given by

$$\begin{aligned} \boldsymbol{\lambda} &= \left(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6, \lambda_7\right) \\ &= \left(A_0, A_1, B_1, C_0, C_1, D_1, \sigma_1^2\right). \end{aligned}$$

The performances of the functional and non-functional models are now assessed by the following steps:

(i) The parameters of the unreplicated JS circular functional relationship model and JS circular regression models are fixed as employed in the previous chapters based on the choice of constant *a*.

(ii) Generate fixed $\boldsymbol{U}$ variable of size *n* from von Mises distribution $VM(\pi/4,3)$.

(iii) For functional model, calculate the observed values of the response variable $\boldsymbol{V}$ using equation (7.2).

(iv) Obtain the variable $\boldsymbol{V}_j = \tan^{-1}\left(\dfrac{\sin \boldsymbol{V}_j}{\cos \boldsymbol{V}_j}\right), \; j=1,\ldots,n$.

(v) Generate δ_j and ε_j of size *n* from $\boldsymbol{N}\left(\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}\sigma_1^2 & 0\\ 0 & \sigma_1^2\end{pmatrix}\right)$ and $\boldsymbol{N}\left(\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}\sigma_1^2 & 0\\ 0 & \sigma_2^2\end{pmatrix}\right)$, respectively.

(vi) Obtain the simulated data containing variables *u* and *v* using equation (7.1).

(vii) Fit the JS circular functional relationship model on the simulated data giving the parameter estimates $\hat{A}_0, \hat{A}_1, \hat{B}_1, \hat{C}_0, \hat{C}_1, \hat{D}_1$ and $\hat{\sigma}_1^2$.

(viii) For non-functional model, generate $\boldsymbol{\varepsilon}_1$ and $\boldsymbol{\varepsilon}_2$ of size *n* from $\boldsymbol{N}\left(\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}\sigma_1^2 & 0\\ 0 & \sigma_2^2\end{pmatrix}\right)$. Then, calculate V_{1j} and V_{2j}, $j=1,\ldots,n$ using equation (3.6).

(ix) Obtain the variable $v_j = \arctan\left(\dfrac{V_{2j}}{V_{1j}}\right), \; j=1,\ldots,n$.

(x) The generated circular data (u_j, v_j) above is fitted to the JS circular regression model to give the parameter estimates $\hat{A}_0, \hat{A}_1, \hat{B}_1, \hat{C}_0, \hat{C}_1, \hat{D}_1$ and $\hat{\sigma}_1^2$.

(xi) Finally, steps (ii)-(vii) for functional model and steps (viii)-(x) non-functional model are repeated for *simu*=1000 times. For each parameter $(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6, \lambda_7) = (A_0, A_1, B_1, C_0, C_1, D_1, \sigma_1^2)$, we calculate the mean, bias, *SE* and *RMSE* as described in Section 3.4.1.

7.5.2 Discussion

The results for mean, bias, *SE* and *RMSE* of each parameter estimates for functional and non-functional models are tabulated in Tables 7.1 - 7.6 for different values of $\sigma_1^2 = \sigma_2^2$. Several results are observed:

1. For functional model, the estimated means for all parameter estimates are consistently close to the true values. The estimation further improves when the sample size *n* increases.
2. The value of bias is consistently small for all parameter estimates and σ_1^2. The value is closer to zero when the sample size *n* increases. Meanwhile, when the value of σ_1^2 increase, the bias also increases. This is because the data set is more dispersed around the unit circle for larger σ_1^2.
3. The *SE* for all parameter estimates is generally small.
4. The value for *RMSE* of each parameter estimates decreases when the value of *n* increases.

5. For non-functional model, the values of estimated mean also consistently close to the true values. The value of bias, *SE* and *RMSE* also small for all parameter estimates.
6. The value for *RMSE* of each parameter estimates also decreases when the value of n increases. Generally, the values of non-functional model are smaller than functional model for all the parameter estimates.

By looking the above results, the maximum likelihood estimation method performs well in estimating the parameters of the JS circular functional relationship model based on mean, bias, *SE* and *RMSE*. However, the method is affected when the value of σ_1^2 gets larger.

Table 7.1: Parameter estimates of functional model for $\sigma_1^2 = 0.01$.

n	20					30				70			
Estimates	true value	mean	bias	*RMSE*	*SE*	mean	bias	*RMSE*	*SE*	mean	bias	*RMSE*	*SE*
$\hat{A}_0$	0.0000	-0.0148	-0.0148	0.4693	0.1758	-0.0099	-0.0099	0.3145	0.0999	-0.0020	-0.0020	0.0617	0.0501
$\hat{A}_1$	-0.4161	-0.4035	0.0126	0.3985	0.1394	-0.4062	0.0099	0.3125	0.0741	-0.4126	0.0035	0.1120	0.0385
$\hat{B}_1$	-0.9093	-0.8899	0.0194	0.6129	0.1330	-0.8934	0.0159	0.5039	0.0883	-0.9000	0.0093	0.2955	0.0455
$\hat{C}_0$	0.0000	-0.0333	-0.0333	1.0518	0.2396	-0.0178	-0.0178	0.5643	0.1282	-0.0086	-0.0086	0.2705	0.0380
$\hat{C}_1$	0.9093	0.9320	0.0227	0.7170	0.2017	0.9228	0.0135	0.4278	0.1080	0.9140	0.0047	0.1492	0.0376
$\hat{D}_1$	-0.4161	-0.3851	0.0310	0.9788	0.1730	-0.3967	0.0194	0.6132	0.1004	-0.4039	0.0122	0.3843	0.0375
$\hat{\sigma}_1^2$	0.0100	0.0207	0.0107	0.0248	0.0190	0.0225	0.0125	0.1582	0.0157	0.0239	0.0139	0.2596	0.0102
n	100					130							
Estimates	true value	mean	bias	*RMSE*	*SE*	mean	bias	*RMSE*	*SE*				
$\hat{A}_0$	0.0000	-0.0030	-0.0030	0.0934	0.0399	-0.0045	-0.0045	0.1006	0.0355				
$\hat{A}_1$	-0.4161	-0.4132	0.0029	0.0923	0.0315	-0.4103	0.0058	0.1300	0.0260				
$\hat{B}_1$	-0.9093	-0.8988	0.0105	0.3331	0.0364	-0.8986	0.0107	0.2394	0.0343				
$\hat{C}_0$	0.0000	-0.0088	-0.0088	0.2796	0.0270	-0.0064	-0.0064	0.1431	0.0234				
$\hat{C}_1$	0.9093	0.9139	0.0046	0.1455	0.0281	0.9127	0.0034	0.0764	0.0240				
$\hat{D}_1$	-0.4161	-0.4048	0.0113	0.3585	0.0277	-0.4057	0.0104	0.2326	0.0241				
$\hat{\sigma}_1^2$	0.0100	0.0242	0.0142	0.2801	0.0083	0.0245	0.0145	0.2121	0.0073				

Table 7.2: Parameter estimates of functional model for $\sigma_1^2 = 0.09$.

n		20				30				70			
Estimates	true value	mean	bias	*RMSE*	*SE*	mean	bias	*RMSE*	*SE*	mean	bias	*RMSE*	*SE*
$\hat{A}_0$	0.0000	-0.0217	-0.0217	0.4851	0.8318	-0.0489	-0.0489	1.5474	0.7453	-0.0084	-0.0084	0.2670	0.2419
$\hat{A}_1$	-0.4161	-0.3838	0.0323	0.7216	0.6243	-0.3657	0.0504	1.5928	0.5685	-0.3963	0.0198	0.6269	0.1893
$\hat{B}_1$	-0.9093	-0.8115	0.0978	2.1877	0.6285	-0.7883	0.1210	3.8249	0.5799	-0.8250	0.0843	2.6655	0.2117
$\hat{C}_0$	0.0000	-0.0954	-0.0954	2.1333	0.7625	-0.0787	-0.0787	2.4878	1.1126	-0.0544	-0.0544	1.7190	0.2261
$\hat{C}_1$	0.9093	0.9521	0.0428	0.9560	0.6310	0.9345	0.0252	0.7964	0.8762	0.9155	0.0062	0.1968	0.2083
$\hat{D}_1$	-0.4161	-0.3181	0.0980	2.1917	0.5966	-0.3113	0.1048	3.3128	0.8249	-0.3451	0.0710	2.2450	0.2041
$\hat{\sigma}_1^2$	0.0900	0.1360	0.0460	0.9367	0.0581	0.1394	0.0494	1.2211	0.0615	0.1542	0.0642	0.7954	0.0398

n		100				130			
Estimates	true value	mean	bias	*RMSE*	*SE*	mean	bias	*RMSE*	*SE*
$\hat{A}_0$	0.0000	-0.0293	-0.0293	0.9261	0.1935	-0.0285	-0.0285	0.6375	0.1589
$\hat{A}_1$	-0.4161	-0.3821	0.0340	1.0743	0.1489	-0.3742	0.0419	0.9380	0.1191
$\hat{B}_1$	-0.9093	-0.8104	0.0989	3.1274	0.1712	-0.8167	0.0926	2.0700	0.1518
$\hat{C}_0$	0.0000	-0.0520	-0.0520	1.6438	0.1583	-0.0444	-0.0444	0.9928	0.1246
$\hat{C}_1$	0.9093	0.9127	0.0034	0.1076	0.1522	0.9088	-0.0005	0.0101	0.1208
$\hat{D}_1$	-0.4161	-0.3493	0.0668	2.1138	0.1453	-0.3494	0.0667	1.4908	0.1159
$\hat{\sigma}_1^2$	0.0900	0.1560	0.0660	0.7436	0.0332	0.1593	0.0693	0.4603	0.0277

Table 7.3: Parameter estimates of functional model for $\sigma_1^2 = 0.36$.

n		20				30				70			
Estimates	true value	mean	bias	*RMSE*	*SE*	mean	bias	*RMSE*	*SE*	mean	bias	*RMSE*	*SE*
$\hat{A}_0$	0.0000	-0.0152	-0.0152	0.3397	0.1639	0.1445	0.1445	4.5694	4.7053	-0.0157	-0.0157	0.4972	0.8385
$\hat{A}_1$	-0.4161	-0.4037	0.0124	0.2777	0.1289	-0.4268	-0.0107	0.3378	4.0166	-0.2749	0.1412	4.4658	0.6892
$\hat{B}_1$	-0.9093	-0.8884	0.0209	0.4684	0.1274	-0.6570	0.2523	7.9784	2.5972	-0.5853	0.3240	10.2465	0.6882
$\hat{C}_0$	0.0000	-0.0227	-0.0227	0.5070	0.2402	-0.2130	-0.2130	6.7370	2.9078	-0.0687	-0.0687	2.1709	1.1877
$\hat{C}_1$	0.9093	0.9243	0.0150	0.3351	0.2035	0.7794	-0.1299	4.1093	2.2002	0.6586	-0.2507	7.9273	0.9052
$\hat{D}_1$	-0.4161	-0.3928	0.0233	0.5203	0.1744	-0.1129	0.3032	9.5880	2.1408	-0.2123	0.2038	6.4452	1.0030
$\hat{\sigma}_1^2$	0.3600	0.0211	-0.3389	0.0351	0.0195	0.4257	0.0657	4.5310	0.1089	0.4640	0.1040	3.8957	0.0859

n		100				130			
Estimates	true value	mean	bias	*RMSE*	*SE*	mean	bias	*RMSE*	*SE*
$\hat{A}_0$	0.0000	-0.0030	-0.0030	0.0955	0.6787	-0.0349	-0.0349	0.7805	0.4496
$\hat{A}_1$	-0.4161	-0.2904	0.1257	3.9759	0.5726	-0.2441	0.1720	3.8454	0.3742
$\hat{B}_1$	-0.9093	-0.5815	0.3278	10.3661	0.5426	-0.5801	0.3292	7.3615	0.4041
$\hat{C}_0$	0.0000	-0.0573	-0.0573	1.8118	0.5770	-0.0640	-0.0640	1.4319	0.4566
$\hat{C}_1$	0.9093	0.6675	-0.2418	7.6453	0.5045	0.6762	-0.2331	5.2124	0.3953
$\hat{D}_1$	-0.4161	-0.2452	0.1709	5.4042	0.5096	-0.2248	0.1913	4.2766	0.4102
$\hat{\sigma}_1^2$	0.3600	0.4741	0.1141	3.7320	0.0713	0.4946	0.1346	2.4083	0.0618

Table 7.4: Parameter estimates of non-functional model for $\sigma_1^2 = 0.01$.

n	20					30				70			
Estimates	true value	mean	bias	*RMSE*	*SE*	mean	bias	*RMSE*	*SE*	mean	bias	*RMSE*	*SE*
$\hat{A}_0$	0.0000	-0.0001	-0.0001	0.0026	0.0083	0.0001	0.0001	0.0046	0.0066	0.0000	0.0000	0.0016	0.0035
$\hat{A}_1$	-0.4161	-0.4161	0.0000	0.0000	0.0064	-0.4163	-0.0002	0.0051	0.0051	-0.4162	0.0000	0.0002	0.0028
$\hat{B}_1$	-0.9093	-0.9091	0.0002	0.0057	0.0073	-0.9094	-0.0001	0.0020	0.0058	-0.9091	0.0001	0.0047	0.0032
$\hat{C}_0$	0.0000	0.0001	0.0001	0.0021	0.0080	-0.0001	-0.0001	0.0039	0.0053	-0.0002	-0.0002	0.0049	0.0027
$\hat{C}_1$	0.9093	0.9092	-0.0001	0.0028	0.0073	0.9093	-0.0001	0.0007	0.0052	0.9094	0.0001	0.0031	0.0028
$\hat{D}_1$	-0.4161	-0.4161	0.0000	0.0003	0.0070	-0.4160	0.0001	0.0040	0.0048	-0.4160	0.0001	0.0034	0.0026
$\hat{\sigma}_1^2$	0.0100	0.0001	0.2486	13.3361	0.0015	0.0001	-0.0099	13.3375	0.0012	0.0001	-0.0099	13.3410	0.0008

n	100					130			
Estimates	true value	mean	bias	*RMSE*	*SE*	mean	bias	*RMSE*	*SE*
$\hat{A}_0$	0.0000	0.0001	0.0001	0.0022	0.0028	0.0000	0.0000	0.0014	0.0024
$\hat{A}_1$	-0.4161	-0.4162	0.0000	0.0003	0.0023	-0.4162	0.0000	0.0005	0.0020
$\hat{B}_1$	-0.9093	-0.9093	0.0000	0.0002	0.0026	-0.9093	0.0000	0.0006	0.0022
$\hat{C}_0$	0.0000	-0.0001	-0.0001	0.0036	0.0026	0.0000	0.0000	0.0011	0.0018
$\hat{C}_1$	0.9093	0.9094	0.0001	0.0023	0.0023	0.9092	-0.0001	0.0018	0.0019
$\hat{D}_1$	-0.4161	-0.4161	0.0001	0.0024	0.0021	-0.4162	0.0000	0.0005	0.0018
$\hat{\sigma}_1^2$	0.0100	0.0001	-0.0099	13.3409	0.0006	0.0001	-0.0099	13.3432	0.0006

Table 7.5: Parameter estimates of non-functional model for $\sigma_1^2 = 0.09$.

n	20					30				70			
Estimates	true value	mean	bias	*RMSE*	*SE*	mean	bias	*RMSE*	*SE*	mean	bias	*RMSE*	*SE*
$\hat{A}_0$	0.0000	-0.0103	-0.0103	0.3259	0.0941	-0.0047	-0.0047	0.1482	0.0663	-0.0050	-0.0050	0.1569	0.0365
$\hat{A}_1$	-0.4161	-0.4097	0.0064	0.2033	0.0721	-0.4133	0.0029	0.0907	0.0499	-0.4126	0.0035	0.1121	0.0280
$\hat{B}_1$	-0.9093	-0.8924	0.0169	0.5332	0.0820	-0.8985	0.0108	0.3415	0.0594	-0.8982	0.0111	0.3498	0.0342
$\hat{C}_0$	0.0000	-0.0157	-0.0157	0.4954	0.0986	-0.0120	-0.0120	0.3791	0.0588	-0.0073	-0.0073	0.2319	0.0289
$\hat{C}_1$	0.9093	0.9197	0.0104	0.3290	0.0832	0.9161	0.0068	0.2157	0.0545	0.9130	0.0037	0.1168	0.0284
$\hat{D}_1$	-0.4161	-0.4012	0.0150	0.4733	0.0882	-0.4035	0.0127	0.4001	0.0538	-0.4076	0.0085	0.2688	0.0293
$\hat{\sigma}_1^2$	0.0100	0.0054	0.2539	14.7884	0.0145	0.0056	-0.0844	14.8134	0.0118	0.0058	-0.0842	14.8385	0.0077

n	100					130			
Estimates	true value	mean	bias	*RMSE*	*SE*	mean	bias	*RMSE*	*SE*
$\hat{A}_0$	0.0000	-0.0031	-0.0031	0.0978	0.0293	-0.0033	-0.0033	0.1044	0.0254
$\hat{A}_1$	-0.4161	-0.4133	0.0028	0.0889	0.0230	-0.4132	0.0029	0.0932	0.0200
$\hat{B}_1$	-0.9093	-0.8998	0.0095	0.2997	0.0268	-0.8999	0.0094	0.2983	0.0235
$\hat{C}_0$	0.0000	-0.0070	-0.0070	0.2216	0.0223	-0.0056	-0.0056	0.1756	0.0184
$\hat{C}_1$	0.9093	0.9134	0.0041	0.1303	0.0227	0.9121	0.0028	0.0877	0.0193
$\hat{D}_1$	-0.4161	-0.4089	0.0073	0.2300	0.0226	-0.4094	0.0067	0.2127	0.0184
$\hat{\sigma}_1^2$	0.0100	0.0058	-0.0842	14.8476	0.0062	0.0059	-0.0841	14.8604	0.0054

Table 7.6: Parameter estimates of non-functional model for $\sigma_1^2 = 0.36$.

n	20					30				70			
Estimates	true value	mean	bias	*RMSE*	*SE*	mean	bias	*RMSE*	*SE*	mean	bias	*RMSE*	*SE*
$\hat{A}_0$	0.0000	-0.0455	-0.0455	1.4397	0.6277	-0.0441	-0.0441	1.3951	0.4193	-0.0269	-0.0269	0.8505	0.2069
$\hat{A}_1$	-0.4161	-0.3657	0.0505	1.5962	0.4819	-0.3561	0.0600	1.8974	0.3190	-0.3707	0.0455	1.4380	0.1628
$\hat{B}_1$	-0.9093	-0.7562	0.1531	4.8426	0.5130	-0.7670	0.1423	4.5008	0.3646	-0.7808	0.1285	4.0627	0.1889
$\hat{C}_0$	0.0000	-0.1295	-0.1295	4.0967	0.7919	-0.0712	-0.0712	2.2527	0.4199	-0.0521	-0.0521	1.6470	0.1732
$\hat{C}_1$	0.9093	0.9478	0.0385	1.2189	0.6347	0.8966	-0.0127	0.4018	0.3749	0.8802	-0.0291	0.9211	0.1710
$\hat{D}_1$	-0.4161	-0.2786	0.1375	4.3485	0.6616	-0.3175	0.0987	3.1207	0.3565	-0.3327	0.0834	2.6373	0.1612
$\hat{\sigma}_1^2$	0.0100	0.1130	0.4777	20.5999	0.1027	0.1227	-0.2373	20.9131	0.0892	0.1280	-0.2320	21.0802	0.0548
n	100					130							
Estimates	true value	mean	bias	*RMSE*	*SE*	mean	bias	*RMSE*	*SE*				
$\hat{A}_0$	0.0000	-0.0224	-0.0224	0.7090	0.1628	-0.0243	-0.0243	0.7694	0.1361				
$\hat{A}_1$	-0.4161	-0.3692	0.0470	1.4858	0.1319	-0.3681	0.0480	1.5184	0.1087				
$\hat{B}_1$	-0.9093	-0.7846	0.1247	3.9428	0.1461	-0.7845	0.1248	3.9464	0.1273				
$\hat{C}_0$	0.0000	-0.0466	-0.0466	1.4749	0.1335	-0.0415	-0.0415	1.3134	0.1125				
$\hat{C}_1$	0.9093	0.8797	-0.0296	0.9363	0.1314	0.8726	-0.0367	1.1620	0.1096				
$\hat{D}_1$	-0.4161	-0.3441	0.0720	2.2774	0.1275	-0.3429	0.0733	2.3178	0.1086				
$\hat{\sigma}_1^2$	0.0100	0.1299	-0.2301	21.1385	0.0463	0.2550	-0.3450	21.2250	0.0417				

7.6 Practical Example: Eye Data

In this section, we will consider the eye data again as given in Section 3.5. We consider the eye data consisting of 23 observations of glaucoma patients (unit in radians) at the University Malaya Medical Centre (UMMC). The measurements selected are the angle of the posterior corneal curvature (u) and the angle of the eye (between posterior corneal curvature to iris) (v). However, we assume now that both the dependent variable v and the exploratory variable u are subject to errors. That is, we assume that, for observation j, v_j is assumed to be the observed angle of the eye, $\boldsymbol{V}_j$, recorded with some random error, $\boldsymbol{\varepsilon}_j$. Similarly, u_j measures the underlying direction measured by posterior corneal curvature, $\boldsymbol{U}_j$, with some random error, $\boldsymbol{\delta}_j$.

The JS circular functional relationship model is described by equation (7.1) such that

$$(\cos u_j + i\sin u_j) = (\cos U_j + i\sin U_j) + (\delta_{jR} + i\,\delta_{jI})$$

$$(\cos v_j + i\sin v_j) = (\cos V_j + i\sin V_j) + (\varepsilon_{jR} + i\,\varepsilon_{jI}).$$

The unreplicated JS circular functional relationship model is used to fit the eye data. The final fitted model is obtained giving

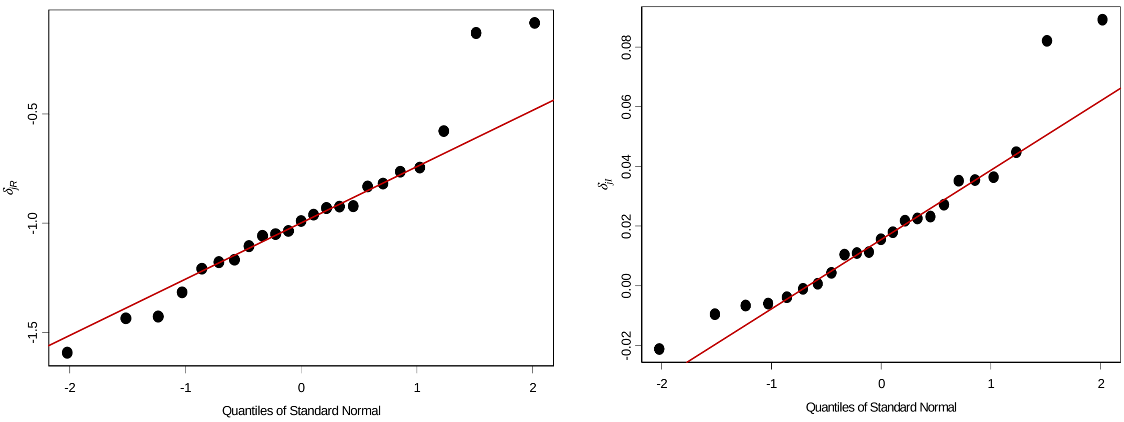
$$\cos V_j = 1.0819 - 0.1571\cos U_j - 0.3716\sin U_j$$

$$\sin V_j = 0.0948 + 0.2334\cos U_j + 0.5955\sin U_j$$

by assuming that the ratio of error variances is known and equal to one. The parameter estimates for eye data is given in Table 7.7. Note that the ML estimates of the JS circular functional relationship model are quite close to the LS estimates of the JS circular regression model given in Section 3.5.1. However, the estimate $\hat{\sigma}_1^2$ = 0.0446 for our present model is much smaller than that of the JS circular regression model, which is 0.16. In other word, allowing the explanatory variable *u* to be subjected to error has reduced the variability of the model. On the other hand, the standard errors of the parameters are obtained using the bootstrap method and are given in the forth column of Table 7.4. The values are reasonably small compared to LS estimates.

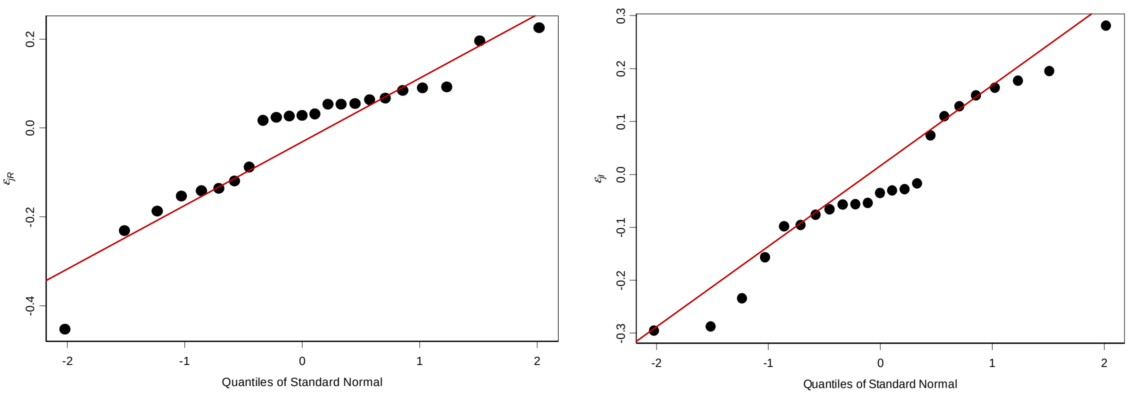
Table 7.7: Estimates for eye data with two variables

Estimates	**LS estimates**	**Standard error**	**ML estimates**	**Standard error**
$\hat{A}_0$	1.0822	0.2664	1.0819	0.2113
$\hat{A}_1$	-0.1497	0.1026	-0.1571	0.0188
$\hat{B}_1$	-0.3836	0.2873	-0.3716	0.0068
$\hat{C}_0$	0.0986	0.2776	0.0948	0.1964
$\hat{C}_1$	0.2534	0.1070	0.2334	0.0317
$\hat{D}_1$	0.5935	0.2994	0.5955	0.0092
$\hat{\sigma}_1^2$	0.16	0.1635	0.0446	0.7726



(a) δ_{jR} (c) δ_{jI}

Figure 7.1: Q-Q plot for residuals of $\boldsymbol{\delta}_j$



(a) ε_{jR} (b) ε_{jI}

Figure 7.2: Q-Q plot for residuals of $\boldsymbol{\varepsilon}_j$

The Q-Q plots for the residuals of the JS circular functional relationship model from equation (7.1) is shown in Figures 7.1 - 7.2. The plot for δ_{jR} and δ_{jI} shows that most of the points are close to the straight line except two points at the top right of the plot. Meanwhile, for ε_{jR} and ε_{jI}, the plots shows that most of the points are close to the straight line except one point at the lower left of the plot for ε_{jR}. In fact, these plots are better than that seen in Figure 3.3 for the JS circular regression model. We conclude that the model has fitted the data quite well by assuming both variables of the unreplicated JS circular functional relationship model are subject to errors.

7.7 Summary

In this chapter, we extend the linear functional relationship model to include the JS circular regression models in the set up. Here, both variables v and u are subject to errors. The maximum likelihood estimates are obtained by assuming a known ratio of error variances. The performance of maximum likelihood estimates performs well in estimating the parameters of JS circular functional relationship model. The standard errors of the parameters are obtained using bootstrap method. As an application, the eye data set has been fitted using the proposed model and give a very encouraging results.

CHAPTER EIGHT

CONCLUSIONS

8.1 Summary

This study looks at some problems related to circular regression models. Few published work can be found on the problem of outliers and none on the study of the relationship between a dependent circular variable and multiple explanatory circular variables. In this study, we specifically choose the JS circular regression models proposed by Jammalamadaka & Sarma (1993) due to its interesting properties which have very close resemblance to that of the multiple linear regression models. We look at three problems associated to the JS circular regression models.

Firstly, we look at the problem of identifying outliers in the JS circular regression model. The parameter estimates of the model can be obtained using the least squares method. However, the LS estimates are shown to be sensitive to the occurrence of outliers. Hence, we use three different statistics to detect outliers based on the row deletion approach; the *COVRATIO*, *DMCEc* and *DMCEs* statistics. The cut-off points are obtained via simulation. Due to large number of factors to be taken care of, the cut-off point for real data set can be obtained by running the special program prepared for each approach as given in Appendix 5 and 6 respectively. From the cases considered, we conclude that the performance of the outlier detection procedures are good, with the procedure of *DMCEc* is found to be superior than that based on *DMCEs* and for small sample size. *DMCEc* also is found to be the superior methods compared to *DMCEs* and

COVRATIO statistics. For illustration, we apply the procedures on two real data sets; the wind data set and the eye data set.

Secondly, we propose the generalized version of the JS circular regression model to include more than one explanatory variables. We call the extended model as the generalized JS circular regression model. The real and imaginary parts of the model are now estimated using the trigonometric polynomial with more than one variables. Via simulation, we show that the generalized model is estimated well by the LS estimation method. We then investigate the problem of multicollinearity in the generalized JS circular regression model. A new procedure of detecting the problem based on *VIF* is proposed to suit the nature of the model. We then extend the idea of ridge regression approach in linear case to the circular case. Related theory and procedures to obtain the ridge estimates of the generalized JS circular regression model are presented. Next, we propose to use the ridge regression approach to overcome the multicollinearity problem in the generalized JS circular regression models. We illustrate the proposed method using a multivariate eye data set obtained from the University of Malaya Medical Centre, Malaysia.

Lastly, we develop the theory of the circular functional relationship model for JS circular regression model when both variables are subject to errors. The parameters of the model are derived using maximum likelihood estimation method. Due to a large number of parameters in the model set-up, the bootstrap method is used to obtain the variance-covariance matrix of the parameters. We then illustrate the application of the model using the eye data set.

In conclusion, we have looked at three main research problems that are related to the JS circular regression model proposed by Jammalamadaka & Sarma (1993). The work is very significant in providing more flexible circular modelling options for the practitioners who are working on circular data.

8.2 Contributions

This study has contributed to circular data analysis in the following ways:

1. We have shown that the least squares estimates of JS circular regression models are not robust toward the occurrence of outliers. Thus, it is important to develop relevant methods to identify outliers for further investigation purposes.
2. We have considered three outlier detection procedures to detect outliers in JS circular regression models using row deletion approach, namely, *COVRATIO*, *DMCEc* and *DMCEs* statistics. Via simulation, the procedures perform well in identifying the outliers that exist in the data set.
3. We have proposed a generalized JS circular regression model for accommodating two or more explanatory circular variables in the models. The relevant theory is presented and, via simulation, the model is found to be well estimated by the least squares estimation method.
4. We have looked at the problem of multicollinearity in the generalized JS circular regression model. The relevant procedure of indentifying the problem in the model has been presented. We extend the idea of ridge regression approach in multiple linear regression case to the generalized JS circular regression case to give the ridge estimates of the model.
5. We have developed a new functional relationship model by using the JS circular regression model in the setup. The parameters of the JS circular functional model are estimated using maximum likelihood estimation method. We show that the estimation method perform well when investigated via simulation.

6. We illustrated the works carried out in this study using three different real data sets; the wind data set, eye data set with two variables and another multivariate eye data set with four variables.

8.3 Further Research

There are various possibilities for further research in this area. Some suggestions are given as follows:

(i) to develop some effective procedures to detect multiple outliers as in circular regression models.

(ii) to extend the procedures of the detection of outliers in multiple circular regression models.

(iii) to extend the procedures of the detection of outliers to the circular functional relationship model.

(iv) to extend the procedures of the detection of outliers to the robust techniques in circular cases.

We recognize that there are still many problems ready to be explored in circular statistics for future works.

REFERENCES

Abuzaid, A. H., Mohamed, I. and Hussin, A. G. (2008). Identifying single outlier in linear circular regression model based on circular distance. *Journal of Applied Probability and Statistics,* **3** (1), 107 – 117.

Abuzaid, A. H., Mohamed, I. and Hussin, A. G. (2009). A New Test of Discordancy in Circular Data. *Journal of Communications in Statistics - Simulation and Computation,* **38** (4), 682 **–** 691.

Abuzaid, A. H. (2009). Some problems of outliers in circular data. (PhD thesis). Institute of Mathematical Science, University of Malaya.

Abuzaid, A. H, Mohamed, I, and Hussin, A. G. And Rambli, A. (2011). *COVRATIO* Statistic for Simple Circular Regression Model, *Chiang Mai J. Sci*., **38**(3), 321-330.

Abuzaid, A. H, Mohamed, I, and Hussin, A. G., (2012). Boxplot for Circular Variables. *Computational Statistics*, **27** (3), pp 381-392.

Abuzaid, A. H., Hussin, A. G. and Mohamed, I. (2013). Detection of Outliers in Simple Circular Regression Model Using the Mean Circular Error Statistic, *Journal of Statistical Computation and Simulation*, **83** (2), 269-277.

Adcock, R. J. (1877). Note on the Method of Least Squares, *Analyst*, **4**, 183-184

Adcock, R. J. (1878). A Problem in Least Squares, *Analyst*, **5**, 53-54.

Agostinelli, C. (2006). Notes on Pearson Residuals and Weighted Likelihood Estimating Equations. *Statist. Probab. Lett*, **76** (17), 1930-1934.

Barnett, V. (1978). The study of outliers: Purpose and model. *Applied Statistics*, **27** (3), 242-250.

Barnett, V. and Lewis, T. (1994). *Outliers in Statistical Data*. John Wiley & Sons, Chichesters.

Batschelet, E. (1981). *Circular Statistics in Biology*. Academic Press, London.

Beckman, R. J. and Cook, R. D. (1983). Outlier……….s. *Technometrics*, **25** (2), 119-149.

Belsley, D. A., Kuh, E. and Welsch, R. E. (1980). *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. John Wiley & Sons, New York.

Best, D. J. and Fisher, N. I. (1981). The bias of the maximum likelihood estimators of the von Mises-Fisher concentration parameters. *Communication in Statistics – Simulations and Computations*. **B10** (5), 394-502.

Binkley, S. (1990). *The clockwork sparrow: time, clock and calendars in Biological Organisms*. Upper Saddle River, NJ: Prentice Hall.

Bowtell, P. and Patefield, M. (1999). Circular functional relationships. *Communication in Statistics- Theory and Methods*, **28** (9), 2103-2118.

Brunsdon, C. and Corcoran, J. (2006). Using Circular Statistics to Analyse Time patterns in crime incidence. *Computer, Environment and Urban Systems*, **30**, 300-319.

Buonaccorsi, J. P. (1996). A Modified Estimating Equation Approach to Correcting for Measurement Error in Regression, *Biometrika*, 433-440.

Cai, T. and Hall, P. (2006). Prediction in functional linear regression. *The Annals of Statistics* **34**, 2159-2179.

Caires, S. and Wyatt, L. R. (2003). A linear functional relationship model for circular data with an application to the assessment of ocean wave measurements. *Journal of Agricultural, Biological and Environmental Statistics*, **8** (2), 153-169.

Chatterjee. S and Hadi. A. S. (1986). Influential Observations, High Leverage Points, and Outliers in Linear Regression. *Statistical Science*, **1** (3), pp. 379-393.

Chatterjee, S., Hadi, A. S. and Price, B. (2000). *Regression Analysis by Example*. 3rd ed., John Wiley & Sons, New York.

Collett, D. (1980). Outliers in circular data. *Applied Statistics*, **29** (1), 50-57.

Cook, R. D. (1977). Detection of influential observations in linear regression. *Technometrics*, **19** (1), 15-19.

Cook, R. D. (1979). Influential observations in linear regression. *Journal of the American Statistical Association*, **74** (365), 169-174.

David, H. A. (1970). *Order Statistics*. Wiley, New York and London.

Dandona, L., Dandona, R., Mandal, P., et al. (2000). Angle-closure glaucoma in an urban population in southern India: the Andhra Pradesh Eye Disease Study. *Ophthalmology*, **107**, 1710–6.

Downs, T., Leibman, J. & Mackay, W. (1970). Statistical methods for vectorcardiogram orientations. In *Proc. XI Int. Vector Cardiography Symposium*, Ed. I. Hoffman, 216-222. Amsterdam: North Holland

Downs, T. (1974). Rotational Angular Correlation. In *Biorhythms and Human Reproduction*, Ed. M. Ferin, F. Halberg and L. van der Wiele, 97-104. New York: Wiley.

Downs, T. D. and Mardia, K. V. (2002). Circular regression. *Biometrika*, **89** (3), 683-697.

Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife Annal of Statistic. **7** (1), 1-26.

Efron, B. and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap.* San Francisco: Chapman & Hall.

Farrar, D. E, and Glauber, R. R., (1967). Multicollinearity in regression analysis: *The problem revisited. Review of Economics and Statistics,* **49**, pp. 92-107

Ferguson, D. E., Landreth, H. F. and Mckeown, J. P. (1967). Sun compass orientation of the northern cricket frog, Acris crepitans. *Animal Behaviour,* **15** (1), 45-53.

Fisher, N. I., Lewis, T. and Embleton, B. J. J. (1987). *Statistical Analysis of Spherical Data.* Cambridge University Press, Cambridge.

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

Fisher, N. I. (1993). *Statistical Analysis of Circular Data*. Cambridge University Press, London.

Foster P. J., Baasanhu J., Alsbirk, P. H., et al. (1996). Glaucoma in Mongolia: a population-based survey in Hövsgöl Province, Northern Mongolia. *Arch Ophthalmol*, **114**, 1235– 41.

Foster P. J., Oen, F.T., Machin, D., et al. (2000). The prevalence of glaucoma in Chinese residents of Singapore: a cross-sectional population survey of the Tanjong Pagar district. *Arch Ophthalmol*, **118**, 1105–11.

Fuller, W.A. (1987). *Measurement error models.* Wiley, New York

Gatto, R. and Jammalamadaka, S. R. (2007). The generalized von Mises distribution. *Statistical Methodology*, 4, 341-353.

Goodman, N. R. (1963). Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction), *Ann. Math. Stat.* **34,** 152-165.

Gordon, A. D., Jupp, P. E. and Byrne, R. W. (1989). Construction and assessment of mental maps. *British Journal of Mathematical and Statistical Psychology*, **42**,169-182.

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

Haan C. T. (2002*) Statistical Methods in Hydrology,* second edition. Iowa State University Press, Ames, Iowa.

Hassan, S. F. Hussin, A. G. and Zubairi, Y. Z. (2010). Estimation of Functional Relationship Model for Circular Variables and Its Application in Measurement Problems. *Chiang Mai J. Sci*., **37** (2), 195-205

He, X. and Simpson, D. G. (1992). Robust Direction Estimation. *The Annals of Statistics*, **20** (1), 351-369

Hoerl, E. and Kennard, R. W. (1970a). Ridge Regression: Biased Estimation for to northogonal Problems. *Technometrics*, **12** (1), pp. 55-67.

Hoerl, E. & Kennard, R. W. (1970b). Ridge Regression. Applications to nonorthogonal problems. *Technometrics* **12** (1), pp. 69-82.

Hoerl, A. E. (1962). Application of Ridge Analysis to Regression Problems. *Chemical Engineering Progress*, **58**, 54-59

Hoerl,A. E. and Kennarrd., W. (1968). On Regression Analysis and Diased Estimation. *Technometrics*, **10**, 422-423. Abstract.

Huber, P. J. (1981). *Robust Statistics*. John Wiley & Sons, New York.

Hussin, A. G., Fieller, N. R. J. and Stillman, E. C. (2004). Linear regression for circular variables with application to directional data. *Journal of Applied Science & Technology*, **8**, (1 & 2), 1-6.

Hussin, A. G. (2004), Linear Regression Model for Circular Variables with Application to Directional Data. *Journal of Applied Science and Technology*, **9** (1 & 2), 1 -6.

Hussin, A. G. (1997). *Pesudo-replication in Functional Relationships with Environmental Applications*. Ph.D. thesis, University of Sheffield.

Hussin, A. G., (1998). The unreplicated complex linear functional relationship model and its application, Bull. Malaysian Math. Soc. (Second Series) **21**, 79-86.

Hussain F, Zubairi, Y. Z. and Hussin, A. G. (2007). An alternative diagrammatical representation of wind data. *Proceedings of the WSEAS Conference on Applied Mathematics*. December, Egypt.

Jacob, A., Thomas, R., Koshi, S.P., et al. (1998). Prevalence of primary glaucoma in an urban south Indian population. *Indian J Ophthalmol*, **46**, 81– 6.

Jammalamadaka, S. R., Bhadra, N., Chaturvedi, D., Kutty, T. K., Majumdar, P. P. and Poduval, G. (1986). Functional assessment of knee and ankle during level walking. In Matusita, K., editor. *Data Analysis in Life Science*, 21-54. Indian Statistical Institute, Calcutta, India.

Jammalamadaka, S. R. and Sarma, Y. R. (1988). A correlation coefficient for angular variables. *Statistical Theory and Data Analysis II*, 349-364, North Holand, Amesterdam.

Jammalamadaka, S. R. and Sarma, Y. R. (1993). Circular regression. In Matusita, K., editor, *Statistical Theory and Data Analysis*, 109-128. VSP, Utrecht.

Jammalamadaka, S. R. and SenGupta, A. (2001). *Topics in Circular Statistics*. World Scientific Press, Singapore.

Johnson, R. A. and Wehrly, T. E. (1977). Measures and models for angular correlation and angular- linear correlation. *Journal of the Royal Statistical Society, Series B*, **39**, 222-229.

Johnson, R. A. and Wehrly, T. E. (1978). Some angular-linear distributions and related regression models. *Journal of the American Statistical Association*, **73** (363), 602-606.

Jones, T. A. and James, W. R. (1969). Analysis of bimodal orientation data. *Mathematical Geology*, **1** (2), 129-135.

Jupp, P. E. and Mardia, K. V. (1989). A unified view of the theory of directional statistics, 1975-1988. *Internat. Statist. Rev.*, **57**, 261-294.

Kato, S., Shimizu, K. and Shieh, G. S. (2008). A circular-circular regression model. *Statistica Sinica*, **18** (2), 633-643.

Kendall, M. G. (1951). Regression, structural and functional relationship I, *Biometrika*, **38**, pp11-25.

Kendall, M. G. (1952). Regression, structural and functional relationship I, *Biometrika*, **38**, pp96-108.

Kendall, M.G. and Stuart, A. (1973) *The Advance Theory of Statistics*, Vol.2 London: Griffin

Kufner, A. and Kadlec, J. (1971). *Fourier Series*. Academia,Prague.

Lawrence, K. D. and Arthur, J. L. (1990). *Robust Regression: Analysis and Applications*. Marcel Dekker Inc., New York.

Laycock, P. J. (1975). Optimal design: Regression model for directions. *Biometrika*, **62** (2), 305-311.

Lemieux, P. H. (1978). A Note on the Detection of Multicillinearity. *American Journal of Political Science*, **22 (**1), pp. 183-186

Lenth, R. V. (1981). On finding the source of a signal. *Technometrics*, **23** (2), 149-154.

Liu, K. (1993). A New Class of Biased Estimate in Linear Regression. *Comm Statistics. Theory Methods*, **22** (2), 393-402.

Lund, U. (1999). Least circular distance regression for directional data. *Journal of Applied Statistics*, **26** (6), 723-733.

Madansky, A. (1959). The fitting of straight lines when both variables are subject to error. *J. Am. Statist. Ass.*, **54**: 173-205.

Mansfield, E. R. and Helms, B. P. (1982). Detecting Multicollinearity. *The American Statistician*, **36 (**3), Part 1, pp. 158-160

Mardia, K. V. (1972). *Statistics of Directional Data*. Academic Press, London.

Mardia, K. V. (1975). Statistics of directional data. *Journal of the Royal Statistical Society, Series* B, **37**, 349-393.

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

Montgomery, D. C. and Peck, E. A. (1992). *Introduction to Linear Regression Analysis*, 2nd ed., John Wiley & Sons, New York.

Moore-Ede, M. C., Sulzman, F. M. & Fuller, C. A. (1982). *The Clocks that Time Us: Physiology of the Circadian Timing System*. Cambridge, MA: Harvard University Press

Moran, P. A. P. (1971). Estimating structural and functional relationships. *J. Multiv. Anal.*, 1: 232-255.

Nolan, W. P., FRCOphth, See, J. L., et al. (2007). Detection of Primary Angle Closure Using Anterior Segment Optical Coherence Tomography in Asian Eye, *American Academy of Ophthalmology*, **114** (1), pp 33-39.

Rambli, A. (2011). Outlier Detection in Circular Data and Circular-circular Regression Model, (MSc thesis). Institute of Mathematical Science, University of Malaya.

Ramsay, J. O. and Silverman, B. W. (2005). *Functional Data Analysis*. 2nd ed. Springer Series in Statistics, Springer, New York.

Rao, J. S. (1969). *Some Contributions to the Analysis of Circular Data.* Ph.D. thesis, Indian Statistical Institute, Calcutta, India.

Rivest, L.-P. (1997). A decentred predictor for circular –circular regression. *Biometrika*, **84** (3), 717-726.

Swindel, F. F. (1976). Good ridge estimators based on prior information. Comm Statistics. Theory Methods, **A5** (11): 1065-1075.

Tukey, J. W. (1977). *Exploratory Data Analysis*. Addison-Wesley, Reading, MA.

Von Mises, R. (1918). Uber die “Ganzzahligkeit” der atomgewichte und vewandte fragen. *Physikal. Z.* , 19,490-500.

Wehrly, T. E. and Shine, E. P. (1981). Influence curves of estimators for directional data. *Biometrika*, **68** (1), 334-335.

Whittaker, E. T. and Watson, G. N. (1994). *A course in Modern Analysis,* Cambridge University Press, Cambridge.

APPENDIX 1

Wind direction data

Obs. No.	Radar Time	Radar Obs.	Anchored Buoy Time	Anchored Buoy Obs.	Obs. No.	Radar Time	Radar Obs.	Anchored Buoy Time	Anchored Buoy Obs.
1	1.615	0.79	1.618	1.154	33	3.823	5.406	3.826	5.744
2	1.656	0.715	1.66	1.154	34	3.865	5.472	3.868	5.547
3	1.698	0.975	1.701	1.007	35	3.906	5.401	3.91	5.498
4	1.74	0.97	1.743	1.178	36	3.948	5.42	3.951	5.4
5	1.781	0.993	1.785	0.859	37	3.99	5.276	3.993	5.449
6	1.823	0.902	1.826	1.007	38	4.031	1.728	4.035	4.786
7	1.837	0.943	1.847	1.056	39	4.406	5.512	4.41	5.449
8	2.406	1.728	2.41	1.4	40	4.448	5.486	4.451	5.178
9	2.448	1.445	2.451	1.497	41	4.49	5.444	4.493	5.62
10	2.49	1.679	2.493	1.693	42	4.531	5.518	4.535	5.13
11	2.531	1.703	2.535	2.012	43	4.559	5.505	4.576	4.541
12	2.573	1.862	2.576	1.792	44	9.573	5.558	9.576	5.571
13	2.615	1.726	2.618	1.766	45	9.615	5.42	9.618	5.62
14	2.656	1.79	2.66	1.669	46	9.656	5.398	9.66	5.473
15	2.698	1.831	2.701	1.4	47	9.698	5.334	9.701	5.327
16	2.726	1.719	2.743	1.4	48	9.781	5.418	9.785	4.835
17	2.781	1.646	2.785	1.375	49	9.823	5.418	9.826	5.032
18	2.823	1.622	2.826	1.056	50	9.892	5.338	9.91	5.842
19	2.865	1.342	2.868	1.178	51	9.948	5.47	9.951	5.571
20	2.906	1.176	2.91	1.276	52	9.99	5.455	9.993	5.522
21	2.948	1.325	2.951	1.693	53	10.073	5.555	10.076	5.473
22	2.99	1.103	2.993	1.325	54	10.115	5.462	10.118	5.522
23	3.406	6.131	3.41	6.062	55	10.156	5.401	10.16	5.522
24	3.448	5.719	3.451	5.988	56	10.198	5.316	10.201	5.376
25	3.49	5.713	3.493	5.988	57	10.24	5.439	10.243	5.081
26	3.531	5.487	3.535	5.498	58	10.406	5.408	10.41	5.473
27	3.573	5.742	3.576	5.276	59	10.448	5.431	10.451	5.449
28	3.615	5.728	3.618	5.302	60	10.49	5.473	10.493	5.915
29	3.656	5.61	3.66	5.62	61	10.531	5.46	10.535	5.351
30	3.698	5.463	3.701	5.744	62	10.573	5.364	10.576	5.571
31	3.74	5.427	3.743	5.644	63	10.615	5.444	10.618	5.376
32	3.781	5.418	3.785	5.669	64	10.656	5.35	10.66	5.327
65	10.698	5.202	10.701	4.983	70	10.906	5.238	10.91	4.417
66	10.74	5.161	10.743	4.786	71	10.948	4.97	10.951	5.007
67	10.781	5.062	10.785	4.908	72	10.99	4.947	10.993	5.473
68	10.823	5.145	10.826	4.517	73	11.031	4.887	11.035	5.4
69	10.865	5.212	10.868	4.835	74	11.073	4.872	11.076	4.859

Appendix1, continued.

Obs. No.	Radar		Anchored Buoy		Obs. No.	Radar		Anchored Buoy	
	Time	Obs.	Time	Obs.		Time	Obs.	Time	Obs.
75	11.115	4.589	11.118	4.859	103	20.906	0.237	20.91	0.171
76	11.156	4.51	11.16	4.761	104	20.948	0.045	20.951	0.295
77	11.281	4.319	11.285	4.639	105	20.99	6.241	20.993	6.259
78	11.323	4.427	11.326	4.664	106	21.031	0.248	21.035	0.319
79	11.337	4.436	11.347	4.664	107	21.073	0.578	21.076	0.539
80	11.406	4.451	11.41	4.074	108	21.087	0.627	21.097	0.81
81	12.198	3.84	12.201	4.295	109	21.406	0.251	21.41	6.161
82	12.24	3.819	12.243	4.098	110	21.448	5.299	21.451	5.473
83	12.281	4.159	12.285	4.173	111	21.49	3.749	21.493	5.62
84	12.323	3.987	12.326	4.122	112	21.531	1.876	21.535	2.012
85	19.823	5.506	19.826	5.817	113	21.573	1.776	21.576	1.963
86	19.865	5.509	19.868	5.571	114	21.615	1.786	21.618	1.841
87	19.906	5.643	19.91	5.571	115	21.656	1.658	21.66	1.89
88	19.948	5.707	19.951	5.596	116	21.684	1.377	21.701	1.497
89	19.99	5.727	19.993	5.964	117	21.74	1.305	21.743	1.669
90	20.031	5.685	20.035	5.547	118	21.781	1.309	21.785	1.325
91	20.073	5.696	20.076	6.161	119	21.823	1.337	21.826	1.644
92	20.115	5.745	20.118	6.037	120	21.865	1.198	21.868	1.571
93	20.142	5.837	20.16	5.915	121	21.906	1.15	21.91	1.08
94	20.531	1.146	20.535	1.546	122	21.948	1.047	21.951	1.129
95	20.573	1.074	20.576	1.866	123	21.99	0.97	21.993	0.466
96	20.615	1.201	20.618	1.717	124	22.031	0.998	22.035	0.981
97	20.656	1.253	20.66	1.89	125	22.073	1.071	22.076	1.007
98	20.698	1.032	20.701	1.89	126	22.531	0.793	22.535	0.834
99	20.74	1.093	20.743	1.988	127	22.573	0.753	22.576	1.056
100	20.781	0.505	20.785	6.137	128	22.615	0.573	22.618	0.932
101	20.823	0.234	20.826	0.393	129	22.656	0.437	22.66	0.761
102	20.865	0.275	20.868	0.271					

APPENDIX 2

The 1% upper percentiles of the $\left| COVRATIO_{(-j)} - 1 \right|$ statistic at a = -2

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
10	**0.03**	0.9758	0.9709	0.9764	0.9773	0.9768	0.9709
	0.05	0.9819	0.9768	0.9772	0.9743	0.9790	0.9896
	0.08	0.9706	0.9747	0.9822	0.9756	0.9768	0.9885
	0.1	0.9720	0.9692	0.9812	0.9768	0.9688	0.9881
	0.3	0.9740	0.9755	0.9880	0.9867	0.9749	0.9869
	0.6	0.9799	0.9733	0.9810	0.9807	0.9771	0.9718
20	**0.03**	0.7739	0.7882	0.9209	0.8308	0.8827	0.8921
	0.05	0.7569	0.7756	0.7842	0.7908	0.8857	0.8903
	0.08	0.7781	0.7473	0.7597	0.7449	0.8786	0.8888
	0.1	0.7805	0.7769	0.7650	0.7678	0.8562	0.8827
	0.3	0.8469	0.9315	0.8168	0.8353	0.8168	0.8072
	0.6	0.8136	0.8316	0.8220	0.8300	0.8214	0.7709
30	**0.03**	0.7282	0.7935	0.8089	0.8048	0.8778	0.9049
	0.05	0.7306	0.7454	0.7653	0.7781	0.8693	0.9299
	0.08	0.7363	0.7350	0.7393	0.7782	0.8770	0.8932
	0.1	0.7334	0.7346	0.7367	0.7623	0.8771	0.8987
	0.3	0.7638	0.7796	0.8282	0.8377	0.8752	0.8930
	0.6	0.8260	0.8435	0.8675	0.8784	0.9228	0.9533
40	**0.03**	0.6106	0.6474	0.7062	0.7220	0.8572	0.8571
	0.05	0.5885	0.6253	0.6786	0.6894	0.8579	0.8757
	0.08	0.6169	0.5963	0.6588	0.7140	0.8339	0.8849
	0.1	0.6258	0.6306	0.6543	0.6968	0.8204	0.8914
	0.3	0.7178	0.7219	0.7665	0.7833	0.8267	0.8477
	0.6	0.7586	0.7563	0.7680	0.7924	0.8484	0.8238
50	**0.03**	0.5176	0.5706	0.6358	0.6618	0.7809	0.7432
	0.05	0.5284	0.5472	0.6061	0.6405	0.7725	0.7430
	0.08	0.5411	0.5466	0.5807	0.6077	0.7754	0.7503
	0.1	0.5644	0.5632	0.5674	0.5977	0.7634	0.7545
	0.3	0.6988	0.3686	0.6785	0.6685	0.7486	0.7633
	0.6	0.6695	0.6876	0.7142	0.7181	0.7346	0.7138
60	**0.03**	0.4327	0.4582	0.5191	0.5281	0.7834	0.6839
	0.05	0.4416	0.4295	0.4772	0.4995	0.7623	0.6954
	0.08	0.4659	0.4555	0.4666	0.5010	0.7716	0.6932
	0.1	0.4775	0.4764	0.4584	0.4730	0.7434	0.6752
	0.3	0.5801	0.5711	0.5625	0.5436	0.6416	0.6754
	0.6	0.6585	0.6566	0.6754	0.6585	0.6566	0.6612
70	**0.03**	0.4097	0.4700	0.5237	0.5508	0.7158	0.6050
	0.05	0.4211	0.4334	0.4965	0.5238	0.7174	0.6044
	0.08	0.4580	0.4190	0.4495	0.4942	0.6986	0.6038
	0.1	0.4717	0.4649	0.4705	0.4715	0.6977	0.6122
	0.3	0.5681	0.5833	0.5767	0.5685	0.6307	0.5799
	0.6	0.5822	0.5647	0.5832	0.5863	0.6385	0.5288

The 1% upper percentiles of the $\left|COVRATIO_{(-j)} - 1\right|$ statistic at a = -2, continued.

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
80	**0.03**	0.3726	0.4242	0.4781	0.5161	0.6903	0.6024
	0.05	0.3686	0.3744	0.4373	0.4659	0.6822	0.5894
	0.08	0.3995	0.3954	0.4067	0.4047	0.6573	0.5882
	0.1	0.4030	0.4026	0.4185	0.4186	0.6598	0.5798
	0.3	0.4824	0.5134	0.5313	0.5349	0.5985	0.5324
	0.6	0.5537	0.5627	0.5839	0.6100	0.5839	0.4753
90	**0.03**	0.3274	0.3666	0.4316	0.4594	0.7177	0.4919
	0.05	0.3291	0.3234	0.3866	0.4183	0.7226	0.5065
	0.08	0.3583	0.3449	0.3723	0.3800	0.6988	0.5192
	0.1	0.3570	0.3517	0.3615	0.3752	0.6371	0.5218
	0.3	0.5377	0.5502	0.5336	0.5156	0.5866	0.5167
	0.6	0.5059	0.5000	0.5312	0.5253	0.5907	0.4354
100	**0.03**	0.3170	0.3673	0.4230	0.4371	0.5833	0.4318
	0.05	0.3279	0.3256	0.3739	0.4206	0.5653	0.4256
	0.08	0.3500	0.3325	0.3352	0.3558	0.5684	0.4243
	0.1	0.3649	0.3445	0.3387	0.3492	0.5469	0.4416
	0.3	0.5006	0.5021	0.4974	0.4979	0.5557	0.4562
	0.6	0.4141	0.4351	0.4467	0.4562	0.5193	0.3907
110	**0.03**	0.2808	0.3093	0.3581	0.3805	0.6443	0.4377
	0.05	0.2719	0.2995	0.3334	0.3614	0.6439	0.4519
	0.08	0.3278	0.2946	0.3015	0.3105	0.6310	0.4670
	0.1	0.3432	0.3298	0.3100	0.3083	0.6165	0.4572
	0.3	0.4492	0.4397	0.4331	0.4297	0.5266	0.3884
	0.6	0.4547	0.4494	0.4618	0.4669	0.4528	0.3730
130	**0.03**	0.2438	0.2844	0.3271	0.3515	0.4783	0.3835
	0.05	0.2625	0.2458	0.2932	0.3237	0.4670	0.4020
	0.08	0.3036	0.2537	0.2681	0.4578	0.4578	0.3916
	0.1	0.3030	0.2603	0.2559	0.2829	0.4433	0.3998
	0.3	0.4269	0.4497	0.4535	0.4480	0.5194	0.3429
	0.6	0.3938	0.3931	0.3982	0.3933	0.4105	0.2790
150	**0.03**	0.2396	0.2810	0.2898	0.3008	0.5290	0.3120
	0.05	0.1639	0.2229	0.2667	0.2866	0.5201	0.3139
	0.08	0.2333	0.2068	0.2480	0.2659	0.5162	0.3216
	0.1	0.2356	0.2164	0.2274	0.2560	0.5037	0.3303
	0.3	0.3804	0.3659	0.3657	0.3737	0.4544	0.2975
	0.6	0.3293	0.3335	0.3223	0.3188	0.3913	0.2376

The 5% upper percentiles of the $\left| COVRATIO_{(-j)} - 1 \right|$ statistic at a = -2

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
10	**0.03**	0.9364	0.9371	0.9485	0.9490	0.9501	0.9371
	0.05	0.9391	0.9501	0.9422	0.9484	0.9500	0.9714
	0.08	0.9370	0.9478	0.9513	0.9487	0.9518	0.9646
	0.1	0.9421	0.9389	0.9428	0.9452	0.9453	0.9610
	0.3	0.9430	0.9450	0.9487	0.9471	0.9418	0.9629
	0.6	0.9556	0.9557	0.9484	0.9375	0.9271	0.9351
20	**0.03**	0.6761	0.7066	0.8290	0.7288	0.8147	0.7917
	0.05	0.6698	0.6666	0.6976	0.7201	0.8185	0.7893
	0.08	0.6782	0.6657	0.6555	0.6814	0.8072	0.7766
	0.1	0.6986	0.6753	0.6686	0.6705	0.7890	0.7614
	0.3	0.7435	0.8478	0.7341	0.7252	0.7341	0.7063
	0.6	0.7134	0.7146	0.7190	0.7263	0.7030	0.6670
30	**0.03**	0.5802	0.6321	0.6900	0.7017	0.8088	0.8089
	0.05	0.6287	0.6124	0.6444	0.6753	0.7996	0.8310
	0.08	0.6437	0.6326	0.6416	0.6502	0.7964	0.8170
	0.1	0.6502	0.6505	0.6593	0.6572	0.8064	0.8140
	0.3	0.6899	0.7042	0.7056	0.7151	0.7953	0.8335
	0.6	0.7590	0.7706	0.7851	0.7903	0.8188	0.8530
40	**0.03**	0.4808	0.5454	0.5930	0.6135	0.7229	0.7615
	0.05	0.5047	0.5168	0.5699	0.5981	0.7151	0.7617
	0.08	0.5222	0.5315	0.5476	0.5704	0.7198	0.7749
	0.1	0.5498	0.5489	0.5679	0.5723	0.7012	0.7650
	0.3	0.5832	0.5910	0.5983	0.6091	0.7078	0.7246
	0.6	0.6137	0.6296	0.6606	0.6663	0.7542	0.7197
50	**0.03**	0.4392	0.4722	0.5265	0.5511	0.6536	0.6142
	0.05	0.4291	0.451	0.4941	0.4959	0.6596	0.6285
	0.08	0.4557	0.4502	0.4765	0.4968	0.6523	0.6542
	0.1	0.4662	0.4571	0.4747	0.4911	0.6513	0.6481
	0.3	0.5184	0.3019	0.5486	0.5426	0.5965	0.6165
	0.6	0.5757	0.5813	0.5915	0.6024	0.6235	0.6082
60	**0.03**	0.3605	0.4024	0.4400	0.4722	0.6261	0.5574
	0.05	0.3699	0.3686	0.4099	0.4370	0.6273	0.5556
	0.08	0.3941	0.3795	0.3819	0.4088	0.6077	0.5667
	0.1	0.4044	0.3918	0.3828	0.4006	0.6053	0.5622
	0.3	0.4465	0.4598	0.4587	0.4596	0.5534	0.5433
	0.6	0.5290	0.5496	0.5433	0.5290	0.5496	0.5333
70	**0.03**	0.3484	0.3763	0.4290	0.4468	0.5875	0.5070
	0.05	0.3617	0.3615	0.3832	0.4095	0.6014	0.5114
	0.08	0.3710	0.3672	0.3708	0.3848	0.5962	0.5152
	0.1	0.3722	0.3777	0.3698	0.3801	0.5822	0.5228
	0.3	0.4521	0.4593	0.4617	0.4520	0.5004	0.4756
	0.6	0.4775	0.4744	0.4880	0.4838	0.5134	0.4477

The 5% upper percentiles of the $\left| COVRATIO_{(-j)} - 1 \right|$ statistic at a = -2, continued.

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
80	**0.03**	0.3003	0.3384	0.3678	0.3930	0.5494	0.4926
	0.05	0.3019	0.3057	0.3356	0.3567	0.5458	0.4885
	0.08	0.3136	0.3130	0.3211	0.3272	0.5407	0.4881
	0.1	0.3255	0.3187	0.3219	0.3235	0.5275	0.4916
	0.3	0.4118	0.4062	0.4053	0.4134	0.4663	0.4365
	0.6	0.4507	0.4575	0.4603	0.4836	0.4949	0.3994
90	**0.03**	0.2742	0.3169	0.3547	0.3764	0.5210	0.4302
	0.05	0.2777	0.2827	0.3257	0.3526	0.5108	0.4389
	0.08	0.3020	0.3020	0.2924	0.3105	0.5096	0.4375
	0.1	0.3175	0.3074	0.2999	0.3063	0.4968	0.4369
	0.3	0.4042	0.4006	0.4221	0.4078	0.4465	0.4196
	0.6	0.4102	0.4290	0.4399	0.4315	0.4471	0.3590
100	**0.03**	0.2757	0.3028	0.3367	0.3618	0.5062	0.3721
	0.05	0.2676	0.2755	0.3038	0.3329	0.4953	0.3854
	0.08	0.2800	0.2764	0.2779	0.2971	0.4774	0.3834
	0.1	0.2869	0.2841	0.2853	0.2854	0.4636	0.3762
	0.3	0.3717	0.3760	0.3812	0.3718	0.4327	0.3798
	0.6	0.3768	0.3728	0.3848	0.3846	0.4265	0.3339
110	**0.03**	0.2157	0.2582	0.3068	0.3235	0.4719	0.3808
	0.05	0.2257	0.2318	0.2674	0.2928	0.4737	0.3851
	0.08	0.2488	0.2489	0.2446	0.2627	0.4718	0.3796
	0.1	0.2540	0.2491	0.2461	0.2476	0.4575	0.3804
	0.3	0.3472	0.3467	0.3455	0.3504	0.3882	0.3399
	0.6	0.3740	0.3743	0.3878	0.3872	0.3880	0.2876
130	**0.03**	0.1948	0.2276	0.2719	0.2842	0.4115	0.3203
	0.05	0.2028	0.1989	0.2394	0.2676	0.4102	0.3237
	0.08	0.2281	0.2139	0.2151	0.3950	0.3950	0.3227
	0.1	0.2382	0.2234	0.2235	0.2243	0.3845	0.3211
	0.3	0.3270	0.3227	0.3099	0.2911	0.3639	0.2924
	0.6	0.3187	0.3211	0.3302	0.3428	0.3694	0.2457
150	**0.03**	0.1918	0.2295	0.2519	0.2638	0.3738	0.2651
	0.05	0.1788	0.1837	0.2138	0.2383	0.3719	0.2650
	0.08	0.1956	0.1874	0.1918	0.2080	0.3620	0.2654
	0.1	0.1994	0.1884	0.1927	0.1937	0.3582	0.2706
	0.3	0.3000	0.2967	0.3036	0.2938	0.3334	0.2563
	0.6	0.2771	0.2736	0.2779	0.2773	0.3047	0.2019

The 10% upper percentiles of the $\left|COVRATIO_{(-j)} - 1\right|$ statistic at a = -2

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
10	**0.03**	0.9045	0.9050	0.9196	0.9193	0.9216	0.9050
	0.05	0.9015	0.9216	0.9049	0.9153	0.9177	0.9452
	0.08	0.9049	0.9047	0.9093	0.9139	0.9231	0.9458
	0.1	0.9007	0.9044	0.9163	0.9137	0.9217	0.9432
	0.3	0.9164	0.9125	0.9178	0.9139	0.8938	0.9214
	0.6	0.9235	0.9149	0.9087	0.9105	0.8922	0.8897
20	**0.03**	0.6151	0.6541	0.7696	0.6821	0.7499	0.7310
	0.05	0.6112	0.6174	0.6314	0.6523	0.7518	0.7310
	0.08	0.6230	0.6079	0.6061	0.6184	0.7426	0.7112
	0.1	0.6328	0.6155	0.6129	0.6183	0.7380	0.7064
	0.3	0.6651	0.7800	0.6708	0.6481	0.6708	0.6540
	0.6	0.6607	0.6497	0.6578	0.6573	0.6180	0.6030
30	**0.03**	0.5411	0.5812	0.6275	0.6506	0.7464	0.7549
	0.05	0.5701	0.5574	0.6041	0.6284	0.7466	0.7911
	0.08	0.5726	0.5825	0.5729	0.6034	0.7241	0.7490
	0.1	0.5938	0.6018	0.5767	0.5874	0.7323	0.7485
	0.3	0.6341	0.6579	0.6599	0.6592	0.7347	0.7555
	0.6	0.6879	0.6931	0.7093	0.7180	0.7678	0.7832
40	**0.03**	0.4306	0.4845	0.5368	0.5746	0.6615	0.6884
	0.05	0.4350	0.4585	0.5034	0.5420	0.6578	0.6966
	0.08	0.4583	0.4574	0.4814	0.5097	0.6634	0.6976
	0.1	0.4746	0.4682	0.4882	0.4959	0.6428	0.6912
	0.3	0.5154	0.5356	0.5381	0.5502	0.6221	0.6603
	0.6	0.5558	0.5629	0.6004	0.6045	0.6899	0.6327
50	**0.03**	0.3790	0.4079	0.4648	0.4802	0.5896	0.5667
	0.05	0.3968	0.3904	0.4252	0.4502	0.5962	0.5694
	0.08	0.4039	0.398	0.4164	0.4295	0.5845	0.5790
	0.1	0.4165	0.4033	0.4244	0.4309	0.5764	0.5770
	0.3	0.4691	0.2645	0.4770	0.4715	0.5273	0.5503
	0.6	0.5195	0.5353	0.5471	0.5418	0.5810	0.5365
60	**0.03**	0.3217	0.3576	0.4000	0.4170	0.5470	0.5017
	0.05	0.3247	0.3352	0.3806	0.4041	0.5480	0.5117
	0.08	0.3411	0.3343	0.3493	0.3729	0.5386	0.5099
	0.1	0.3610	0.3471	0.3502	0.3599	0.5376	0.5030
	0.3	0.4125	0.4196	0.4220	0.4293	0.4744	0.4855
	0.6	0.6585	0.4965	0.4855	0.4799	0.4965	0.4684
70	**0.03**	0.3103	0.3383	0.3624	0.3877	0.5261	0.4629
	0.05	0.3048	0.3153	0.3456	0.3593	0.5207	0.4620
	0.08	0.3293	0.3169	0.3282	0.3354	0.5222	0.4558
	0.1	0.3355	0.3221	0.3212	0.3295	0.5112	0.4551
	0.3	0.3978	0.4017	0.3964	0.3857	0.4442	0.4294
	0.6	0.4181	0.4223	0.4311	0.4379	0.4583	0.3909

The 10% upper percentiles of the $\left| COVRATIO_{(-j)} - 1 \right|$ statistic at a = -2, continued.

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
80	**0.03**	0.2606	0.2957	0.3277	0.3360	0.4755	0.4386
	0.05	0.2645	0.2639	0.3039	0.3159	0.4658	0.4417
	0.08	0.2937	0.2697	0.2845	0.2948	0.4420	0.4391
	0.1	0.2976	0.2841	0.2799	0.2905	0.4313	0.4393
	0.3	0.3558	0.3503	0.3497	0.3555	0.4131	0.4006
	0.6	0.3817	0.3882	0.4018	0.4069	0.4246	0.3582
90	**0.03**	0.2508	0.2823	0.3237	0.3405	0.4654	0.3867
	0.05	0.2521	0.2599	0.2951	0.3119	0.4570	0.3958
	0.08	0.2712	0.2621	0.2715	0.2862	0.4568	0.3990
	0.1	0.2793	0.2783	0.2732	0.2793	0.4350	0.3948
	0.3	0.3394	0.3458	0.3421	0.3485	0.3729	0.3569
	0.6	0.3588	0.3690	0.3710	0.3719	0.4157	0.3331
100	**0.03**	0.2360	0.2646	0.2982	0.3184	0.4289	0.3411
	0.05	0.2284	0.2440	0.2693	0.2842	0.4182	0.3398
	0.08	0.2506	0.2435	0.2496	0.2593	0.3957	0.3467
	0.1	0.2620	0.2481	0.2541	0.2616	0.3829	0.3471
	0.3	0.3273	0.3322	0.3355	0.3234	0.3458	0.3409
	0.6	0.3310	0.3383	0.3474	0.3448	0.3659	0.2952
110	**0.03**	0.1997	0.2253	0.2592	0.2770	0.3923	0.3277
	0.05	0.2083	0.2060	0.2349	0.2572	0.3916	0.3334
	0.08	0.2217	0.2128	0.2141	0.2310	0.3860	0.3327
	0.1	0.2287	0.2191	0.2185	0.2243	0.3860	0.3361
	0.3	0.3052	0.3122	0.3155	0.3106	0.3352	0.3044
	0.6	0.3286	0.3331	0.3385	0.3347	0.3493	0.2596
130	**0.03**	0.1805	0.2046	0.2381	0.2508	0.3618	0.2868
	0.05	0.1773	0.1875	0.2125	0.2303	0.3617	0.2857
	0.08	0.1975	0.1851	0.1928	0.3552	0.3552	0.2811
	0.1	0.2120	0.2006	0.2012	0.1966	0.3446	0.2820
	0.3	0.2743	0.2712	0.2680	0.2656	0.3005	0.2630
	0.6	0.2764	0.2788	0.2803	0.2848	0.3053	0.2144
150	**0.03**	0.1641	0.1925	0.2190	0.2256	0.3236	0.2424
	0.05	0.2131	0.1637	0.1888	0.2084	0.3239	0.2438
	0.08	0.1722	0.1683	0.1737	0.1889	0.3156	0.2467
	0.1	0.1772	0.1722	0.1738	0.1802	0.3091	0.2476
	0.3	0.2528	0.2565	0.2588	0.2551	0.2797	0.2317
	0.6	0.2489	0.2505	0.2522	0.2596	0.2688	0.1866

The 1% upper percentiles of the $\left| COVRATIO_{(-j)} - 1 \right|$ statistic at a = -1

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
10	**0.03**	0.9729	0.9796	0.9851	0.9753	0.9898	0.9872
	0.05	0.9642	0.9766	0.9872	0.9723	0.9904	0.9875
	0.08	0.9615	0.9658	0.9791	0.9732	0.9840	0.9846
	0.1	0.9746	0.9686	0.9692	0.9685	0.9837	0.9843
	0.3	0.9768	0.9744	0.9697	0.9758	0.9869	0.9810
	0.6	0.9851	0.9842	0.9828	0.9811	0.9836	0.9681
20	**0.03**	0.7672	0.7992	0.8346	0.8509	0.8883	0.9024
	0.05	0.8075	0.8216	0.8003	0.8258	0.9592	0.9004
	0.08	0.8639	0.8059	0.8659	0.8036	0.9544	0.8848
	0.1	0.8706	0.8120	0.8754	0.7847	0.9604	0.9069
	0.3	0.8375	0.8284	0.8359	0.8158	0.8535	0.8216
	0.6	0.8573	0.8319	0.8375	0.8210	0.8272	0.7421
30	**0.03**	0.7196	0.7377	0.7377	0.7865	0.9021	0.8935
	0.05	0.7152	0.7339	0.7494	0.7838	0.8994	0.8897
	0.08	0.7323	0.7325	0.7512	0.7607	0.8992	0.8853
	0.1	0.7328	0.7487	0.7560	0.7552	0.8664	0.8976
	0.3	0.8085	0.7960	0.8248	0.8360	0.8953	0.9173
	0.6	0.8541	0.8795	0.8894	0.8847	0.9494	0.9636
40	**0.03**	0.5767	0.6548	0.7212	0.7376	0.8802	0.8315
	0.05	0.6197	0.6278	0.6823	0.7020	0.8572	0.8196
	0.08	0.6123	0.6209	0.6395	0.6693	0.8459	0.8622
	0.1	0.6200	0.6298	0.6471	0.6411	0.8393	0.8551
	0.3	0.7449	0.7271	0.7359	0.7449	0.8271	0.8256
	0.6	0.7296	0.7547	0.7808	0.7689	0.8319	0.8303
50	**0.03**	0.5638	0.6176	0.6432	0.6556	0.7816	0.7835
	0.05	0.5717	0.5955	0.6121	0.6562	0.7698	0.7848
	0.08	0.6094	0.5839	0.6123	0.6262	0.7565	0.7765
	0.1	0.6107	0.6098	0.5819	0.6107	0.7449	0.7603
	0.3	0.7212	0.7256	0.7126	0.7061	0.7816	0.7472
	0.6	0.6866	0.7008	0.7238	0.7174	0.7397	0.7308
60	**0.03**	0.4603	0.5017	0.5640	0.5551	0.7205	0.6567
	0.05	0.4888	0.4691	0.4962	0.5088	0.7305	0.6607
	0.08	0.5256	0.4938	0.4847	0.4876	0.7316	0.6764
	0.1	0.5594	0.5241	0.4850	0.4793	0.7196	0.6569
	0.3	0.6812	0.6569	0.6869	0.6549	0.6902	0.6946
	0.6	0.5751	0.5763	0.6081	0.6184	0.6960	0.6344
70	**0.03**	0.4274	0.5037	0.5110	0.5293	0.6566	0.6588
	0.05	0.4489	0.4366	0.4870	0.5245	0.6975	0.6333
	0.08	0.4959	0.4758	0.4639	0.4658	0.6685	0.6426
	0.1	0.4929	0.4802	0.4739	0.4650	0.6432	0.6317
	0.3	0.6025	0.6120	0.6029	0.5929	0.6383	0.5736
	0.6	0.5736	0.5661	0.5697	0.5850	0.6722	0.5731

The 1% upper percentiles of the $\left|COVRATIO_{(-j)} - 1\right|$ statistic at a = -1, continued.

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
80	**0.03**	0.4274	0.4406	0.5141	0.5200	0.6338	0.5667
	0.05	0.3847	0.4012	0.4552	0.4905	0.6387	0.5763
	0.08	0.4242	0.3873	0.4178	0.4484	0.6244	0.5693
	0.1	0.4500	0.4128	0.4095	0.4426	0.6020	0.5684
	0.3	0.5045	0.5156	0.5484	0.5667	0.5655	0.5463
	0.6	0.5574	0.3886	0.5592	0.5478	0.5739	0.4934
90	**0.03**	0.3667	0.4008	0.4524	0.4728	0.6042	0.5528
	0.05	0.3745	0.3587	0.4226	0.4470	0.6031	0.5356
	0.08	0.4219	0.3992	0.3924	0.3883	0.6010	0.5408
	0.1	0.4309	0.4162	0.3996	0.4302	0.5901	0.5449
	0.3	0.5687	0.5677	0.5608	0.5601	0.5248	0.5248
	0.6	0.5448	0.5569	0.5351	0.5299	0.5385	0.4809
100	**0.03**	0.3502	0.3846	0.4407	0.4503	0.5702	0.5167
	0.05	0.3677	0.3682	0.3777	0.4251	0.5502	0.5235
	0.08	0.3989	0.3932	0.4044	0.4149	0.5501	0.5170
	0.1	0.4159	0.4066	0.3970	0.4037	0.5446	0.5335
	0.3	0.5500	0.5247	0.5351	0.5063	0.5681	0.4883
	0.6	0.4373	0.4355	0.4421	0.4420	0.4665	0.3965
110	**0.03**	0.3039	0.3685	0.4137	0.4352	0.5811	0.4389
	0.05	0.2974	0.3184	0.3758	0.3975	0.5695	0.4342
	0.08	0.3095	0.3002	0.3272	0.3666	0.5589	0.4423
	0.1	0.3251	0.3080	0.3194	0.353	0.5682	0.4364
	0.3	0.4642	0.4635	0.4656	0.4604	0.4981	0.3784
	0.6	0.3946	0.4088	0.4134	0.4256	0.4785	0.3203
130	**0.03**	0.2563	0.2866	0.3200	0.3521	0.5502	0.4072
	0.05	0.2919	0.2645	0.2962	0.3094	0.2919	0.3980
	0.08	0.3152	0.3044	0.2717	0.2900	0.4951	0.3915
	0.1	0.3255	0.3231	0.2891	0.2830	0.4768	0.3866
	0.3	0.4540	0.4516	0.4362	0.4389	0.4433	0.3645
	0.6	0.4209	0.4098	0.3838	0.3992	0.4184	0.3066
150	**0.03**	0.2280	0.2514	0.2908	0.2514	0.5125	0.3436
	0.05	0.2405	0.2377	0.2738	0.2842	0.5191	0.3441
	0.08	0.2498	0.2429	0.2673	0.2523	0.4892	0.3505
	0.1	0.2671	0.2517	0.2523	0.2739	0.4790	0.3480
	0.3	0.4570	0.4656	0.4670	0.4574	0.4208	0.2424
	0.6	0.3502	0.3400	0.3487	0.3404	0.3657	0.2756

The 5% upper percentiles of the $\left|COVRATIO_{(-j)} - 1\right|$ statistic at a = -1

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
10	**0.03**	0.9274	0.9380	0.9519	0.9531	0.9658	0.9618
	0.05	0.9319	0.9300	0.9618	0.9376	0.9555	0.9619
	0.08	0.9332	0.9338	0.9396	0.9437	0.9478	0.9633
	0.1	0.9323	0.9390	0.9447	0.9457	0.9440	0.9650
	0.3	0.9377	0.9321	0.9305	0.9360	0.9497	0.9522
	0.6	0.9581	0.9516	0.9480	0.9434	0.9490	0.9476
20	**0.03**	0.6977	0.6989	0.7372	0.7577	0.8064	0.8052
	0.05	0.6904	0.6862	0.7094	0.7237	0.9106	0.7980
	0.08	0.7680	0.7043	0.7990	0.6981	0.9122	0.7889
	0.1	0.7580	0.7319	0.7992	0.6732	0.9049	0.8127
	0.3	0.7577	0.7640	0.7146	0.7375	0.7328	0.7119
	0.6	0.7491	0.7687	0.7397	0.7609	0.7128	0.6624
30	**0.03**	0.5959	0.6274	0.6274	0.7168	0.8049	0.8043
	0.05	0.6014	0.6149	0.6465	0.6932	0.8122	0.8173
	0.08	0.6477	0.6266	0.6418	0.6796	0.8047	0.8192
	0.1	0.6653	0.6400	0.6526	0.6650	0.8019	0.8258
	0.3	0.6994	0.7195	0.7279	0.7229	0.7996	0.8461
	0.6	0.7608	0.7657	0.7778	0.7816	0.8576	0.8845
40	**0.03**	0.4685	0.5169	0.5695	0.6015	0.7461	0.6876
	0.05	0.4608	0.4839	0.5353	0.5688	0.7357	0.6940
	0.08	0.4983	0.4992	0.5191	0.5490	0.7156	0.6872
	0.1	0.5110	0.5211	0.5282	0.5381	0.7198	0.6982
	0.3	0.6166	0.6241	0.6379	0.6166	0.6805	0.7297
	0.6	0.6414	0.6452	0.6559	0.6493	0.7379	0.7062
50	**0.03**	0.4404	0.5039	0.5546	0.5747	0.6323	0.6435
	0.05	0.4322	0.4673	0.5185	0.5436	0.6569	0.6443
	0.08	0.4595	0.4626	0.4702	0.4998	0.6359	0.6429
	0.1	0.4847	0.4726	0.4810	0.4872	0.6214	0.6427
	0.3	0.5649	0.5722	0.5771	0.5871	0.6323	0.6401
	0.6	0.6019	0.6013	0.6093	0.6068	0.6506	0.6274
60	**0.03**	0.3861	0.4065	0.4498	0.4665	0.6062	0.5661
	0.05	0.3772	0.3977	0.4149	0.4373	0.6054	0.5682
	0.08	0.4122	0.3986	0.4068	0.4095	0.5849	0.5785
	0.1	0.4313	0.4061	0.4016	0.4101	0.5743	0.5757
	0.3	0.5292	0.5757	0.5444	0.5398	0.5458	0.5593
	0.6	0.5101	0.5205	0.5263	0.5372	0.5691	0.5161
70	**0.03**	0.3552	0.3900	0.4253	0.4382	0.5525	0.5124
	0.05	0.3691	0.3767	0.3970	0.4158	0.5683	0.5050
	0.08	0.4003	0.3800	0.3802	0.3913	0.5454	0.5221
	0.1	0.4112	0.4048	0.3790	0.3833	0.5421	0.5106
	0.3	0.4630	0.4694	0.4700	0.4537	0.4887	0.5048
	0.6	0.5048	0.5008	0.5151	0.5151	0.5201	0.4490

The 5% upper percentiles of the $\left|COVRATIO_{(-j)} - 1\right|$ statistic at a = -1, continued.

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
80	**0.03**	0.3552	0.3495	0.3967	0.4119	0.5092	0.4775
	0.05	0.3104	0.3151	0.3587	0.3918	0.5159	0.4695
	0.08	0.3385	0.3260	0.3247	0.3484	0.5007	0.4830
	0.1	0.3496	0.3474	0.3395	0.3375	0.4927	0.4941
	0.3	0.4369	0.4391	0.4559	0.4432	0.4633	0.4667
	0.6	0.4469	0.3329	0.4722	0.4741	0.4640	0.3756
90	**0.03**	0.2866	0.3257	0.3826	0.4012	0.4980	0.4302
	0.05	0.2987	0.2921	0.3313	0.3571	0.5051	0.4336
	0.08	0.3291	0.2984	0.3040	0.3178	0.4923	0.4601
	0.1	0.3327	0.3186	0.3026	0.3154	0.4928	0.4548
	0.3	0.4285	0.4383	0.4377	0.4412	0.4273	0.4273
	0.6	0.4091	0.4190	0.4280	0.4333	0.4537	0.3806
100	**0.03**	0.2684	0.3070	0.3378	0.3639	0.4528	0.4109
	0.05	0.2820	0.2932	0.3063	0.3354	0.4456	0.4214
	0.08	0.3045	0.2933	0.2991	0.3084	0.4458	0.4163
	0.1	0.3141	0.3019	0.3072	0.3129	0.4405	0.4178
	0.3	0.3988	0.3924	0.4088	0.3973	0.4329	0.3789
	0.6	0.3776	0.3831	0.3790	0.3830	0.4027	0.3194
110	**0.03**	0.2438	0.2742	0.3075	0.3296	0.4403	0.3602
	0.05	0.2398	0.2578	0.2878	0.3067	0.4349	0.3556
	0.08	0.2564	0.2437	0.2666	0.2870	0.4371	0.3670
	0.1	0.2693	0.2531	0.2646	0.2707	0.4176	0.3683
	0.3	0.3536	0.3530	0.3456	0.3384	0.3766	0.3377
	0.6	0.3444	0.3453	0.3505	0.3547	0.3903	0.2720
130	**0.03**	0.2097	0.2309	0.2634	0.2849	0.4027	0.3166
	0.05	0.2264	0.2207	0.2399	0.2594	0.2264	0.3128
	0.08	0.2470	0.2336	0.2326	0.2351	0.3989	0.3213
	0.1	0.2597	0.2539	0.2422	0.2385	0.388	0.3213
	0.3	0.3475	0.3473	0.3587	0.3533	0.3446	0.3112
	0.6	0.3180	0.3313	0.3411	0.3411	0.3476	0.2519
150	**0.03**	0.1909	0.2195	0.2420	0.2195	0.3937	0.2847
	0.05	0.1961	0.1972	0.2195	0.2266	0.3970	0.2812
	0.08	0.2145	0.2025	0.2060	0.2147	0.3914	0.2849
	0.1	0.2197	0.2124	0.2049	0.2092	0.3875	0.2874
	0.3	0.3504	0.3644	0.3461	0.3335	0.3263	0.2744
	0.6	0.2996	0.3032	0.2925	0.2897	0.3052	0.2244

The 10% upper percentiles of the $\left| COVRATIO_{(-j)} - 1 \right|$ statistic at $a = -1$

n	σ_1	σ_2 0.03	0.05	0.08	0.1	0.3	0.6
10	**0.03**	0.8954	0.9078	0.9153	0.9249	0.9389	0.9380
	0.05	0.9006	0.8997	0.9380	0.9086	0.9354	0.9325
	0.08	0.8991	0.9088	0.9033	0.9123	0.9287	0.9319
	0.1	0.8980	0.9083	0.9200	0.9125	0.9214	0.9260
	0.3	0.9087	0.9080	0.9032	0.8985	0.9187	0.9280
	0.6	0.9148	0.9198	0.9144	0.9086	0.9164	0.9019
20	**0.03**	0.6318	0.6510	0.6770	0.6869	0.7495	0.7301
	0.05	0.6304	0.6288	0.6419	0.6648	0.8677	0.7276
	0.08	0.6971	0.6325	0.7499	0.6268	0.8603	0.7182
	0.1	0.7009	0.6537	0.7361	0.6305	0.8580	0.7552
	0.3	0.7168	0.6986	0.6705	0.6822	0.6577	0.6488
	0.6	0.6917	0.7022	0.6774	0.6996	0.6519	0.5843
30	**0.03**	0.5288	0.5757	0.5757	0.6477	0.7477	0.7522
	0.05	0.5490	0.5572	0.5929	0.6236	0.7540	0.7630
	0.08	0.5790	0.5801	0.5887	0.6083	0.7483	0.7684
	0.1	0.5899	0.5995	0.6020	0.6094	0.7315	0.7736
	0.3	0.6427	0.6558	0.6554	0.6507	0.7279	0.7843
	0.6	0.6968	0.7067	0.7235	0.7154	0.7893	0.8009
40	**0.03**	0.4295	0.4677	0.5096	0.5301	0.6675	0.6414
	0.05	0.4209	0.445	0.4762	0.5021	0.6706	0.6508
	0.08	0.4483	0.4437	0.4699	0.4810	0.6609	0.6434
	0.1	0.4723	0.4665	0.4662	0.4775	0.6367	0.6480
	0.3	0.5340	0.5452	0.5508	0.5340	0.6212	0.6328
	0.6	0.5868	0.5898	0.5833	0.5956	0.6687	0.6480
50	**0.03**	0.3935	0.4378	0.4923	0.4986	0.5705	0.5703
	0.05	0.3854	0.4102	0.4515	0.4803	0.5656	0.5699
	0.08	0.4133	0.3993	0.4279	0.4412	0.5574	0.5684
	0.1	0.4233	0.4178	0.4124	0.4377	0.5473	0.5665
	0.3	0.5092	0.5135	0.5106	0.5006	0.5705	0.5828
	0.6	0.5596	0.5578	0.5552	0.5518	0.5643	0.5433
60	**0.03**	0.3222	0.3567	0.3862	0.4034	0.5401	0.5042
	0.05	0.3442	0.3359	0.3735	0.3849	0.5283	0.5121
	0.08	0.3714	0.3580	0.3481	0.3516	0.5088	0.5143
	0.1	0.3857	0.3688	0.3623	0.3589	0.5066	0.5187
	0.3	0.4563	0.5187	0.4582	0.4517	0.4750	0.4992
	0.6	0.4624	0.4724	0.4727	0.4805	0.5164	0.4437
70	**0.03**	0.3137	0.3504	0.3813	0.3928	0.4924	0.4517
	0.05	0.3243	0.3280	0.3664	0.3834	0.4995	0.4562
	0.08	0.3356	0.3296	0.3433	0.3517	0.4902	0.4653
	0.1	0.3425	0.3312	0.3379	0.3502	0.4782	0.4633
	0.3	0.4035	0.4035	0.3985	0.3949	0.4434	0.4300
	0.6	0.4300	0.4318	0.4450	0.4448	0.4504	0.3956

The 10% upper percentiles of the $\left| COVRATIO_{(-j)} - 1 \right|$ statistic at a = -1, continued.

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
80	**0.03**	0.3137	0.3087	0.3421	0.3568	0.4592	0.4223
	0.05	0.2751	0.2829	0.3099	0.3337	0.4451	0.4296
	0.08	0.3006	0.2948	0.3014	0.3057	0.4360	0.4304
	0.1	0.3095	0.3082	0.2968	0.3039	0.4279	0.4418
	0.3	0.3803	0.3890	0.3957	0.3875	0.3878	0.4198
	0.6	0.3960	0.3053	0.4192	0.4078	0.4105	0.3483
90	**0.03**	0.2517	0.2841	0.3256	0.3387	0.4397	0.3832
	0.05	0.2527	0.2581	0.2940	0.3139	0.4404	0.3876
	0.08	0.2800	0.2669	0.2690	0.2915	0.4342	0.4012
	0.1	0.3021	0.2801	0.2710	0.2787	0.4215	0.4055
	0.3	0.3633	0.3630	0.3658	0.3607	0.3787	0.3787
	0.6	0.3650	0.3670	0.3674	0.3722	0.3999	0.3208
100	**0.03**	0.2379	0.2620	0.2923	0.3079	0.3818	0.3554
	0.05	0.2423	0.2488	0.2774	0.2955	0.3928	0.3587
	0.08	0.2690	0.2568	0.2623	0.2757	0.3865	0.3634
	0.1	0.2789	0.2674	0.2645	0.2703	0.3819	0.3628
	0.3	0.3529	0.3567	0.3502	0.3426	0.3563	0.3443
	0.6	0.3495	0.3556	0.3540	0.3539	0.3815	0.2935
110	**0.03**	0.1867	0.2065	0.2407	0.2502	0.3536	0.2910
	0.05	0.1886	0.1899	0.2159	0.2305	0.1886	0.2905
	0.08	0.2142	0.1987	0.2049	0.2158	0.3277	0.2904
	0.1	0.2230	0.2128	0.2054	0.2096	0.3154	0.2948
	0.3	0.2934	0.2953	0.2891	0.2824	0.2897	0.2792
	0.6	0.2800	0.2948	0.2974	0.2940	0.3028	0.2310
130	**0.03**	0.1713	0.1901	0.2086	0.1901	0.3381	0.2612
	0.05	0.1689	0.1776	0.1972	0.2101	0.3335	0.2605
	0.08	0.1898	0.1751	0.1803	0.1908	0.3267	0.2576
	0.1	0.1976	0.1864	0.1812	0.1918	0.3188	0.2589
	0.3	0.2845	0.2843	0.2774	0.2769	0.2906	0.2424
	0.6	0.2658	0.2674	0.2647	0.2586	0.2633	0.2045
150	**0.03**	0.1867	0.2065	0.2407	0.2502	0.3536	0.2910
	0.05	0.1886	0.1899	0.2159	0.2305	0.1886	0.2905
	0.08	0.2142	0.1987	0.2049	0.2158	0.3277	0.2904
	0.1	0.2230	0.2128	0.2054	0.2096	0.3154	0.2948
	0.3	0.2934	0.2953	0.2891	0.2824	0.2897	0.2792
	0.6	0.2800	0.2948	0.2974	0.2940	0.3028	0.2310

APPENDIX 3

The 1% upper percentiles of the *DMCEc* statistic at a = -1

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
10	**0.03**	0.0004	0.0009	0.0019	0.0031	0.0367	0.2080
	0.05	0.0010	0.0012	0.0022	0.0032	0.0010	0.1907
	0.08	0.0029	0.0026	0.0026	0.0034	0.0361	0.2111
	0.1	0.0037	0.0042	0.0051	0.0058	0.0459	0.2091
	0.3	0.0349	0.0337	0.0339	0.0340	0.0782	0.2096
	0.6	0.1923	0.2044	0.1844	0.2201	0.2115	0.2432
20	**0.03**	0.0002	0.0005	0.0013	0.0020	0.0297	0.0962
	0.05	0.0006	0.0007	0.0014	0.0022	0.0006	0.0966
	0.08	0.0015	0.0014	0.0016	0.0023	0.0290	0.0990
	0.1	0.0022	0.0023	0.0025	0.0027	0.0367	0.0969
	0.3	0.0254	0.0239	0.0253	0.0243	0.0457	0.0981
	0.6	0.0985	0.0940	0.0966	0.0949	0.0990	0.1148
30	**0.03**	0.0002	0.0005	0.0012	0.0019	0.0262	0.1148
	0.05	0.0005	0.0006	0.0012	0.0018	0.0251	0.1245
	0.08	0.0011	0.0012	0.0016	0.0022	0.0276	0.1215
	0.1	0.0018	0.0018	0.0022	0.0026	0.0262	0.1167
	0.3	0.0259	0.0254	0.0260	0.0264	0.0560	0.1586
	0.6	0.1122	0.1159	0.0807	0.1216	0.1476	0.2073
40	**0.03**	0.0002	0.0003	0.0008	0.0015	0.0175	0.0813
	0.05	0.0004	0.0005	0.0009	0.0015	0.0180	0.0854
	0.08	0.0010	0.0011	0.0013	0.0016	0.0194	0.0879
	0.1	0.0017	0.0017	0.0020	0.0022	0.0206	0.0919
	0.3	0.0185	0.0184	0.0183	0.0186	0.0606	0.1062
	0.6	0.0823	0.0827	0.0862	0.0877	0.1069	0.1677
50	**0.03**	0.0001	0.0002	0.0007	0.0010	0.0249	0.0685
	0.05	0.0003	0.0003	0.0007	0.0011	0.0250	0.0692
	0.08	0.0009	0.0009	0.0010	0.0012	0.0273	0.0730
	0.1	0.0014	0.0013	0.0016	0.0017	0.0273	0.0738
	0.3	0.0170	0.0173	0.0177	0.0181	0.0476	0.0887
	0.6	0.0584	0.0614	0.0638	0.0644	0.0892	0.1307
60	**0.03**	0.0001	0.0002	0.0006	0.0010	0.0131	0.0532
	0.05	0.0003	0.0003	0.0006	0.0010	0.0135	0.0536
	0.08	0.0007	0.0007	0.0008	0.0011	0.0140	0.0564
	0.1	0.0011	0.0011	0.0012	0.0013	0.0148	0.0565
	0.3	0.0169	0.0180	0.0184	0.0191	0.0311	0.0753
	0.6	0.0517	0.0540	0.0570	0.0587	0.0713	0.1013
70	**0.03**	0.0001	0.0002	0.0005	0.0009	0.0128	0.0474
	0.05	0.0002	0.0003	0.0005	0.0009	0.0132	0.0473
	0.08	0.0006	0.0006	0.0008	0.0010	0.0137	0.0498
	0.1	0.0009	0.0010	0.0010	0.0014	0.0139	0.0493
	0.3	0.0132	0.0141	0.0145	0.0148	0.0305	0.0597
	0.6	0.0437	0.0454	0.0463	0.0484	0.0640	0.0852

The 1% upper percentiles of the *DMCEc* statistic at *a* = -1, continued.

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
80	**0.03**	0.0001	0.0002	0.0005	0.0009	0.0137	0.0415
	0.05	0.0002	0.0002	0.0006	0.0009	0.0138	0.0425
	0.08	0.0005	0.0005	0.0007	0.0010	0.0136	0.0431
	0.1	0.0008	0.0008	0.0010	0.0012	0.0144	0.0433
	0.3	0.0110	0.0115	0.0125	0.0128	0.0319	0.0527
	0.6	0.0384	0.0407	0.0430	0.0442	0.0542	0.0679
90	**0.03**	0.0001	0.0002	0.0005	0.0007	0.0147	0.0363
	0.05	0.0002	0.0002	0.0005	0.0007	0.0159	0.0380
	0.08	0.0005	0.0005	0.0006	0.0008	0.0184	0.0389
	0.1	0.0008	0.0008	0.0009	0.0010	0.0193	0.0392
	0.3	0.0117	0.0116	0.0111	0.0111	0.0248	0.0461
	0.6	0.0346	0.0364	0.0374	0.0403	0.0497	0.0548
100	**0.03**	0.0001	0.0002	0.0004	0.0007	0.0197	0.0318
	0.05	0.0002	0.0002	0.0005	0.0008	0.0201	0.0328
	0.08	0.0004	0.0004	0.0006	0.0010	0.0219	0.0342
	0.1	0.0007	0.0007	0.0008	0.0011	0.0220	0.0356
	0.3	0.0098	0.0106	0.0111	0.0113	0.0297	0.0408
	0.6	0.0311	0.0325	0.0337	0.0345	0.0437	0.0546
110	**0.03**	0.0001	0.0001	0.0004	0.0006	0.0124	0.0288
	0.05	0.0002	0.0002	0.0004	0.0006	0.0133	0.0293
	0.08	0.0004	0.0004	0.0005	0.0007	0.0140	0.0309
	0.1	0.0006	0.0006	0.0007	0.0009	0.0142	0.0309
	0.3	0.0091	0.0092	0.0095	0.0114	0.0242	0.0388
	0.6	0.0279	0.0294	0.0299	0.0305	0.0383	0.0486
130	**0.03**	0.0001	0.0001	0.0003	0.0006	0.0099	0.0245
	0.05	0.0001	0.0002	0.0003	0.0006	0.0100	0.0247
	0.08	0.0004	0.0004	0.0005	0.0006	0.0101	0.0255
	0.1	0.0006	0.0006	0.0006	0.0007	0.0101	0.0256
	0.3	0.0094	0.0099	0.0108	0.0115	0.0125	0.0319
	0.6	0.0242	0.0249	0.0261	0.0264	0.0306	0.0388
150	**0.03**	0.0001	0.0001	0.0003	0.0006	0.0081	0.0217
	0.05	0.0001	0.0001	0.0003	0.0006	0.0100	0.0219
	0.08	0.0003	0.0003	0.0004	0.0006	0.0101	0.0227
	0.1	0.0005	0.0005	0.0005	0.0007	0.0102	0.0233
	0.3	0.0066	0.0070	0.0071	0.0074	0.0142	0.0256
	0.6	0.0199	0.0209	0.0214	0.0279	0.0299	0.0305

The 5% upper percentiles of the *DMCEc* statistic at a = -1

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
10	**0.03**	0.0003	0.0006	0.0014	0.0022	0.0215	0.1634
	0.05	0.0007	0.0008	0.0015	0.0022	0.0007	0.1459
	0.08	0.0017	0.0018	0.0021	0.0026	0.0237	0.1517
	0.1	0.0026	0.0027	0.0030	0.0037	0.0267	0.1631
	0.3	0.0239	0.0228	0.0229	0.0236	0.0425	0.1824
	0.6	0.0944	0.1423	0.1151	0.1541	0.1554	0.2118
20	**0.03**	0.0002	0.0004	0.0008	0.0013	0.0153	0.0907
	0.05	0.0004	0.0005	0.0010	0.0015	0.0004	0.0904
	0.08	0.0011	0.0011	0.0013	0.0017	0.0163	0.0906
	0.1	0.0017	0.0017	0.0018	0.0022	0.0163	0.0911
	0.3	0.0165	0.0164	0.0176	0.0167	0.0284	0.0938
	0.6	0.0768	0.0778	0.0803	0.0784	0.0807	0.0961
30	**0.03**	0.0001	0.0003	0.0009	0.0013	0.0163	0.0974
	0.05	0.0003	0.0004	0.0009	0.0014	0.0167	0.0918
	0.08	0.0008	0.0009	0.0012	0.0016	0.0164	0.1055
	0.1	0.0013	0.0070	0.0016	0.0019	0.0151	0.1022
	0.3	0.0148	0.0154	0.0156	0.0159	0.0321	0.1232
	0.6	0.0815	0.0834	0.0866	0.0902	0.1067	0.1654
40	**0.03**	0.0001	0.0002	0.0006	0.0009	0.0115	0.0730
	0.05	0.0003	0.0004	0.0007	0.0010	0.0117	0.0734
	0.08	0.0007	0.0007	0.0010	0.0013	0.0119	0.0748
	0.1	0.0011	0.0012	0.0013	0.0016	0.0122	0.0754
	0.3	0.0116	0.0118	0.0127	0.0131	0.0257	0.0906
	0.6	0.0687	0.0703	0.0751	0.0772	0.0882	0.1253
50	**0.03**	0.0001	0.0002	0.0005	0.0008	0.0104	0.0594
	0.05	0.0002	0.0003	0.0005	0.0008	0.0105	0.0605
	0.08	0.0006	0.0006	0.0008	0.0010	0.0112	0.0622
	0.1	0.0008	0.0009	0.0011	0.0012	0.0116	0.0639
	0.3	0.0112	0.0120	0.0129	0.0129	0.0243	0.0776
	0.6	0.0503	0.0522	0.0553	0.0553	0.0724	0.0975
60	**0.03**	0.0001	0.0002	0.0005	0.0007	0.0095	0.0483
	0.05	0.0002	0.0002	0.0005	0.0007	0.0098	0.0491
	0.08	0.0005	0.0005	0.0006	0.0008	0.0104	0.0500
	0.1	0.0008	0.0008	0.0009	0.0010	0.0106	0.0518
	0.3	0.0097	0.0099	0.0104	0.0111	0.0198	0.0630
	0.6	0.0461	0.0471	0.0491	0.0507	0.0630	0.0768
70	**0.03**	0.0001	0.0002	0.0004	0.0006	0.0090	0.0412
	0.05	0.0002	0.0002	0.0004	0.0006	0.0091	0.0416
	0.08	0.0004	0.0004	0.0006	0.0007	0.0092	0.0439
	0.1	0.0007	0.0007	0.0008	0.0010	0.0094	0.0448
	0.3	0.0078	0.0080	0.0083	0.0087	0.0171	0.0515
	0.6	0.0398	0.0413	0.0431	0.0437	0.0531	0.0711

The 5% upper percentiles of the *DMCEc* statistic at a = -1, continued.

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
80	**0.03**	0.0001	0.0002	0.0004	0.0006	0.0084	0.0366
	0.05	0.0002	0.0002	0.0004	0.0006	0.0086	0.0377
	0.08	0.0004	0.0004	0.0005	0.0007	0.0087	0.0391
	0.1	0.0006	0.0006	0.0007	0.0008	0.0089	0.0396
	0.3	0.0073	0.0074	0.0082	0.0084	0.0162	0.0466
	0.6	0.0351	0.0370	0.0384	0.0392	0.0458	0.0582
90	**0.03**	0.0001	0.0001	0.0003	0.0005	0.0075	0.0334
	0.05	0.0001	0.0002	0.0004	0.0005	0.0078	0.0340
	0.08	0.0004	0.0004	0.0005	0.0007	0.0080	0.0351
	0.1	0.0006	0.0006	0.0007	0.0008	0.0080	0.0359
	0.3	0.0067	0.0068	0.0070	0.0074	0.0173	0.0415
	0.6	0.0319	0.0328	0.0337	0.0353	0.0414	0.0434
100	**0.03**	0.0001	0.0001	0.0003	0.0005	0.0078	0.0293
	0.05	0.0001	0.0002	0.0003	0.0005	0.0081	0.0298
	0.08	0.0003	0.0004	0.0004	0.0006	0.0083	0.0308
	0.1	0.0005	0.0005	0.0006	0.0007	0.0085	0.0313
	0.3	0.0063	0.0064	0.0068	0.0071	0.0152	0.0367
	0.6	0.0283	0.0296	0.0306	0.0311	0.0382	0.0460
110	**0.03**	0.0001	0.0001	0.0003	0.0005	0.0069	0.0270
	0.05	0.0001	0.0002	0.0003	0.0005	0.0069	0.0274
	0.08	0.0003	0.0003	0.0004	0.0005	0.0070	0.0283
	0.1	0.0005	0.0005	0.0006	0.0007	0.0072	0.0289
	0.3	0.0054	0.0055	0.0060	0.0068	0.0137	0.0347
	0.6	0.0254	0.0264	0.0272	0.0285	0.0329	0.0415
130	**0.03**	0.0000	0.0001	0.0003	0.0004	0.0059	0.0227
	0.05	0.0001	0.0001	0.0003	0.0004	0.0061	0.0232
	0.08	0.0003	0.0003	0.0004	0.0005	0.0063	0.0240
	0.1	0.0004	0.0004	0.0005	0.0006	0.0064	0.0244
	0.3	0.0053	0.0054	0.0056	0.0058	0.0069	0.0274
	0.6	0.0219	0.0224	0.0236	0.0243	0.0288	0.0283
150	**0.03**	0.0001	0.0001	0.0003	0.0004	0.0056	0.0198
	0.05	0.0001	0.0001	0.0003	0.0004	0.0056	0.0201
	0.08	0.0002	0.0002	0.0003	0.0004	0.0059	0.0208
	0.1	0.0004	0.0004	0.0004	0.0005	0.0062	0.0210
	0.3	0.0048	0.0049	0.0049	0.0051	0.0106	0.0241
	0.6	0.0186	0.0191	0.0199	0.0219	0.0239	0.0288

The 10% upper percentiles of the *DMCEc* statistic at a = -1

n	σ_1	σ_2 0.03	0.05	0.08	0.1	0.3	0.6
10	**0.03**	0.0002	0.0005	0.0011	0.0016	0.0164	0.1174
	0.05	0.0005	0.0007	0.0012	0.0017	0.0005	0.1073
	0.08	0.0013	0.0014	0.0017	0.0022	0.0187	0.1130
	0.1	0.0021	0.0022	0.0025	0.0029	0.0192	0.1275
	0.3	0.0187	0.0178	0.0186	0.0188	0.0320	0.1379
	0.6	0.0615	0.0893	0.0678	0.0931	0.1024	0.1921
20	**0.03**	0.0002	0.0003	0.0007	0.0010	0.0122	0.0756
	0.05	0.0003	0.0004	0.0008	0.0012	0.0003	0.0786
	0.08	0.0008	0.0008	0.0011	0.0014	0.0122	0.0857
	0.1	0.0013	0.0014	0.0016	0.0018	0.0127	0.0813
	0.3	0.0129	0.0123	0.0135	0.0133	0.0223	0.0863
	0.6	0.0625	0.0630	0.0705	0.0634	0.0678	0.0873
30	**0.03**	0.0001	0.0003	0.0006	0.0010	0.0118	0.0844
	0.05	0.0003	0.0003	0.0007	0.0011	0.0127	0.0779
	0.08	0.0007	0.0008	0.0010	0.0013	0.0122	0.0926
	0.1	0.0011	0.0012	0.0014	0.0016	0.0122	0.0890
	0.3	0.0118	0.0121	0.0132	0.0132	0.0246	0.1093
	0.6	0.0601	0.0611	0.0637	0.0651	0.0848	0.1410
40	**0.03**	0.0001	0.0002	0.0005	0.0008	0.0098	0.0650
	0.05	0.0002	0.0003	0.0006	0.0009	0.0100	0.0658
	0.08	0.0006	0.0006	0.0008	0.0011	0.0101	0.0684
	0.1	0.0009	0.0010	0.0010	0.0013	0.0103	0.0919
	0.3	0.0092	0.0094	0.0100	0.0106	0.0204	0.0849
	0.6	0.0586	0.0586	0.0632	0.0643	0.0779	0.1142
50	**0.03**	0.0001	0.0002	0.0004	0.0007	0.0082	0.0545
	0.05	0.0002	0.0002	0.0005	0.0007	0.0085	0.0561
	0.08	0.0005	0.0005	0.0007	0.0008	0.0087	0.0583
	0.1	0.0008	0.0008	0.0009	0.0011	0.0090	0.0595
	0.3	0.0086	0.0088	0.0091	0.0092	0.0182	0.0696
	0.6	0.0445	0.0452	0.0460	0.0487	0.0641	0.0883
60	**0.03**	0.0001	0.0002	0.0004	0.0006	0.0077	0.0453
	0.05	0.0002	0.0002	0.0004	0.0006	0.0079	0.0463
	0.08	0.0004	0.0004	0.0006	0.0008	0.0081	0.0476
	0.1	0.0006	0.0007	0.0008	0.0009	0.0083	0.0484
	0.3	0.0076	0.0080	0.0084	0.0087	0.0152	0.0581
	0.6	0.0416	0.0432	0.0451	0.0466	0.0570	0.0731
70	**0.03**	0.0001	0.0001	0.0003	0.0005	0.0073	0.0387
	0.05	0.0001	0.0002	0.0004	0.0006	0.0076	0.0395
	0.08	0.0004	0.0004	0.0005	0.0007	0.0078	0.0399
	0.1	0.0006	0.0006	0.0007	0.0008	0.0081	0.0408
	0.3	0.0065	0.0066	0.0070	0.0072	0.0144	0.0488
	0.6	0.0370	0.0381	0.0397	0.0407	0.0479	0.0628

The 10% upper percentiles of the *DMCEc* statistic at *a* = -1, continued.

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
80	**0.03**	0.0001	0.0001	0.0003	0.0005	0.0069	0.0343
	0.05	0.0001	0.0002	0.0003	0.0005	0.0069	0.0349
	0.08	0.0003	0.0003	0.0005	0.0006	0.0071	0.0361
	0.1	0.0005	0.0005	0.0006	0.0007	0.0073	0.0367
	0.3	0.0055	0.0057	0.0060	0.0064	0.0129	0.0434
	0.6	0.0325	0.0336	0.0352	0.0362	0.0420	0.0539
90	**0.03**	0.0001	0.0001	0.0003	0.0005	0.0057	0.0310
	0.05	0.0001	0.0002	0.0003	0.0005	0.0059	0.0315
	0.08	0.0003	0.0003	0.0004	0.0006	0.0062	0.0333
	0.1	0.0005	0.0005	0.0006	0.0007	0.0064	0.0332
	0.3	0.0054	0.0056	0.0059	0.0059	0.0121	0.0395
	0.6	0.0292	0.0303	0.0316	0.0323	0.0384	0.0443
100	**0.03**	0.0000	0.0001	0.0003	0.0004	0.0059	0.0279
	0.05	0.0001	0.0001	0.0003	0.0005	0.0060	0.0287
	0.08	0.0003	0.0003	0.0004	0.0006	0.0062	0.0294
	0.1	0.0004	0.0005	0.0005	0.0007	0.0064	0.0301
	0.3	0.0047	0.0048	0.0050	0.0052	0.0087	0.0349
	0.6	0.0269	0.0278	0.0286	0.0294	0.0350	0.0439
110	**0.03**	0.0000	0.0001	0.0002	0.0004	0.0057	0.0256
	0.05	0.0001	0.0001	0.0003	0.0004	0.0057	0.0264
	0.08	0.0003	0.0003	0.0004	0.0005	0.0059	0.0270
	0.1	0.0004	0.0005	0.0005	0.0006	0.0060	0.0276
	0.3	0.0043	0.0045	0.0045	0.0059	0.0107	0.0327
	0.6	0.0238	0.0247	0.0259	0.0262	0.0311	0.0388
130	**0.03**	0.0000	0.0001	0.0002	0.0003	0.0047	0.0218
	0.05	0.0001	0.0001	0.0002	0.0004	0.0049	0.0222
	0.08	0.0002	0.0002	0.0003	0.0004	0.0051	0.0230
	0.1	0.0004	0.0004	0.0004	0.0005	0.0052	0.0234
	0.3	0.0043	0.0045	0.0046	0.0047	0.0059	0.0276
	0.6	0.0205	0.0214	0.0219	0.0226	0.0267	0.0327
150	**0.03**	0.0000	0.0001	0.0002	0.0004	0.0044	0.0188
	0.05	0.0001	0.0001	0.0002	0.0004	0.0045	0.0193
	0.08	0.0002	0.0002	0.0003	0.0004	0.0047	0.0197
	0.1	0.0003	0.0003	0.0004	0.0005	0.0048	0.0202
	0.3	0.0035	0.0036	0.0037	0.0039	0.0081	0.0232
	0.6	0.0181	0.0185	0.0191	0.0205	0.0214	0.0267

The 1% upper percentiles of the *DMCEs* statistic at a = -1

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
10	**0.03**	0.0061	0.0085	0.0136	0.0156	0.0624	0.1178
	0.05	0.0086	0.0092	0.0135	0.0156	0.0091	0.1198
	0.08	0.0129	0.0134	0.0150	0.0175	0.0549	0.1231
	0.1	0.0155	0.0171	0.0173	0.0210	0.0596	0.1302
	0.3	0.0567	0.0513	0.0525	0.0476	0.0791	0.1212
	0.6	0.1139	0.1302	0.1179	0.1364	0.1139	0.1409
20	**0.03**	0.0026	0.0040	0.0063	0.0077	0.0281	0.0577
	0.05	0.0039	0.0039	0.0058	0.0077	0.0039	0.0555
	0.08	0.0062	0.0059	0.0074	0.0084	0.0289	0.0595
	0.1	0.0078	0.0077	0.0082	0.0088	0.0299	0.0554
	0.3	0.0283	0.0275	0.0287	0.0268	0.0377	0.0527
	0.6	0.0608	0.0554	0.0635	0.0559	0.0791	0.0822
30	**0.03**	0.0019	0.0033	0.0056	0.0075	0.0309	0.0720
	0.05	0.0030	0.0036	0.0057233	0.0071	0.0286	0.0728
	0.08	0.0050	0.0052	0.0063	0.0079	0.0315	0.0750
	0.1	0.0062	0.0070	0.0072	0.0084	0.0373	0.0686
	0.3	0.0247	0.0265	0.0288	0.0299	0.0469	0.0876
	0.6	0.0675	0.0697	0.0723	0.0718	0.0812	0.1176
40	**0.03**	0.0015	0.0025	0.0038	0.0045	0.0195	0.0474
	0.05	0.0026	0.0028	0.0039	0.0049	0.0200	0.0491
	0.08	0.0039	0.0041	0.0050	0.0059	0.0203	0.0497
	0.1	0.0049	0.0052	0.0059	0.0063	0.0218	0.0507
	0.3	0.0174	0.0187	0.0199	0.0196	0.0352	0.0633
	0.6	0.0474	0.0499	0.0528	0.0548	0.0598	0.0863
50	**0.03**	0.0011	0.0016	0.0029	0.0037	0.0175	0.0376
	0.05	0.0018	0.0021	0.0027	0.0035	0.0175	0.0395
	0.08	0.0030	0.0033	0.0034	0.0041	0.0179	0.0422
	0.1	0.0037	0.0041	0.0044	0.0047	0.0185	0.0403
	0.3	0.0165	0.0164	0.0164	0.0197	0.0261	0.0494
	0.6	0.0370	0.0380	0.0408	0.0428	0.0498	0.0657
60	**0.03**	0.0010	0.0015	0.0025	0.0031	0.0129	0.0321
	0.05	0.0015	0.0016	0.0025	0.0032	0.0127	0.0323
	0.08	0.0024	0.0025	0.0028	0.0033	0.0141	0.0329
	0.1	0.0030	0.0032	0.0033	0.0037	0.0139	0.0329
	0.3	0.0137	0.0141	0.0151	0.0155	0.0188	0.0395
	0.6	0.0303	0.0311	0.0328	0.0319	0.0386	0.0532
70	**0.03**	0.0008	0.0013	0.0023	0.0030	0.0112	0.0266
	0.05	0.0013	0.0015	0.0022	0.0031	0.0114	0.0263
	0.08	0.0022	0.0024	0.0026	0.0032	0.0118	0.0266
	0.1	0.0027	0.0028	0.0031	0.0033	0.0125	0.0271
	0.3	0.0110	0.0118	0.0125	0.0125	0.0182	0.0307
	0.6	0.0251	0.0257	0.0263	0.0269	0.0370	0.0439

The 1% upper percentiles of the *DMCEs* statistic at a = -1, continued.

n	σ_1	σ_2 0.03	0.05	0.08	0.1	0.3	0.6
80	0.03	0.0007	0.0012	0.0019	0.0025	0.0112	0.0219
	0.05	0.0011	0.0013	0.0020	0.0025	0.0113	0.0225
	0.08	0.0018	0.0019	0.0021	0.0027	0.0117	0.0237
	0.1	0.0022	0.0024	0.0024	0.0027	0.0122	0.0237
	0.3	0.0097	0.0103	0.0109	0.0113	0.0165	0.0272
	0.6	0.0237	0.0242	0.0250	0.0260	0.0277	0.0365
90	0.03	0.0007	0.0010	0.0016	0.0022	0.0105	0.0207
	0.05	0.0010	0.0012	0.0016	0.0022	0.0111	0.0209
	0.08	0.0016	0.0017	0.0020	0.0022	0.0115	0.0203
	0.1	0.0021	0.0021	0.0023	0.0025	0.0113	0.0220
	0.3	0.0083	0.0084	0.0087	0.0089	0.0149	0.0241
	0.6	0.0207	0.0206	0.0234	0.0231	0.0259	0.0297
100	0.03	0.0006	0.0010	0.0016	0.0020	0.0115	0.0171
	0.05	0.0009	0.0010	0.0016	0.0021	0.0118	0.0178
	0.08	0.0014	0.0015	0.0019	0.0022	0.0120	0.0176
	0.1	0.0019	0.0019	0.0023	0.0024	0.0116	0.0184
	0.3	0.0084	0.0085	0.0087	0.0088	0.0138	0.0207
	0.6	0.0205	0.0206	0.0213	0.0213	0.0252	0.0283
110	0.03	0.0005	0.0008	0.0013	0.0017	0.0082	0.0164
	0.05	0.0008	0.0009	0.0014	0.0018	0.0084	0.0167
	0.08	0.0014	0.0014	0.0016	0.0018	0.0097	0.0173
	0.1	0.0017	0.0018	0.0019	0.0021	0.0096	0.0174
	0.3	0.0073	0.0072	0.0072	0.0095	0.0126	0.0190
	0.6	0.0174	0.0156	0.0170	0.0176	0.0204	0.0250
130	0.03	0.0005	0.0008	0.0012	0.0016	0.0071	0.0132
	0.05	0.0007	0.0008	0.0013	0.0017	0.0070	0.0138
	0.08	0.0012	0.0012	0.0014	0.0016	0.0072	0.0140
	0.1	0.0014	0.0015	0.0016	0.0017	0.0073	0.0142
	0.3	0.0074	0.0077	0.0085	0.0086	0.0095	0.0167
	0.6	0.0143	0.0151	0.0159	0.0159	0.0165	0.0173
150	0.03	0.0004	0.0006	0.0011	0.0014	0.0062	0.0117
	0.05	0.0006	0.0007	0.0011	0.0015	0.0063	0.0119
	0.08	0.0010	0.0010	0.0012	0.0014	0.0063	0.0124
	0.1	0.0013	0.0013	0.0013	0.0015	0.0064	0.0125
	0.3	0.0056	0.0057	0.0062	0.0061	0.0082	0.0134
	0.6	0.0119	0.0121	0.0130	0.0143	0.0149	0.0155

The 5% upper percentiles of the *DMCEs* statistic at a = -1

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
10	**0.03**	0.0043	0.0063	0.0095	0.0116	0.0405	0.0969
	0.05	0.0063	0.0072	0.0100	0.0116	0.0065	0.0956
	0.08	0.0102	0.0108	0.0112	0.0132	0.0396	0.0996
	0.1	0.0133	0.0135	0.0146	0.0161	0.0429	0.1036
	0.3	0.0401	0.0386	0.0403	0.0372	0.0527	0.0956
	0.6	0.0822	0.0989	0.0897	0.0995	0.0906	0.1177
20	**0.03**	0.0021	0.0030	0.0049	0.0059	0.0214	0.0509
	0.05	0.0030	0.0035	0.0048	0.0060	0.0030	0.0483
	0.08	0.0049	0.0050	0.0056	0.0062	0.0219	0.0499
	0.1	0.0060	0.0063	0.0067	0.0072	0.0213	0.0491
	0.3	0.0216	0.0213	0.0219	0.0212	0.0264	0.0470
	0.6	0.0458	0.0458	0.0459	0.0463	0.0479	0.0409
30	**0.03**	0.0016	0.0025	0.0043	0.0057	0.0217	0.0581
	0.05	0.0025	0.0028	0.0045	0.0057	0.0220	0.0527
	0.08	0.0041	0.0043	0.0053	0.0059	0.0232	0.0602
	0.1	0.0051	0.0054	0.0059	0.0067	0.0213	0.0575
	0.3	0.0194	0.0197	0.0210	0.0215	0.0335	0.0676
	0.6	0.0512	0.0544	0.0550	0.0558	0.0669	0.0912
40	**0.03**	0.0013	0.0019	0.0029	0.0039	0.0146	0.0397
	0.05	0.0019	0.0022	0.0031	0.0039	0.0149	0.0397
	0.08	0.0031	0.0032	0.0037	0.0043	0.0151	0.0408
	0.1	0.0039	0.0040	0.0044	0.0049	0.0155	0.0408
	0.3	0.0140	0.0146	0.0148	0.0151	0.0249	0.0468
	0.6	0.0390	0.0398	0.0424	0.0417	0.0505	0.0651
50	**0.03**	0.0010	0.0014	0.0024	0.0031	0.0123	0.0323
	0.05	0.0015	0.0017	0.0024	0.0031	0.0128	0.0339
	0.08	0.0025	0.0025	0.0030	0.0034	0.0134	0.0339
	0.1	0.0031	0.0032	0.0036	0.0039	0.0136	0.0345
	0.3	0.0125	0.0133	0.0132	0.0137	0.0197	0.0384
	0.6	0.0304	0.0312	0.0325	0.0329	0.0384	0.0497
60	**0.03**	0.0008	0.0012	0.0020	0.0025	0.0108	0.0273
	0.05	0.0012	0.0014	0.0020	0.0026	0.0108	0.0281
	0.08	0.0020	0.0021	0.0023	0.0028	0.0108	0.0284
	0.1	0.0026	0.0027	0.0029	0.0031	0.0109	0.0289
	0.3	0.0105	0.0109	0.0111	0.0112	0.0160	0.0316
	0.6	0.0255	0.0263	0.0271	0.0279	0.0321	0.0401
70	**0.03**	0.0007	0.0011	0.0018	0.0023	0.0096	0.0225
	0.05	0.0010	0.0011	0.0018	0.0023	0.0097	0.0224
	0.08	0.0018	0.0018	0.0021	0.0024	0.0098	0.0238
	0.1	0.0022	0.0023	0.0025	0.0028	0.0098	0.0240
	0.3	0.0087	0.0092	0.0096	0.0097	0.0134	0.0258
	0.6	0.0221	0.0229	0.0235	0.0239	0.0272	0.0351

The 5% upper percentiles of the *DMCEs* statistic at a = -1, continued.

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
80	**0.03**	0.0006	0.0009	0.0016	0.0020	0.0087	0.0196
	0.05	0.0009	0.0011	0.0016	0.0020	0.0087	0.0199
	0.08	0.0016	0.0016	0.0018	0.0022	0.0087	0.0200
	0.1	0.0019	0.0020	0.0021	0.0024	0.0088	0.0204
	0.3	0.0075	0.0080	0.0084	0.0083	0.0121	0.0237
	0.6	0.0197921	0.0206	0.0208	0.0210	0.0234	0.0277
90	**0.03**	0.0006	0.0008	0.0014	0.0018	0.0075	0.0178
	0.05	0.0009	0.0010	0.0014	0.0018	0.0076	0.0185
	0.08	0.0014	0.0015	0.0017	0.0019	0.0077	0.0183
	0.1	0.0018	0.0018	0.0020	0.0022	0.0078	0.0190
	0.3	0.0069	0.0068	0.0072	0.0073	0.0115	0.0199
	0.6	0.0173	0.0179	0.0194	0.0187	0.0220	0.0239
100	**0.03**	0.0005	0.0008	0.0013	0.0017	0.0072	0.0155
	0.05	0.0008	0.0009	0.0014	0.0017	0.0075	0.0160
	0.08	0.0013	0.0013	0.0015	0.0018	0.0073	0.0162
	0.1	0.0016	0.0016	0.0018	0.0019	0.0076	0.0167
	0.3	0.0065	0.0065	0.0066	0.0068	0.0105	0.0180
	0.6	0.0160	0.0169	0.0173	0.0176	0.0201	0.0244
110	**0.03**	0.0005	0.0007	0.0012	0.0015	0.0063	0.0144
	0.05	0.0007	0.0008	0.0011	0.0015	0.0062	0.0145
	0.08	0.0012	0.0012	0.0014	0.0016	0.0066	0.0147
	0.1	0.0015	0.0016	0.0016	0.0018	0.0065	0.0149
	0.3	0.0054	0.0057	0.0060	0.0064	0.0097	0.0165
	0.6	0.0135	0.0141	0.0148	0.0150	0.0168	0.0207
130	**0.03**	0.0004	0.0006	0.0010	0.0013	0.0055	0.0123
	0.05	0.0006	0.0007	0.0010	0.0013	0.0055	0.0122
	0.08	0.0010	0.0010	0.0012	0.0014	0.0054	0.0127
	0.1	0.0013	0.0013	0.0014	0.0015	0.0056	0.0127
	0.3	0.0050	0.0053	0.0054	0.0056	0.0069	0.0149
	0.6	0.0119	0.0122	0.0129	0.0135	0.0142	0.0165
150	**0.03**	0.0003	0.0005	0.0009	0.0013	0.0050	0.0106
	0.05	0.0005	0.0006	0.0009	0.0012	0.0051	0.0106
	0.08	0.0009	0.0009	0.0010	0.0012	0.0052	0.0108
	0.1	0.0011	0.0011	0.0012	0.0013	0.0053	0.0109
	0.3	0.0043	0.0044	0.0047	0.0047	0.0066	0.0113
	0.6	0.0102	0.0106	0.0109	0.0119	0.0122	0.0135

The 10% upper percentiles of the *DMCEs* statistic at a = -1

n	σ_1	σ_2					
		0.03	**0.05**	**0.08**	**0.1**	**0.3**	**0.6**
10	**0.03**	0.0036	0.0051	0.0082	0.0098	0.0332	0.0851
	0.05	0.0055	0.0063	0.0083	0.0098	0.0056	0.0823
	0.08	0.0088	0.0092	0.0099	0.0117	0.0332	0.0816
	0.1	0.0113	0.0118	0.0126	0.0137	0.0351	0.0866
	0.3	0.0351	0.0339	0.0337	0.0338	0.0441	0.0855
	0.6	0.0668	0.0790	0.0685	0.0818	0.0784	0.1041
20	**0.03**	0.0019	0.0026	0.0042	0.0053	0.0181	0.0468
	0.05	0.0028	0.0031	0.0044	0.0053	0.0028	0.0438
	0.08	0.0044	0.0044	0.0050	0.0057	0.0186	0.0461
	0.1	0.0054	0.0054	0.0060	0.0066	0.0188	0.0445
	0.3	0.0187	0.0185	0.0185	0.0183	0.0239	0.0436
	0.6	0.0410	0.0408	0.0427	0.0407	0.0445	0.0495
30	**0.03**	0.0015	0.0022	0.0037	0.0047	0.0180	0.0501
	0.05	0.0022	0.0026	0.0037	0.0047	0.0189	0.0474
	0.08	0.0036	0.0038	0.0045	0.0053	0.0194	0.0534
	0.1	0.0045	0.0047	0.0053	0.0058	0.0183	0.0511
	0.3	0.0167	0.0174	0.0181	0.0181	0.0274	0.0600
	0.6	0.0430	0.0432	0.0446	0.0455	0.0546	0.0807
40	**0.03**	0.0011	0.0016	0.0027	0.0034	0.0130	0.0363
	0.05	0.0016	0.0019	0.0028	0.0035	0.0134	0.0368
	0.08	0.0027	0.0029	0.0033	0.0038	0.0134	0.0374
	0.1	0.0034	0.0035	0.0039	0.0042	0.0136	0.0382
	0.3	0.0124	0.0128	0.0133	0.0136	0.0212	0.0427
	0.6	0.0340	0.0344	0.0355	0.0357	0.0450	0.0594
50	**0.03**	0.0009	0.0013	0.0021	0.0027	0.0106	0.0297
	0.05	0.0014	0.0016	0.0022	0.0028	0.0110	0.0302
	0.08	0.0022	0.0023	0.0027	0.0032	0.0114	0.0312
	0.1	0.0029	0.0030	0.0031	0.0036	0.0116	0.0314
	0.3	0.0108	0.0111	0.0114	0.0114	0.0165	0.0350
	0.6	0.0274	0.0273	0.0284	0.0298	0.0335	0.0439
60	**0.03**	0.0007	0.0011	0.0018	0.0023	0.0093	0.0244
	0.05	0.0011	0.0013	0.0018	0.0023	0.0095	0.0248
	0.08	0.0018	0.0019	0.0022	0.0025	0.0096	0.0255
	0.1	0.0023	0.0024	0.0026	0.0028	0.0097	0.0262
	0.3	0.0090	0.0091	0.0094	0.0095	0.0138	0.0289
	0.6	0.0230	0.0234	0.0246	0.0252	0.0294	0.0350
70	**0.03**	0.0006	0.0009	0.0016	0.0020	0.0082	0.0207
	0.05	0.0010	0.0011	0.0016	0.0021	0.0083	0.0211
	0.08	0.0016	0.0017	0.0019	0.0022	0.0085	0.0217
	0.1	0.0020	0.0021	0.0022	0.0025	0.0085	0.0219
	0.3	0.0076	0.0079	0.0081	0.0083	0.0120	0.0242
	0.6	0.0201	0.0209	0.0215	0.0219	0.0236	0.0315

The 10% upper percentiles of the *DMCEs* statistic at a = -1, continued.

n	σ_1	σ_2 0.03	0.05	0.08	0.1	0.3	0.6
80	**0.03**	0.0005	0.0008	0.0014	0.0018	0.0077	0.0179
	0.05	0.00084	0.0010	0.0014	0.0018	0.0078	0.0182
	0.08	0.0014	0.0014	0.0017	0.0019	0.0077	0.0187
	0.1	0.0018	0.0018	0.0019	0.0022	0.0077	0.0190
	0.3	0.0067	0.0069	0.0072	0.0073	0.0108	0.0217
	0.6	0.0175	0.0181	0.0190	0.0196	0.0215	0.0255
90	**0.03**	0.0005	0.0008	0.0013	0.0017	0.0065	0.0165
	0.05	0.0008	0.0009	0.0013	0.0017	0.0066	0.0165
	0.08	0.0013	0.0013	0.0015	0.0018	0.0068	0.0172
	0.1	0.0016	0.0016	0.0018	0.0019	0.0070	0.0174
	0.3	0.0060	0.0061	0.0063	0.0065	0.0094	0.0190
	0.6	0.0158	0.0166	0.0177	0.0174	0.0195	0.0238
100	**0.03**	0.0004	0.0007	0.0012	0.0015	0.0062	0.0148
	0.05	0.0007	0.0008	0.0012	0.0015	0.0064	0.0150
	0.08	0.0011	0.0012	0.0014	0.0017	0.0063	0.0154
	0.1	0.0015	0.0015	0.0016	0.0018	0.0066	0.0157
	0.3	0.0054	0.0056	0.0058	0.0059	0.0105	0.0169
	0.6	0.0150	0.0153	0.0158	0.0160	0.0179	0.0214
110	**0.03**	0.0004	0.0006	0.0010	0.0013	0.0055	0.0133
	0.05	0.0007	0.0007	0.0011	0.0014	0.0056	0.0138
	0.08	0.0011	0.0011	0.0012	0.0015	0.0056	0.0140
	0.1	0.0014	0.0014	0.0015	0.0016	0.0058	0.0143
	0.3	0.0049	0.0049	0.0051	0.0059	0.0081	0.0154
	0.6	0.0125	0.0132	0.0138	0.0138	0.0154	0.0184
130	**0.03**	0.0004	0.0006	0.0009	0.0012	0.0048	0.0115
	0.05	0.0006	0.0006	0.0009	0.0012	0.0049	0.0116
	0.08	0.0009	0.0009	0.0011	0.0013	0.0050	0.0119
	0.1	0.0012	0.0012	0.0013	0.0014	0.0050	0.0120
	0.3	0.0043	0.0045	0.0046	0.0046	0.0059	0.0143
	0.6	0.0112	0.0115	0.0121	0.0124	0.0132	0.0154
150	**0.03**	0.0003	0.0005	0.0008	0.0011	0.0043	0.0099
	0.05	0.0005	0.0005	0.0008	0.0011	0.0045	0.0101
	0.08	0.0008	0.0008	0.0009	0.0011	0.0045	0.0102
	0.1	0.0010	0.0011	0.0011	0.0012	0.0048	0.0104
	0.3	0.0039	0.0040	0.0041	0.0040	0.0058	0.0109
	0.6	0.0097	0.0100	0.0102	0.0104	0.0109	0.0124

APPENDIX 4
Splus Subroutine for Introduce Outlier in the Data set

```
## Introduce outlier ##
## 10%, 20%, 30% , 40%, 50% contaminated ##
## set.seed(00)

Raosimu<-function(n,sig1,sig2,sig1s,sig2s,per,a,simu) {

       a01=matrix(0,nrow=simu)
       a11=matrix(0,nrow=simu)
       b11=matrix(0,nrow=simu)
       c01=matrix(0,nrow=simu)
       c11=matrix(0,nrow=simu)
       d11=matrix(0,nrow=simu)
       rho1=matrix(0,nrow=simu)
       s1=matrix(0,nrow=simu)
       s2=matrix(0,nrow=simu)

       a02=matrix(0,nrow=simu)
       a12=matrix(0,nrow=simu)
       b12=matrix(0,nrow=simu)
       c02=matrix(0,nrow=simu)
       c12=matrix(0,nrow=simu)
       d12=matrix(0,nrow=simu)
       rho2=matrix(0,nrow=simu)
       s1o=matrix(0,nrow=simu)
       s2o=matrix(0,nrow=simu)

       for (i in 1:simu){

              AD=Rao(n,sig1,sig2,sig1s,sig2s,per,a)
              a0true=AD$a0t
              a1true=AD$a1t
              b1true=AD$b1t
              c0true=AD$c0t
              c1true=AD$c1t
              d1true=AD$d1t
              s1true=AD$sig1
              s2true=AD$sig2
              s1strue=AD$sig1s
              s2strue=AD$sig2s
              data=AD$data

              AE=cov(data,per)
              a01[i]=AD$A01
              a11[i]=AD$A11
              b11[i]=AD$B11
              c01[i]=AD$C01
              c11[i]=AD$C11
              d11[i]=AD$D11
              rho1[i]=AD$rhoEst1
              s1[i]=AD$s11
              s2[i]=AD$s21

              a02[i]=AD$A02
              a12[i]=AD$A12
              b12[i]=AD$B12
              c02[i]=AD$C02
              c12[i]=AD$C12
              d12[i]=AD$D12
              rho2[i]=AD$rhoEst2
              s1o[i]=AD$s12
              s2o[i]=AD$s22
       }

              a01mean=mean(a01)
```

```
a11mean=mean(a11)
b11mean=mean(b11)
c01mean=mean(c01)
c11mean=mean(c11)
d11mean=mean(d11)
rhomean1=mean(rho1)
s1mean=mean(s1)
s2mean=mean(s2)

a01bias=a01mean-a0true
a11bias=a11mean-a1true
b11bias=b11mean-b1true
c01bias=c01mean-c0true
c11bias=c11mean-c1true
d11bias=d11mean-d1true
s1bias=s1mean-s1true
s2bias=s2mean-s2true

a01RMSE=sqrt(sum(a01-a0true)^2/simu)
a11RMSE=sqrt(sum(a11-a1true)^2/simu)
b11RMSE=sqrt(sum(b11-b1true)^2/simu)
c01RMSE=sqrt(sum(c01-c0true)^2/simu)
c11RMSE=sqrt(sum(c11-c1true)^2/simu)
d11RMSE=sqrt(sum(d11-d1true)^2/simu)

a01SE=stdev(a01)
a11SE=stdev(a11)
b11SE=stdev(b11)
c01SE=stdev(c01)
c11SE=stdev(c11)
d11SE=stdev(d11)
rho1SE=stdev(rho1)
s1SE=stdev(s1)
s2SE=stdev(s2)

a02mean=mean(a02)
a12mean=mean(a12)
b12mean=mean(b12)
c02mean=mean(c02)
c12mean=mean(c12)
d12mean=mean(d12)
rhomean2=mean(rho2)
s1omean=mean(s1o)
s2omean=mean(s2o)

a02bias=a02mean-a0true
a12bias=a12mean-a1true
b12bias=b12mean-b1true
c02bias=c02mean-c0true
c12bias=c12mean-c1true
d12bias=d12mean-d1true
s1obias=s1omean-s1strue
s2obias=s2omean-s2strue

a02RMSE=sqrt(sum(a02-a0true)^2/simu)
a12RMSE=sqrt(sum(a12-a1true)^2/simu)
b12RMSE=sqrt(sum(b12-b1true)^2/simu)
c02RMSE=sqrt(sum(c02-c0true)^2/simu)
c12RMSE=sqrt(sum(c12-c1true)^2/simu)
d12RMSE=sqrt(sum(d12-d1true)^2/simu)


a02SE=stdev(a02)
a12SE=stdev(a12)
b12SE=stdev(b12)
c02SE=stdev(c02)
c12SE=stdev(c12)
d12SE=stdev(d12)
rho2SE=stdev(rho2)
s1oSE=stdev(s1o)
s2oSE=stdev(s2o)
```

```
        mean1=c(a01=a01mean,a11=a11mean,b11=b11mean,c01=c01mean,c11=c11mean,d11=
        d11mean,rho1=rhomean1)
        bias1=c(a01=a01bias,a11=a11bias,b11=b11bias,c01=c01bias,c11=c11bias,d11=
        d11bias)
        RMSE1=c(a01=a01RMSE,a11=a11RMSE,b11=b11RMSE,c01=c01RMSE,c11=c11RMSE,d11=
        d11RMSE)
        SE1=c(a01=a01SE,a11=a11SE,b11=b11SE,c01=c01SE,c11=c11SE,d11=d11SE,rho1=r
        ho1SE)

        mean2=c(a02=a02mean,a12=a12mean,b12=b12mean,c02=c02mean,c1
        2=c12mean,d12=d12mean,rho2=rhomean2)
        bias2=c(a02=a02bias,a12=a12bias,b12=b12bias,c02=c02bias,c12=c12bias,d12=
        d12bias)
        RMSE2=c(a02=a02RMSE,a12=a12RMSE,b12=b12RMSE,c02=c02RMSE,c12=c12RMSE,d12=
        d12RMSE)
        SE2=c(a02=a02SE,a12=a12SE,b12=b12SE,c02=c02SE,c12=c12SE,d12=d12SE,rho2=r
        ho2SE)

        NoOutlier1=cbind(mean1,SE1)
        NoOutlier2=cbind(RMSE1,bias1)
        ContainOutlier1=cbind(mean2,SE2)
        ContainOutlier2=cbind(RMSE2,bias2)

        list(NoOutlier1=NoOutlier1,NoOutlier2=NoOutlier2,ContainOutlier1=Contain
        Outlier1,ContainOutlier2=ContainOutlier2)
        }

##-----------calculate the covariance----------##


cov=function(data,per){

                n=nrow(data)
                u=data[,1]
                v1=data[,2]
                v3=data[,3]

        calculate the covariance matrix

                order=1.
                order.matrix=t(matrix(rep(c(1.:order),n),ncol=n))

        #without outlier
        #---------------

                cos.u=cos(u*order.matrix)
                sin.u=sin(u*order.matrix)
                V11=cos(v1)
                V21=sin(v1)
                ones<-matrix(1.,n,1.)
                U1<-cbind(ones,cos.u,sin.u)
                M<-U1%*%ginverse(t(U1)%*%U1)%*%t(U1)
                I<-diag(n)

                R01<-matrix(0,nrow=2,ncol=2)
                R01[1,1]<-t(V11)%*%(I-M)%*%V11
                R01[2,2]<-t(V21)%*%(I-M)%*%V21
                R01[1,2]<-t(V11)%*%(I-M)%*%V21
                R01[2,1]<-t(V21)%*%(I-M)%*%V11

                covraoNoOutlier=(1/(n-2*(order+1)))*R01

                s11=sqrt(covraoNoOutlier[1,1])
                s11=as.vector(s11)
                s21=sqrt(covraoNoOutlier[2,2])
                s21=as.vector(s21)

        #with outlier
        #-------------
```

```
		cos.u=cos(u*order.matrix)
		sin.u=sin(u*order.matrix)
		V12=cos(v3)
		V22=sin(v3)
		ones<-matrix(1.,n,1.)
		U2<-cbind(ones,cos.u,sin.u)
		M<-U2%*%ginverse(t(U2)%*%U2)%*%t(U2)
		I<-diag(n)

		R02<-matrix(0,nrow=2,ncol=2)
		R02[1,1]<-t(V12)%*%(I-M)%*%V12
		R02[2,2]<-t(V22)%*%(I-M)%*%V22
		R02[1,2]<-t(V12)%*%(I-M)%*%V22
		R02[2,1]<-t(V22)%*%(I-M)%*%V12

		covraoOutlier=(1/(n-2*(order+1)))*R02

		s12=sqrt(covraoOutlier[1,1])
		s12=as.vector(s12)
		s22=sqrt(covraoOutlier[2,2])
		s22=as.vector(s22)

	list(covraoNoOutlier=covraoNoOutlier,covraoOutlier=covraoOutlier,s11=s11
	,s21=s21,s12=s12,s22=s22)
}

##----------------- generate the data set --------------------##

Rao<-function(n,sig1,sig2,sig1s,sig2s,per,a){

#uncontaminated
#--------------

# step 1:	generate u variable

		u<-rvm(n,pi,2)

# step 2: generate e1 & e2 variable

		e1=rnorm(n,0,sig1)
		e2=rnorm(n,0,sig2)
#step 3: calculate v

		cv1=cos(a+u)
		sv1=sin(a+u)
		ccv1=cv1 + e1
		ccv1=as.matrix(ccv1)
		ssv1=sv1 + e2
		ssv1=as.matrix(ssv1)

		for(i in 1:n){

		if (ccv1[i,] > 1 | ccv1[i,] < -1){ccv1[i,]=NA}
			if (ssv1[i,] >1 | ssv1[i,] < -1){ssv1[i,]=NA}
		}

		tt=matrix(0,nrow=n,ncol=6)

		tt[,1]=ccv1
		tt[,2]=ssv1
		tt[,3]=atan(ssv1,ccv1)
		tt[,4]=u
		tt[,5]=e1
		tt[,6]=e2

		tt=na.exclude(tt)

		u=tt[,4]
		v1=tt[,3]
		v1 = v1 %% (2. *pi)
		e1=tt[,5]
```

```
              e2=tt[,6]

# step 5: get the parameter estimation

              u=as.vector(u)
              v1=as.vector(v1)
              CirReg1<-circ.reg(u,v1)    # without outlier #
              A=CirReg1$coef
              A01=A[1,1]
              A01=as.vector(A01)
              A11=A[2,1]
              A11=as.vector(A11)
              B11=A[3,1]
              B11=as.vector(B11)
              C01=A[1,2]
              C01=as.vector(C01)
              C11=A[2,2]
              C11=as.vector(C11)
              D11=A[3,2]
              D11=as.vector(D11)

              rhoEst1=CirReg1$rho

#contaminated
#------------

              h=length(v1)

# step 6: calculate the % error

              m=(per/100)*h
              m=ceiling(m)
              t=u[1:m]

# step 7: generate e1 & e2 variable

              e1s=rnorm(m,0,sig1s)
              e2s=rnorm(m,0,sig2s)

#step 8: calculate v new

              cv2=cos(a+t)
              sv2=sin(a+t)
              ccv2=cv2 + e1s
              ccv2=as.matrix(ccv2)
              ssv2=sv2 + e2s
              ssv2=as.matrix(ssv2)

              for(i in 1:m){

              if (ccv2[i,] > 1 | ccv2[i,] < -1){ccv2[i,]=NA}
                     if (ssv2[i,] >1 | ssv2[i,] < -1){ssv2[i,]=NA}

              }

              ss=matrix(0,nrow=m,ncol=5)

              ss[,1]=ccv2
              ss[,2]=ssv2
              ss[,3]=atan(ssv2,ccv2)
              ss[,4]=e1s
              ss[,5]=e2s

              ss=na.exclude(ss)

              newe3=ss[,4]
              newe4=ss[,5]
              v2=ss[,3]
              v2 = v2 %% (2. *pi)
              r=nrow(ss)
              v3=c(v1[1:(h-r)],v2)
```

```
# step 9: get the parameter estimation

        u=as.vector(u)
        v3=as.vector(v3)

        CirReg2<-circ.reg(u,v3)

        B=CirReg2$coef

        A02=B[1,1]
        A02=as.vector(A02)
        A12=B[2,1]
        A12=as.vector(A12)
        B12=B[3,1]
        B12=as.vector(B12)
        C02=B[1,2]
        C02=as.vector(C02)
        C12=B[2,2]
        C12=as.vector(C12)
        D12=B[3,2]
        D12=as.vector(D12)
        rhoEst2=CirReg2$rho

#step 10: true value of parameter estimation

        a0t=0
        a1t=cos(a)
        b1t=-sin(a)
        c0t=0
        c1t=sin(a)
        d1t=cos(a)
        sig1=sig1
        sig2=sig2
        sig1s=sig1s
        sig2s=sig2s

        data=cbind(u,v1,v3)

    list(data=data,a0t=a0t,a1t=a1t,b1t=b1t,c0t=c0t,c1t=c1t,d1t=d1t,A01=A01,A
    11=A11,B11=B11,C01=C01,C11=C11,D11=D11,rhoEst1=rhoEst1,A02=A02,A12=A12,B
    12=B12,C02=C02,C12=C12,D12=D12,rhoEst2=rhoEst2,sig1=sig1,sig2=sig2,sig1s
    =sig1s,sig2s=sig2s)

}
```

APPENDIX 5

Splus Subroutine for *COVRATIO* Statistic of real data set: wind direction data

```
simu=function(n,sig1,sig2,simu){
    #simu(n=130,sig1=0.3,sig2=0.3,simu=500)

    pe=matrix(0,nrow=simu,ncol=9)

    for (i in 1:simu){

        f=DataIN(n,sig1,sig2)
        aa=covi(n,sig1,sig2)

        pe[i,1]=aa$a0E              #simu estimate parameter
        pe[i,2]=aa$a1E
        pe[i,3]=aa$b1E
        pe[i,4]=aa$c0E
        pe[i,5]=aa$c1E
        pe[i,6]=aa$d1E
        pe[i,7]=aa$kE
        pe[i,8]=aa$rhoE
        pe[i,9]=aa$maxP

    }

        a0e=pe[,1]
        a1e=pe[,2]
        b1e=pe[,3]
        c0e=pe[,4]
        c1e=pe[,5]
        d1e=pe[,6]
        kappae=pe[,7]
        rhoe=pe[,8]
        MaxP=pe[,9]

        pe=cbind(a0e,a1e,b1e,c0e,c1e,d1e,kappae,rhoe)


        a0mean=mean(a0e)
        a1mean=mean(a1e)
        b1mean=mean(b1e)
        c0mean=mean(c0e)
        c1mean=mean(c1e)
        d1mean=mean(d1e)
        kappamean=mean(kappae)
        rhomean=mean(rhoe)

        a0bias=a0mean-a0tr
        a1bias=a1mean-a1tr
        b1bias=b1mean-b1tr
        c0bias=c0mean-c0tr
        c1bias=c1mean-c1tr
        d1bias=d1mean-d1tr

        a0RMSE=sqrt(sum(a0e-a0tr)^2/simu)
        a1RMSE=sqrt(sum(a1e-a1tr)^2/simu)
        b1RMSE=sqrt(sum(b1e-b1tr)^2/simu)
        c0RMSE=sqrt(sum(c0e-c0tr)^2/simu)
        c1RMSE=sqrt(sum(c1e-c1tr)^2/simu)
        d1RMSE=sqrt(sum(d1e-d1tr)^2/simu)

        a0SE= stdev (a0e)
        a1SE= stdev (a1e)
        b1SE= stdev (b1e)
        c0SE= stdev (c0e)
        c1SE= stdev (c1e)
        d1SE= stdev (d1e)
        kappaSE= stdev (kappae)
        rhoSE= stdev (rhoe)
```

```
mean=c(a0e=a0mean,a1e=a1mean,b1e=b1mean,c0e=c0mean,c1e=c1mean,d1e=d1mean,k
   appae=kappamean,rhoe=rhomean)
bias=c(a0e=a0bias,a1e=a1bias,b1e=b1bias,c0e=c0bias,c1e=c1bias,d1e=d1bias)
RMSE=c(a0e=a0RMSE,a1e=a1RMSE,b1e=b1RMSE,c0e=c0RMSE,c1e=c1RMSE,d1e=d1RMSE)
SE=c(a0e=a0SE,a1e=a1SE,b1e=b1SE,c0e=c0SE,c1e=c1SE,d1e=d1SE,kappae=kappaSE,
rhoe=rhoSE)

    result1=cbind(mean,SE)
    result2=cbind(bias,RMSE)

    mp=sort(MaxP)

    l1=99/100*simu
    l1=ceiling(l1)
    cp01=mp[l1]

    l2=95/100*simu
    l2=ceiling(l2)
    cp05=mp[l2]

    l3=90/100*simu
    l3=ceiling(l3)
    cp10=mp[l3]

    pe=cbind(a0e,a1e,b1e,c0e,c1e,d1e,kappae,rhoe,MaxP)

list(pt=pt,result1=result1,result2=result2,cp01=cp01,cp05=cp05,cp10=cp10,M
axP=MaxP)

}

#---------------------data & cov------------------#

covi=function(n,sig1,sig2){

    dd=DataIN(n,sig1,sig2)
    data=dd$datain
    n=nrow(data)

    a0E=dd$a01
    a1E=dd$a11
    b1E=dd$b11
    c0E=dd$c01
    c1E=dd$c11
    d1E=dd$d11
    kE=dd$kappaEst1
    rhoE=dd$rhoEst1
    ehat1E=dd$ehat1
    ehat2E=dd$ehat2

    C=cov(data,sig1,sig2)
    COVFullData=C$covrao
    DeterFullData=C$Deter


            #-------deleted i row data-----------#

    DeterE=matrix(0,nrow=n)

    for(i in 1:n){

        Newdata=remove.row(data,r=[i],1)
        DeterE[i]=cov(Newdata,sig1,sig2)$Deter

    }

    COVRATIO=matrix(0,nrow=n)

    for (i in 1:n){

        COVRATIO[i]=DeterE[i]/DeterFullData
```

```
        }

        p=abs(COVRATIO-1)
        maxP=max(p)

    list(a0E=a0E,a1E=a1E,b1E=b1E,c0E=c0E,c1E=c1E,d1E=d1E,kE=kE,rhoE=rhoE,ehat1
    E=ehat1E,ehat2E=ehat2E,COVFullData=COVFullData,DeterFullData=DeterFullData
    ,COVRATIO=COVRATIO,p=p,maxP=maxP)

    }


    #-------cov method----------#

        cov=function(data,sig1,sig2){

        n=nrow(data)
        u=data[,1]
        v=data[,2]

        order=1.
        order.matrix=t(matrix(rep(c(1.:order),n),ncol=n))
        cos.u=cos(u*order.matrix)
        sin.u=sin(u*order.matrix)
        V1=cos(v)
        V2=sin(v)
        ones<-matrix(1.,n,1.)
        U<-cbind(ones,cos.u,sin.u)
        M<-U%*%ginverse(t(U)%*%U)%*%t(U)

        I<-diag(n)

        R0<-matrix(0,nrow=2,ncol=2)
        R0[1,1]<-t(V1)%*%(I-M)%*%V1
        R0[2,2]<-t(V2)%*%(I-M)%*%V2
        R0[1,2]<-t(V1)%*%(I-M)%*%V2
        R0[2,1]<-t(V2)%*%(I-M)%*%V1

        covrao=(1/(n-2*(2*order+1)))*R0
        Deter=det(covrao)

        list(covrao=covrao,Deter=Deter)
    }

#--------- Generate Von mises data set-----------------#

    DataIN<-function(n,sig1,sig2){
        #DataIN(n=130,sig1=0.3,sig2=0.3)

# step 1:

        u<-rvm(n,pi,2)

# step 2:

        e1=rnorm(n,0,sig1)
        e2=rnorm(n,0,sig2)

#step 3: true value of parameter estimation

        a0 = 0.0674
        a1 = 0.7559
        b1 = -0.0948
        c0 = -0.047
        c1 = 0.1049
        d1 = 0.9762

#step 4: calculate v

        ccv=a0+ a1*cos(u) + b1*sin(u) + e1
        ccv=as.matrix(ccv)
```

```
    ssv=c0+ c1*cos(u) + d1*sin(u) + e2
    ssv=as.matrix(ssv)

    for(i in 1:n){

        if (ccv[i,] > 1 | ccv[i,] < -1)          {ccv[i,]=NA}
            if (ssv[i,] >1 | ssv[i,] < -1)       {ssv[i,]=NA}

    }

    tt=matrix(0,nrow=n,ncol=6)

    tt[,1]=ccv
    tt[,2]=ssv
    tt[,3]=atan(ssv,ccv)
    tt[,4]=u
    tt[,5]=e1
    tt[,6]=e2

    tt=na.exclude(tt)

    u=tt[,4]
    v=tt[,3]
    v = v %% (2. *pi)
    e1=tt[,5]
    e2=tt[,6]

    datain=cbind(u,v,e1,e2)

    u=as.vector(u)
    v=as.vector(v)

    CirReg<-circ.reg(u,v)       # get the parameter estimation

    A=CirReg$coef

    a01=A[1,1]
    a01=as.vector(a01)
    a11=A[2,1]
    a11=as.vector(a11)
    b11=A[3,1]
    b11=as.vector(b11)
    c01=A[1,2]
    c01=as.vector(c01)
    c11=A[2,2]
    c11=as.vector(c11)
    d11=A[3,2]
    d11=as.vector(d11)
    kappaEst1=CirReg$kappa
    rhoEst1=CirReg$rho
    vhat=CirReg$fitted
    resid=CirReg$residuals
    ehat1=tt[,1]-cos(vhat)
    ehat2=tt[,2]-sin(vhat)

list(datain=datain,a01=a01,a11=a11,b11=b11,c01=c01,c11=c11,d11=d11,kappaEs
t1=kappaEst1,rhoEst1=rhoEst1,ehat1=ehat1,ehat2=ehat2)

}
```

APPENDIX 6

Splus Subroutine for *DMCE* Statistic of real data set: Eye data

```
DMCE=function(n=23,sig1=0.16,sig2=0.16,simu=500){
             #DMCE(n=23,sig1=0.16,sig2=0.16,simu=500)

             pe=matrix(0,nrow=simu,ncol=9)

             for (i in 1:simu){

                    f=DataIN(n,sig1,sig2)
                    aa=com(n,sig1,sig2)

                    pe[i,1]=aa$a0E             #simu estimate parameter
                    pe[i,2]=aa$a1E
                    pe[i,3]=aa$b1E
                    pe[i,4]=aa$c0E
                    pe[i,5]=aa$c1E
                    pe[i,6]=aa$d1E
                    pe[i,7]=aa$rhoE
                    pe[i,8]=aa$dmcec
                    pe[i,9]=aa$dmces

                    }

                    a0e=pe[,1]
                    a1e=pe[,2]
                    b1e=pe[,3]
                    c0e=pe[,4]
                    c1e=pe[,5]
                    d1e=pe[,6]
                    rhoe=pe[,7]
                    DMCEc=pe[,8]
                    DMCEs=pe[,9]

       pe=cbind(a0e,a1e,b1e,c0e,c1e,d1e,rhoe,DMCEc,DMCEs)

                    mp1=sort(DMCEc)
                    mp2=sort(DMCEs)

                    l1=99/100*simu
                    l1=ceiling(l1)
                    cp01c=mp1[l1]
                    cp01s=mp2[l1]

                    l2=95/100*simu
                    l2=ceiling(l2)
                    cp05c=mp1[l2]
                    cp05s=mp2[l2]

                    l3=90/100*simu
                    l3=ceiling(l3)
                    cp10c=mp1[l3]
                    cp10s=mp2[l3]

       list(mp1=mp1,mp2=mp2,DMCEc=DMCEc,DMCEs=DMCEs,cp01c=cp01c,cp05c=cp05c,cp1
0c=cp10c,cp01s=cp01s,cp05s=cp05s,cp10s=cp10s)
}

##----------row deletion approach--------##

       com=function(n,sig1,sig2){

       pp=DataIN(n,sig1,sig2)
       data=pp$datain
       n=nrow(data)
       ff=mce(data)

       a0E=ff$a01
       a1E=ff$a11
```

```
       b1E=ff$b11
       c0E=ff$c01
       c1E=ff$c11
       d1E=ff$d11
       rhoE=ff$rhoEst1

       MCEcFullData=ff$MCEc
       MCEsFullData=ff$MCEs

              #-------deleted i row data-----------#

              MCEci=matrix(0,nrow=n)
              MCEsi=matrix(0,nrow=n)

              for(i in 1:n){

                            Newdata=remove.row(data,r=[i],1)
                            jj=mce(Newdata)

                            MCEci[i]=jj$MCEc
                            MCEsi[i]=jj$MCEs

                            }

              dmcec = max(abs(MCEcFullData-MCEci))
              dmces = max(abs(MCEsFullData-MCEsi))

              pc=abs(MCEcFullData-MCEci)
              #plot(abs(MCEcFullData-MCEci))
              #plot(abs(MCEsFullData-MCEsi))
       list(a0E=a0E,a1E=a1E,b1E=b1E,c0E=c0E,c1E=c1E,d1E=d1E,rhoE=rhoE,dmcec=dmc
ec,dmces=dmces,pc=pc)

       }

##------------mce method------------##

       mce=function(data){

              u=data[,1]
              v=data[,2]
              n=nrow(data)

# get the parameter estimation

              CirReg<-circ.reg(u,v)

              A=CirReg$coef

              a01=A[1,1]
              a01=as.vector(a01)
              a11=A[2,1]
              a11=as.vector(a11)
              b11=A[3,1]
              b11=as.vector(b11)
              c01=A[1,2]
              c01=as.vector(c01)
              c11=A[2,2]
              c11=as.vector(c11)
              d11=A[3,2]
              d11=as.vector(d11)
              rhoEst1=CirReg$rho
              vhat=CirReg$fitted
              resid=CirReg$residuals

# var-cov matrix

              order=1.
              order.matrix=t(matrix(rep(c(1.:order),n),ncol=n))
              cos.u=cos(u*order.matrix)
              sin.u=sin(u*order.matrix)
```

```
            V1=cos(v)
            V2=sin(v)
            ones<-matrix(1.,n,1.)
            U<-cbind(ones,cos.u,sin.u)
            M<-U%*%ginverse(t(U)%*%U)%*%t(U)
            I<-diag(n)
            R0<-matrix(0,nrow=2,ncol=2)

            R0[1,1]<-t(V1)%*%(I-M)%*%V1
            R0[2,2]<-t(V2)%*%(I-M)%*%V2
            R0[1,2]<-t(V1)%*%(I-M)%*%V2
            R0[2,1]<-t(V2)%*%(I-M)%*%V1

            covrao=(1/(n-2*(order+1)))*R0

# calculate circular distance

            d=pi-abs(pi-abs(v-vhat))

# Mean circular error statistics

            MCEc=1-(1/n)*(sum(cos(v-vhat)))  #MCEc full data set
            MCEs=(1/n)*(sum(sin(d/2)))             #MCEs full data set

      list(a01=a01,a11=a11,b11=b11,c01=c01,c11=c11,d11=d11,rhoEst1=rhoEst1,vha
t=vhat,covrao=covrao,MCEc=MCEc,MCEs=MCEs)

      }

##---------generate von Mises data set----------##

      DataIN<-function(n,sig1,sig2){

# step 1:

            u<-rvm(5*n,pi,2)

# step 2:

            e1=rnorm(5*n,0,sig1)
            e2=rnorm(5*n,0,sig2)

#step 3: true value of parameter estimation

#step 4: calculate v

            a0t=1.0822
            a1t=-0.1497
            b1t=-0.3837
            c0t=0.0986
            c1t=0.2534
            d1t=0.5935

            ccv=a0t+ a1t*cos(u) + b1t*sin(u)+e1
            ccv=as.matrix(ccv)

            ssv=c0t+ c1t*cos(u) + d1t*sin(u)+e2
            ssv=as.matrix(ssv)


            for(i in 1:(5*n)){

                  if (ccv[i,] > 1 | ccv[i,] < -1)  {ccv[i,]=NA}
                        if (ssv[i,] >1 | ssv[i,] < -1){ssv[i,]=NA}

            }

            tt=matrix(0,nrow=5*n,ncol=6)

            tt[,1]=ccv
            tt[,2]=ssv
```

```
        tt[,3]=atan(ssv,ccv)
        tt[,4]=u
        tt[,5]=e1
        tt[,6]=e2

        tt=na.exclude(tt)

        jk=length(tt)

        if (jk>=n){ppp=n}
        else {ppp=jk}

        u=tt[(1:ppp),4]
        v=tt[(1:ppp),3]
        v = v %% (2. *pi)
        e1=tt[(1:ppp),5]
        e2=tt[(1:ppp),6]

        u=as.vector(u)
        v=as.vector(v)

        datain=cbind(u,v,e1,e2)

list(datain=datain,a0t=a0t,a1t=a1t,b1t=b1t,c0t=c0t,c1t=c1t,d1t=d1t)

}
```

APPENDIX 7

Splus Subroutine for Bootstrapping on JS Circular Functional Relationship Model

```
JSCfunc=function(DATA,B){

	u=DATA[,1]
	v=DATA[,2]
	n=nrow(DATA)

	aa=complexEst(DATA)

	aa1=aa$delRh
	aa2=aa$delIh
	aa3=aa$epsRh
	aa4=aa$epsIh

	err=cbind(aa1,aa2,aa3,aa4)

	for(i in 1:B){

		index=c(1:n)
		Bindex=sample(index,n,replace=T)
		ji1=cbind(err[Bindex,1],err[Bindex,2],err[Bindex,3],err[Bindex,4]
		)[,1]
		ji2=cbind(err[Bindex,1],err[Bindex,2],err[Bindex,3],err[Bindex,4]
		)[,2]
		ji3=cbind(err[Bindex,1],err[Bindex,2],err[Bindex,3],err[Bindex,4]
		)[,3]
		ji4=cbind(err[Bindex,1],err[Bindex,2],err[Bindex,3],err[Bindex,4]
		)[,4]

		Berr=cbind(ji1,ji2,ji3,ji4)

		cU=ji1-cos(u)
		sU=ji2-sin(u)
		cV=ji3-cos(v)
		sV=ji4-sin(v)

		Ub=atan(sU,cU)
		Vb=atan(sV,cV)

		newUV=cbind(u,v,Ub,Vb)

		complexEstN<-function(newUV){

			u=newUV[,1]
			v=newUV[,2]
			U=newUV[,3]

			n<-length(newUV)

			# get initial value #
			#-------------------#

			Estlog=function(newUV){

			loglike1=function(AA0,AA1,BB1,w1,newUV){

			u=newUV[,1]
			v=newUV[,2]
			U=newUV[,3]

			n=length(u)
			omega1=1

			L1=2/omega1*w1*(sum(cos(v)-AA0-AA1*cos(U)-BB1*sin(U)))

			list(L1=L1)
			}
```

```
loglike2=function(CC0,CC1,DD1,w2,newUV){

u=newUV[,1]
v=newUV[,2]
U=newUV[,3]

n=length(u)
omega1=1

L2=2/omega1*w2*(sum(sin(v)-CC0-CC1*cos(U)-DD1*sin(U)))

list(L2=L2)
}

AA0=seq(-1,1,0.3)
AA1=seq(-1,1,0.3)
BB1=seq(-1,1,0.3)
CC0=seq(-1,1,0.3)
CC1=seq(-1,1,0.3)
DD1=seq(-1,1,0.3)
w1=seq(0,1,0.01)
w2=seq(0,1,0.01)

LL1=array(0,dim=c(length(AA0),length(AA1),length(BB1),length(w1)))

for(i in 1:length(AA0)){
    for(j in 1:length(AA1)){
        for(k in 1:length(BB1)){
            for(l in 1:length(w1)){

LL1[i,j,k,l]=loglike1(AA0[i],AA1[j],BB1[k],w1[l],newUV)$L1
}
                                    }
                                }
                            }

LL2=array(0,dim=c(length(CC0),length(CC1),length(DD1),length(w2)))

for(m in 1:length(CC0)){
    for(n in 1:length(CC1)){
        for(o in 1:length(DD1)){
            for(p in 1:length(w2)){

LL2[m,n,o,p]=loglike2(CC0[m],CC1[n],DD1[o],w2[p],newUV)$L2

}
                                    }
                                }
                            }
#plot(LL1)
#plot(LL2)

MaxLL1=max(LL1)

for(i in 1:length(AA0)){
    for(j in 1:length(AA1)){
        for(k in 1:length(BB1)){
            for(l in 1:length(w1)){

            if(LL1[i,j,k,l]==MaxLL1)        {
                a00=AA0[i]
                a10=AA1[j]
                b10=BB1[k]
                w10=w1[l]

}
                                    }
                                }
                            }
                        }
```

```
            mm1=cbind(MaxLL1,a00,a10,b10,w10)

    MaxLL2=max(LL2)

            for(m in 1:length(CC0)){
                for(n in 1:length(CC1)){
                    for(o in 1:length(DD1)){
                        for(p in 1:length(w2)){

                    if(LL2[m,n,o,p]==MaxLL2)        {

                        c00=CC0[m]
                        c10=CC1[n]
                        d10=DD1[o]
                        w20=w2[p]

    }
                                                    }
                                                }
                                            }
                                        }

        mm2=cbind(MaxLL2,c00,c10,d10,w20)

list(mm1=mm1,mm2=mm2,a00=a00,a10=a10,b10=b10,w10=w10,c00=c00,c10=c10,d10
=d10,w20=w20)

}

        #estimate parameters
        #--------------------

        ini=Estlog(newUV)
        n=length(u)

        AA00=ini$a00
        AA10=ini$a10
        BB10=ini$b10
        CC00=ini$c00
        CC10=ini$c10
        DD10=ini$d10
        w0=ini$w10
        omega=1

        MSC=ms(~-(-2*n*logb((pi),base=exp(1))-
        (n*logb((omega*w0),base=exp(1)))-(1/w0*sum((cos(u)-
        cos(U))^2+(sin(u)-sin(U))^2))-(1/omega*w0*sum((cos(v)-A0h-
        A1h*cos(U)-B1h*sin(U))^2+(sin(v)-C0h-C1h*cos(U)-
        D1h*sin(U))^2))),start=
        list(A0h=AA00,A1h=AA10,B1h=BB10,C0h=CC00,C1h=CC10,D1h=DD10
        ))

        maxMSC=MSC$value
        para=MSC$parameters

        A0h=para[1]
        A0h=as.vector(A0h)
        A1h=para[2]
        A1h=as.vector(A1h)
        B1h=para[3]
        B1h=as.vector(B1h)
        C0h=para[4]
        C0h=as.vector(C0h)
        C1h=para[5]
        C1h=as.vector(C1h)
        D1h=para[6]
        D1h=as.vector(D1h)


        #estimate sigma^2
```

```
#----------------

sh=1/n*(sum((cos(u)-cos(U))^2+(sin(u)-
sin(U))^2)+1/omega*sum((cos(v)-A0h-A1h*cos(U)-
B1h*sin(U))^2 + (sin(v)-C0h-C1h*cos(U)-D1h*sin(U))^2))

#estimate U
#----------

C=sum(cos(u)*sin(U))/sin(u)
+sum(sqrt((B1h^4 + 6*A1h^2+ B1h^2 + A1h^4)*sin(U)+((6*A0h
-       6*sin(v))*A1h*B1h^2       +       (2*A0h       -
2*sin(v))*A1h^3)*sin(U)+(A0h^2     -     2*sin(v)*A0h     +
sin(v)^2)*A1h^2)    +(B1h^2    +    A1h^2)*sin(U)+(A0h    -
sin(v))*A1h)/2*A1h*B1h+sum(sqrt((D1h^4    +    6*C1h^2    +
C1h^4)*sin(U)+((6*C0h  -  6*sin(v))*C1h*D1h^2  +  (2*C0h  -
2*sin(v))*C1h^3)*sin(U)+(C0h^2     -     2*sin(v)*C0h     +
sin(v)^2)*C1h^2)+(D1h^2     +     A1h^2)*sin(U)+(C0h     -
sin(v))*C1h)/2*C1h*D1h

S=sum(sin(u)*cos(U))/cos(u)  +sum(sqrt((B1h^4  +  6*A1h^2  +
B1h^2 + A1h^4)*cos(U)+((6*A0h - 6*cos(v))*A1h*B1h^2+(2*A0h
-    2*cos(v))*A1h^3)*cos(U)     +(A0h^2-2*cos(v)*A0h     +
cos(v)^2)*A1h^2)   +((B1h^2)   +   (A1h^2))*cos(U)+(A0h   -
cos(v))*A1h)/2*A1h*B1h+sum(sqrt((D1h^4    +    6*C1h^2    +
C1h^4)*cos(U)+((6*C0h  -  6*cos(v))*C1h*D1h^2  +  (2*C0h  -
2*cos(v))*C1h^3)*cos(U)+(C0h^2     -     2*cos(v)*C0h     +
cos(v)^2)*C1h^2)+(D1h^2     +     A1h^2)*cos(U)+(C0h     -
cos(v))*C1h)/2*C1h*D1h

Uh=atan(S,C)
Uh=Uh %% (2. *pi)

#estimate V
#----------
cosV=A0h + A1h*cos(Uh) + B1h*sin(Uh)
cosV=as.matrix(cosV)

sinV=C0h + C1h*cos(Uh) + D1h*sin(Uh)
sinV=as.matrix(sinV)

Vh=atan(sinV,cosV)


# estimates errors
#-----------------

delRh=cos(u)-cos(Uh)
delIh=sin(u)-sin(Uh)
epsRh=cos(v)-cos(Vh)
epsIh=sin(v)-sin(Vh)

list(u=u,v=v,Uh=Uh,Vh=Vh,A0h=A0h,A1h=A1h,B1h=B1h,C0h=C0h,C
1h=C1h,D1h=D1h,sh=sh,delRh=delRh,delIh=delIh,epsRh=epsRh,e
psIh=epsIh)

}

Es=complexEstN(newUV)

pe=matrix(0,nrow=B,ncol=7)

pe[i,1]=Es$A0h
pe[i,2]=Es$A1h
pe[i,3]=Es$B1h
pe[i,4]=Es$C0h
pe[i,5]=Es$C1h
pe[i,6]=Es$D1h
pe[i,7]=Es$sh

a0e=pe[,1]
```

```
                a1e=pe[,2]
                b1e=pe[,3]
                c0e=pe[,4]
                c1e=pe[,5]
                d1e=pe[,6]
                sige=pe[,7]

                a0emean=mean(a0e)
                a1emean=mean(a1e)
                b1emean=mean(b1e)
                c0emean=mean(c0e)
                c1emean=mean(c1e)
                d1emean=mean(d1e)
                sigemean=mean(sige)

                a0ese=stdev(a0e)
                a1ese=stdev(a1e)
                b1ese=stdev(b1e)
                c0ese=stdev(c0e)
                c1ese=stdev(c1e)
                d1ese=stdev(d1e)
                sigese=stdev(sige)

        meane=c(a0e=a0emean,a1e=a1emean,b1e=b1emean,c0e=c0emean,c1e=c1emean,d1e=
        d1emean,sige=sigemean)
        see=c(a0e=a0ese,a1e=a1ese,b1e=b1ese,c0e=c0ese,c1e=c1ese,d1e=d1ese,sige=s
        igese)

        estfunc=cbind(meane,see)
        }

        list(estfun=estfun)
}

#---unreplicated JS circular functional relationship--#


        complexEst=function(DATA){

                u=DATA[,1]
                v=DATA[,2]
                U=u

                n<-length(DATA)

        #step to get initial values#
        #-------------------------#

        Estlog=function(DATA){

                loglike1=function(AA0,AA1,BB1,w1,DATA){

                u=DATA[,1]
                v=DATA[,2]
                U=u
                n=length(u)
                omega1=1

                L1=2/omega1*w1*(sum(cos(v)-AA0-AA1*cos(U)-BB1*sin(U)))

                list(L1=L1)
                }

                loglike2=function(CC0,CC1,DD1,w2,DATA){
                u=DATA[,1]
                v=DATA[,2]
                U=u
                n=length(u)
                omega1=1

                L2=2/omega1*w2*(sum(sin(v)-CC0-CC1*cos(U)-DD1*sin(U)))
```

```
	list(L2=L2)
	}

	# get the initial value #
	#------------------------#

	AA0=seq(-1,1,0.3)
	AA1=seq(-1,1,0.3)
	BB1=seq(-1,1,0.3)
	CC0=seq(-1,1,0.3)
	CC1=seq(-1,1,0.3)
	DD1=seq(-1,1,0.3)
	w1=seq(0,1,0.01)
	w2=seq(0,1,0.01)

LL1=array(0,dim=c(length(AA0),length(AA1),length(BB1),length(w1)))

	for(i in 1:length(AA0)){
		for(j in 1:length(AA1)){
			for(k in 1:length(BB1)){
				for(l in 1:length(w1)){

			LL1[i,j,k,l]=loglike1(AA0[i],AA1[j],BB1[k],w1[l],
DATA)$L1

								}
							}
						}
					}


	LL2=array(0,dim=c(length(CC0),length(CC1),length(DD1),length(w2))
)

	for(m in 1:length(CC0)){
		for(n in 1:length(CC1)){
			for(o in 1:length(DD1)){
				for(p in 1:length(w2)){

LL2[m,n,o,p]=loglike2(CC0[m],CC1[n],DD1[o],w2[p],DATA)$L2

								}
							}
						}
				}

	#plot(LL1)
	#plot(LL2)

	MaxLL1=max(LL1)

	for(i in 1:length(AA0)){
		for(j in 1:length(AA1)){
			for(k in 1:length(BB1)){
				for(l in 1:length(w1)){


		if(LL1[i,j,k,l]==MaxLL1)		{

			a00=AA0[i]
			a10=AA1[j]
			b10=BB1[k]
			w10=w1[l]
								}
							}
						}
					}
				}
```

```
mm1=cbind(MaxLL1,a00,a10,b10,w10)

MaxLL2=max(LL2)

for(m in 1:length(CC0)){
       for(n in 1:length(CC1)){
              for(o in 1:length(DD1)){
                     for(p in 1:length(w2)){

       if(LL2[m,n,o,p]==MaxLL2)          {

              c00=CC0[m]
              c10=CC1[n]
              d10=DD1[o]
              w20=w2[p]
                                                                 }
                                                           }
                                                   }
                                            }
                                     }


mm2=cbind(MaxLL2,c00,c10,d10,w20)

list(mm1=mm1,mm2=mm2,a00=a00,a10=a10,b10=b10,w10=w10,c00=c00,c10=
c10,d10=d10,w20=w20)

}

#get the parameter estimate based on the initial value
#------------------------------------------------------

ini=Estlog(DATA)
n=length(u)

AA00=ini$a00
AA10=ini$a10
BB10=ini$b10
CC00=ini$c00
CC10=ini$c10
DD10=ini$d10
w0=ini$w10
omega=1

MSC=ms(~-(-2*n*logb((pi),base=exp(1))-
(n*logb((omega*w0),base=exp(1)))-(1/w0*sum((cos(u)-
cos(U))^2+(sin(u)-sin(U))^2))-(1/omega*w0*sum((cos(v)-A0h-
A1h*cos(U)-B1h*sin(U))^2+(sin(v)-C0h-C1h*cos(U)-
D1h*sin(U))^2))),start=
list(A0h=AA00,A1h=AA10,B1h=BB10,C0h=CC00,C1h=CC10,D1h=DD10))

maxMSC=MSC$value

para=MSC$parameters

A0h=para[1]
A0h=as.vector(A0h)
A1h=para[2]
A1h=as.vector(A1h)
B1h=para[3]
B1h=as.vector(B1h)
C0h=para[4]
C0h=as.vector(C0h)
C1h=para[5]
C1h=as.vector(C1h)
D1h=para[6]
D1h=as.vector(D1h)


#estimate U & V
#--------------
```

```
C=sum(cos(u)*sin(U))/sin(u)
+sum(sqrt((B1h^4 + 6*A1h^2+ B1h^2 + A1h^4)*sin(U)
+((6*A0h - 6*sin(v))*A1h*B1h^2 + (2*A0h -
2*sin(v))*A1h^3)*sin(U)+(A0h^2 - 2*sin(v)*A0h +
sin(v)^2)*A1h^2)+(B1h^2 + A1h^2)*sin(U)+(A0h -
sin(v))*A1h)/2*A1h*B1h +sum(sqrt((D1h^4 + 6*C1h^2 + C1h^4)*sin(U)
+((6*C0h - 6*sin(v))*C1h*D1h^2 + (2*C0h -
2*sin(v))*C1h^3)*sin(U)+(C0h^2 - 2*sin(v)*C0h + sin(v)^2)*C1h^2)
+(D1h^2 + A1h^2)*sin(U) +(C0h - sin(v))*C1h)/2*C1h*D1h

S=sum(sin(u)*cos(U))/cos(u) +sum(sqrt((B1h^4 + 6*A1h^2 +
B1h^2 + A1h^4)*cos(U)+((6*A0h - 6*cos(v))*A1h*B1h^2+(2*A0h -
2*cos(v))*A1h^3)*cos(U) +(A0h^2-2*cos(v)*A0h + cos(v)^2)*A1h^2)
+((B1h^2) + (A1h^2))*cos(U) +(A0h -
cos(v))*A1h)/2*A1h*B1h+sum(sqrt((D1h^4 + 6*C1h^2 + C1h^4)*cos(U)
+((6*C0h - 6*cos(v))*C1h*D1h^2 + (2*C0h - 2*cos(v))*C1h^3)*cos(U)
+(C0h^2 - 2*cos(v)*C0h + cos(v)^2)*C1h^2) +(D1h^2 + A1h^2)*cos(U)
+(C0h - cos(v))*C1h)/2*C1h*D1h

Uh=atan(S,C)
Uh=Uh %% (2. *pi)

cosV=A0h + A1h*cos(Uh) + B1h*sin(Uh)
cosV=as.matrix(cosV)

sinV=C0h + C1h*cos(Uh) + D1h*sin(Uh)
sinV=as.matrix(sinV)

Vh=atan(sinV,cosV)

#estimate sigma^2
#-----------------

sh=1/n*(sum((cos(u)-cos(U))^2+(sin(u)-sin(U))^2)+
1/omega*sum((cos(v)-A0h-A1h*cos(U)-B1h*sin(U))^2+(sin(v)-C0h-
C1h*cos(U)-D1h*sin(U))^2))

# estimate errors
#-----------------

delRh=cos(u)-cos(Uh)
delIh=sin(u)-sin(Uh)
epsRh=cos(v)-cos(Vh)
epsIh=sin(v)-sin(Vh)

list(u=u,v=v,Uh=Uh,Vh=Vh,A0h=A0h,A1h=A1h,B1h=B1h,C0h=C0h,C1h=C1h,
D1h=D1h,sh=sh,delRh=delRh,delIh=delIh,epsRh=epsRh,epsIh=epsIh)
}
```