Influence Diagnostic Methods in Generalized Linear Models with Biased Estimators

Muhammad Amin

A thesis presented in partial fulfillment of the requirement for the degree of Doctor of Philosophy in Statistics at Bahauddin Zakariya University, Multan, Pakistan.

Department of Statistics

October, 2016

Certificate

This is to certify that the thesis entitled, **"Influence Diagnostic Methods in Generalized Linear Models with Biased Estimators."** submitted by **Mr. Muhammad Amin** has been thoroughly studied. It is found to be prolific in scope and quality as a thesis for the award of Degree of Doctor of Philosophy in Statistics.

Dr. Muhammad Aman Ullah,
Supervisor,
Prof. & Chairman,
Department of Statistics,
Bahauddin Zakariya University, Multan Pakistan.

Dr. Muhammad Aslam,
Co-Supervisor,
Associate Professor,
Department of Statistics,
Bahauddin Zakariya University, Multan Pakistan.

This thesis is dedicated to My Family

Especially to My Mother

Acknowledgements

I most grateful pay me appreciation to Almighty **ALLAH**, the Most Gracious and the Most Merciful for bestowing His countless blessings upon me all the way through my life and Who endowed me with courage for completing my research work. I crave to pay my utmost admiration for our dearest Holy Prophet,**Hazrat MUHAMMAD (S.A.W.W)** who is the real role model for all mankind.

I would like to thank and express my gratitude to my Supervisor **Dr. Muhammad Amanullah**, Professor and Chairman, Department of Statistics, Bahauddin Zakariya University, Multan for his consistent guidance, invaluable help and suggestions during this research work.

I am also grateful to my Co-Supervisor **Dr Muhammad Aslam**, Associate Professor at Department of Statistics, Bahauddin Zakariya University, Multan. Who has contributed significantly to my understanding and appreciation of Statistics. I would like to express my gratitude to all the esteemed faculty members and staff members of the Department of Statistics, Bahauddin Zakariya University, Multan.

I also thanks to **Mr. Muhammad Awais Manzoor** for helping me in computational work on statistical softwares.

Finally, I wish to thank to my parents and my wife for their constant support, patience, inspiration, good cheerfulness, and encouragement, especially during financially hard times; and to my Brother for giving up time that belonged to them.

Muhammad Amin

Abstract

This thesis is concerned with the development of influence diagnostic methods for the gamma and the inverse Gaussian regression models. In this thesis, we also developed diagnostics using case omission approaches in the GLM with ridge estimator as an alternative to the diagnostics as computed with IRLS.

Firstly, we present the GLM theory with estimation using MLE and IRLS. Then the formulation of different GLM residuals along with standardized form of each residuals are given. We proposed the different influence diagnostic methods i.e. CD, MCD, AP, CVR, DFFITS, WD and H_d for the GLM with IRLS estimation.

Secondly, we have given the GRM estimation theory along with the GRM residuals. Each GRM residual is further categorized into standardized and adjusted residuals. The derivation of the adjusted Pearson and adjusted deviance residuals for the GRM are developed. The GRM influence diagnostics based on these forms of residuals are proposed. Similarly the derivation of the GRM with ridge estimator is presented. The computation of the GRM residuals with ridge estimator are also proposed. Then the GRM influence diagnostic methods with ridge estimator based on standardized and adjusted residuals are formulated.

The comparison of all proposed GRM influence detection methods with each form of residuals are assessed through simulation study and a real data set. Results show that genuine influential observations are detected by the CVR method with all forms of the GRM residuals. These results also show that the detection of influential observation by all diagnostic methods are found to be better with likelihood residuals and with ridge estimator. However, the influential observation detection by H_d method with all forms of the GRM residuals is very poor. The influence detection performance by CD, MCD and DFFITS are same but not better than the CVR method (See Table 3.4-3.8).

Finally, we conduct a study for testing the influence diagnostic methods performance with different forms of IGRM residuals. Simulation and example results show that again the CVR method is found to be better for the detection of influential observations in the IGRM. But the performance of the IGRM residuals in detecting the influential observations is different as GRM residuals. Result shows that the influence diagnostics with working forms IGRM residuals are better as all other form of IGRM residuals. We also observe by all proposed diagnostic methods that the influential observation detection performance with adjusted forms of all IGRM and IGRRM residuals are approximately identical except the H_d method.

List of Symbols and Abbreviations

Abbreviation/Symbols	**Description**
AP	Andrew's Pregibone
AP_R	Andrew's Pregibone with ridge estimate
Ar_a	Adjusted Anscombe residuals
Ar_{aR}	Adjusted Anscombe residuals with ridge estimate
Ar_d	Adjusted deviance residuals
Ar_{dR}	Adjusted deviance residuals with ridge estimate
Ar_l	Adjusted likelihood residuals
Ar_{lR}	Adjusted likelihood residuals with ridge estimate
Ar_P	Adjusted Pearson residuals
Ar_{PR}	Adjusted Pearson residuals with ridge estimate
Ar_W	Adjusted working residuals
Ar_{WR}	Adjusted working residuals with ridge estimate
CD	Cook's distance
CD_R	Cook's distance with ridge estimate
CI	Condition Index

CVR	Covariance Ratio
CVR_R	Covariance Ratio with ridge estimate
$DFFITS$	Difference of Fits Test
$DFFITS_R$	Difference of Fits Test with ridge estimate
GLM	Generalized linear model
GRM	Gamma regression model
GRRM	Gamma ridge regression model
$\boldsymbol{H}$	Hat Matrix
Hd	Hadi's measure
Hd_R	Hadi's measure with ridge estimate
H_R	Hat Matrix with ridge estimate
h_{ii}	Leverages
h_{Rii}	Leverages with ridge estimate
$\boldsymbol{I}$	Identity Matrix
$\boldsymbol{IG}$	Inverse Gaussian
IGRM	Inverse Gaussian regression model
IGRRM	Inverse Gaussian ridge regression model
$\boldsymbol{K}$	ridge parameter
LM	Linear model
MCD	Modified Cook's Distance
MCD_R	Modified Cook's Distance with ridge estimate

OLS	Ordinary least square
r_a	Anscombe residuals
r_{aR}	Anscombe residuals with ridge estimate
r_d	Deviance residuals
r_{dR}	Deviance residuals with ridge estimate
r_l	Likelihood residuals
r_{lR}	Likelihood residuals with ridge estimate
r_P	Pearson residuals
r_{PR}	Pearson residuals with ridge estimate
r_W	Working residuals
r_{WR}	Working residuals with ridge estimate
sr_a	Standardized Anscombe residuals
sr_{aR}	Standardized Anscombe residuals with ridge estimate
sr_d	Standardized deviance residuals
sr_{dR}	Standardized deviance residuals with ridge estimate
sr_P	Standardized Pearson residuals
sr_{PR}	Standardized Pearson residuals with ridge estimate
sr_W	Standardized working residuals
sr_{WR}	Standardized working residuals with ridge estimate
WD	Welsch distance
WD_R	Welsch distance with ridge estimate

$\boldsymbol{X}$	matrix of independent variables
y	Response vector
y^*	adjusted response vector
$\boldsymbol{Z}$	design matrix of standardized independent variables
ϕ	Dispersion parameter
ϕ_R	Dispersion parameter with ridge estimate
β	Vector of slope coefficients of x
β^*	Vector of slope coefficients of Z
β_R	Vector of slope coefficients with ridge estimate
μ	Mean
$V(\mu)$	Variance function
W	Weight Matrix
λ_j	jth eigen value of X^TWX**matrix**

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Overview

Statistical modeling is useful for decision making because it can capture the relationship (linear or nonlinear) among variables. Primarily, regression models are developed for determining the relationships between the observed response and corresponding covariates. There are two main purposes of regression modeling; prediction, and determination of factors that may change the observed response.

The ordinary least squares (OLS) method is commonly used for plausible estimating the unknown parameters with the assumption that the response variable follow a normal distribution. While in practice, the condition of normality for the residuals violates, and it follows to other exponential family of distribution. Then for modeling of such response variable, we have the generalized linear model (GLM). In this situation, the OLS estimates are unreliable and there is need of some other methods. The GLM has been introduced by Nelder and Weddernburn (1972) when distribution of the response variable is non-normal and follows the exponential family distribution such as, Poisson, binomial, negative binomial, exponential and gamma. These models are also called non-linear models and are most popular since last few decades (Wood, 2005; Wang *et al.*, 2010). Non-linearity in the model arises only due to link function, called canonical link (Kaiser, 1997). Nelder (1968) transformed such type of non-linear models into linear models. The GLM is used to determine the relationship between independent variables and the function of the mean of the dependent variable (Lo and Ronchetti, 2009; Haung and Song, 2011).

According to nature of the response variable, various forms of the GLM are available in the literature, including variance models, log-linear models and multinomial models for count data and the gamma model for continuous data. For further details and examples, see McCullagh and Nelder (1989). A lot of work has already been done by different researchers for some of the exponential family of distributions, including Poisson, binomial, normal distribution using the GLM (Breslow, 1990; Consul and Famoye, 1992; Famoye, 1993, 2004; Beirlant and Goegebeur, 2003; Ferrari *et al.*, 2004; Ortega *et al.*, 2006; Cancho *et al.*, 2009; Hashimoto *et al.*, 2010; Prudenti and Cordeiro, 2010; Vanegas *et al.*, 2012).

These models have many applications in health sciences,such as to predict the effect of animal age on the weight of the dried eye lens (Ratkowsky, 1983), modeling renal failure data to estimate the prevalence of disease by different factors (Staniswalis, 2006), modeling life time data (Ortega *et al.*, 2012), and transportation problems (Wood, 2005). In hydrology, the gamma distribution has advantage of having only positive values, since hydrological variables such as rainfall and runoff are always positive (Markovic, 1965).

1.2 Influence Diagnostics

Residuals and influential analyses are the main regression diagnostics for testing of model validity. The interpretation of statistical analysis is difficult without examining the residuals. Residuals and leverages are used to identify outliers and influential observations (Cordeiro, 2004; Zare and Rasekh, 2012). An observation with large residuals may not be influential while observation with small residuals may affect the model results example given by Andrews and Pregibon (1978). Definition of residuals in the GLM was first given by Williams (1984) and later on by Pierce and Schafer (1986). Cordeiro (2004) studied the asymptotic properties of residuals and found that model diagnostics are necessary for the validity and fitting of the model.

The main objective of influence diagnostics is to determine the strange effect of influential observations on the model fitting and inferences (Xiang *et al.*, 2002). Various approaches for the assessment of influential observations in the LM have been proposed in the literature, i.e. Cook and Weisberg (1982), Atkinson (1985), Cook (1986), Chatterjee and Hadi (1988) and Nurunnabi *et al.* (2014, 2016). Pregibon (1981) was first proposed the influence diagnostics for the GLM as in logistic regression. Lee (1986) proposed a method for assessing the partial influence on the GLM estimates. Williams (1987) also proposed model diagnostics using deviance and single deletion case methods. Landwehr and Pregibon (1993) used partial residuals plots for the GLMs with different canonical links to determine the influential observations. Thomas and Cook (1989) proposed a method to assess the influence on the GLM regression coefficients. Preisser and Qaqish (1996) proposed a deletion diagnostic technique for generalized estimating equations (GEE), which is alternative to the GLM. A similar study was conducted by Venezuela (2007). Zhu and Lee (2001) modified the Cooks method for the assessment of local influence with minor perturbation for latent variable models. Xiang *et al.* (2002) and Xu *et al.* (2006) proposed influential and deletion measures for generalized linear mixed effects models. Xie and Wei (2007) proposed diagnostic methods for censored generalized Poisson regression model. Espinheira *et al.* (2008a and 2008b) proposed residuals and influential observation detection procedures in beta regression. Shi *et al.* (2009) proposed the local influence measure with missing covariates in the GLM. Shi and Chen (2009) proposed an influential measure for the GLM with correlated errors. Xu and Lavalley (2011) studied the relationship between the Pearson residuals and the efficiency of the GLM parameter estimates. Ferri (2011) investigated the diagnostic tools in beta regression with varying dispersion. Vanegas *et al.* (2012) proposed diagnostic procedures for generalized Weibull linear regression model. Villegas *et al* (2013) proposed influential analysis of generalized symmetric linear models.

Influence analysis with error in variables was also studied by various authors (Cook 1986; Fuller, 1987; Galea-Rojas *et al.*, 2002; Kelly, 1984; Lee and Zhao, 1996; Nakamura, 1990; Rasekh and Fieller, 2003; Wellman and Gunst, 2003; Wu and Luo, 1993; Zhao and Lee, 1994-1995). Zhong *et al.* (2000) proposed influential measure for a normal linear model with error in variables. Xie and Wei (2009) proposed diagnostic procedure for generalized Poisson regression models with error in variables.

1.3 Influential Observations and Model Estimates

Finding the values of the independent variables which are greatly influencing the regression estimates is an important task in regression modeling. Extreme values in the independent variables results an influential observation. All observations in any regression modeling are important for the estimation and inferences. Influential observation strongly affects the model estimates and inferences (Cook, 1997). Removing these observation results in a change in the parameter estimates and their standard errors. These may cause the regression results misleading in predicting/ forecasting and other decision purposes. Sometimes the most significant factors are still insignificant due to influential observation. Therefore, it is necessary to consider these influential observations, remove these observations and refit the model to make the results more reliable. Cook (1986) described the process of identification, removing influential observation and then refitting the regression model which is also known as influence analysis.

1.4 Multicollinearity

Multicollinearity, approximate linear dependencies among the covariates, is common problem in the LM and GLM. There are two types of multicollinearity; one is numerical and second is statistical (Belsley *et al.*, 1980; Gunst, 1983). This problem has a direct effect on the estimated parameters and leads high variances and instable parameters thats why the results are unreliable (Weissfeld and Sereika, 1991), and termed as the numerical effect. The effects of multicollinearity are well known and the solution depends on the sources of multicollinearity (Mason and Gunst, 1985). This problem can arise because of the following circumstances

- Model constraints
- The variance function is correctly specified
- Population characteristics that restrict variable values
- Sampling characteristics and
- Over defined models

Recently, the problem of multicollinearity have been studied in the GLM. One can refer to Mackinnon and Puterman (1989) for the consideration of multicollinearity in the GLM. Schaeffer *et al.* (1984) consider multicollinearity in the logistic regression. Lesaffre and Marks (1993) studied ill-conditioning in logistic regression.

1.5 Biased Estimators

Since multicollinearity is the data problem and that causes unreliable estimates in regression analysis, so due to this problem the variances of the estimates are so large and some important independent variables become insignificant. This is only due to multicollinearity, so the solution of the problem is necessary for regression modeling. One solution is the use of biased estimators which can reduce the effect of multicollinearity on the regression model estimates. The biased estimators includes a ridge estimation method, principal component method, Stein estimator and the Liu estimators. Among these methods, the most popular method is the use of ridge estimation method (Fox, 1991).

1.6 Influence Diagnostics with Biased Estimators

There are some situations, where the problem of multicollinearity and influential observation occurs simultaneously. Fox (1991) reported that weak collinearity may cause of influential observation, but not always true. While severe collinearity induced large hat values and are always reason of influential observation. Therefore, the influence diagnostics with multicollinearity is essential for better understanding of the statistical model. Under this situation, biased estimation methods are used to control the problem of multicollinearity. In order to study the influence diagnostics with biased estimators, some of the worthwhile studies are available in the literature. Walker and Birch (1988), Jahufer (2014) and Emami and Emami (2016) proposed the application of case deletion approaches in the ridge regression models. They modified the influence measures with ridge estimators. Jahufer and Jianbao (2009) studied the global influential observation in modified ridge regression. Billor and Loynes (1999) and Shi and Wang (1999) studied the local influence diagnostics with ridge estimator. Jahufer and Jianbao (2011) and Ullah *et al.* (2013) proposed the influence diagnostics in modified ridge regression. Jahufer and Jianbao (2013) study the influence diagnostics in the LM with Liu estimators. Later on Jahufer and Jianbao (2012) and Etras *et al.* (2013) explored the influence diagnostics with Liu and modified Liu estimators. All they studied the influence diagnostics under the assumption that the errors follow a normal distribution. However, there are the situations, where the error of the model follows some exponential family of distribution other than normal and the independent variables are linearly dependent. Under these situations, there is need to model such variables with ridge estimators and then assess the influence diagnostics.

1.7 Motivation

If one wants to model the response variables with two or more independent variables, then it is necessary that there is no exact linear relationship between two independent variables. If this exists, then estimation and inferences are unreliable due to large variances. In this situation, we use some biased estimation methods for the GLM fitting with collinear independent variables. After the GLM fitting under multicolinearity, one should be confident that there is no influential observation in the data set. Now the question arises that how we identify the influential observation in the GLM with multicollinearity. To answer this question, we propose methods for identifying the influential observations with multicollinearity. We have already seen in the literature that influence analysis regarding the GLM was studied for the simple GLM without considering other error assumptions. We have also seen that biased estimation for some of the GLM i.e. inverse Gaussian, exponential, gamma and chi-square have not been yet explored. No work has been done yet for the detection of influential observations in the GLM with multicollinearity.

1.8 Objectives of the Research

This research leads to the assessment of influential observation of the GLM with collinear independent variables. In view of the literature and motivation, our research objectives are as follows;

1. Fitting some of the GLM model by iterative procedures
2. Fitting some of the GLM model by ridge estimator
3. Developing influential observation detection methods for some of the GLM with and without multicollnearity in the independent variables
4. Determining the effect of omission of single influential observations on some of the GLM estimates with and without multicollinearity.
5. Comparison of various types of residuals in some of the GLM with and without multicollinearity to determine the influential observations.

1.9 Thesis Outline

Thesis outline is as follows. In Chapter 2, formulation and estimation of the LM and the GLM is presented. The LM residuals, hat matrix and influence diagnostic measures are presented in this chapter. Also the LM influence diagnostics extended to the GLM influence diagnostics with different forms of the GLM residuals are presented. In Chapter 3, the estimation of the gamma regression model (GRM) and the gamma ridge regression model (GRRM) is presented. Also in this chapter, we modified the different residuals and some influence measures with multicollinearity and without multicollinearity for the GRM. The comparison of influence measures with ridge estimator and the GRM without ridge estimator is illustrated with the help of Monte Carlo simulation and a real example. In Chapter 4, we present the estimation of the inverse Gaussian regression model (IGRM) and the inverse Gaussian ridge regression model (IGRRM). Also in this chapter, we modified the different residuals and some influence diagnostic methods with multicollinearity and without multicollinearity for the IGRM. The comparison of influence measures with and without ridge estimator is illustrated with the help of Monte Carlo simulation and a real example. In Chapter 5, we present some concluding remarks of the Chapter 3 and the Chapter 4 analysis.

Chapter 2

Influence Diagnostics in the Generalized Linear Models

2.1 Introduction

Multiple linear and non-linear models are the richest tools with wide variety of applications in science, social science, health sciences and engineering to explore the relationship between independent and dependent variables. The generalized linear models (GLM) are the models, where the dependent variables follow exponential family of distributions (Hardin and Hilbe, 2012). The iterative reweighted least square (IRLS) is the most popular method to produce the best GLM estimate. However, it is well known that the IRLS is affected by one or few influential observations. The detection of these substantial observations is an important task in the GLM theory and has great attentions in the literature. Pregibon (1981) first extended the work of Cook (1977) in LM to the GLM for detecting the influential observations as in logistic regression case. Later on many studies have been conducted with diverse exponential family of distributions to diagnose the influential observations. For details see (Pregibon, 1981; Lee, 1987,1988; Williams, 1987; Thomas and Cook, 1989; Ferrari and Cribari-Neto, 2008b; Xie and Wei, 2009; Ortega *et al.*, 2012; Venegas *et al.*, 2012).

2.2 The Linear Models

Before discussing the influence diagnostics in the GLM, we first overview the theory of the LM with influence diagnostics. The theory and estimation of the LM actually based on the normal theory. Let y_i be the dependent variable and follow normal probability distribution and let $x_{ij}; i = 1, 2, ..., n; j = 1, 2, ..., p$ are the p independent variables, then the LM with p independent variables and and $p' = p + 1$ coefficients(including intercept) is given as

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + ... + \beta_p x_{ip} + \in_i. \tag{2.1}$$

In matrix form, Eq.(2.1) can be written as

$$y = X\beta + \in, \tag{2.2}$$

where y is an $n \times 1$ vector of response variable, X is the $n \times p'$ design matrix which is constructed from categorical or continuous independent variables of full rank $p < n$, β is the $p' \times 1$ vector of unknown regression parameters and $\in$ be the $n \times 1$ vector of normally distributed random error with zero mean and constant variance σ^2. The model (2.2) is alternatively explained with the help of mean and variance functions of the dependent variable as

$$E\left(y\right)=\mu=X\beta,Var\left(y\right)=\sigma^{2}.$$

For the estimation of unknown parameter β in (2.2), the OLS method or maximum likelihood estimation (MLE) method is used under the assumption that the error are normally distributed. The OLS method minimizes the sum of squares of errors as

$$S\left(\beta\right)=\sum_{i=1}^{n}\in_{i}^{2}=\in^{T}\in=\left(y-X\beta\right)^{T}\left(y-X\beta\right).$$

$$S\left(\beta\right)=\ y^{T}y-2\beta^{T}X^{T}y+\beta^{T}X^{T}X\beta \tag{2.3}$$

Now by the Eq. (2.3), we have

$$\frac{\partial S}{\partial\beta}=-2X^{T}y+2X^{T}X\beta=0.$$

This yield the estimator $\hat{\beta}$ of β as

$$\hat{\beta}=\left(X^{T}X\right)^{-1}X^{T}y. \tag{2.4}$$

It is already verified that the variance of the least square estimator $\hat{\beta}$ may be written as

$$Var\left(\hat{\beta}\right)=\sigma^{2}\left(X^{T}X\right)^{-1}.$$

2.3 Hat Matrix, Leverage and Residuals in the Linear Model

The leverages and residuals are the main tools for the detection of influential observations. Now considering the model (2.2), the fitted values of the dependent variable are defined by

$$\hat{y} = X\hat{\beta}. \tag{2.5}$$

The estimated residuals for (2.2) are defined by

$$\hat{\in} = y - \hat{y} = y - X\hat{\beta}. \tag{2.6}$$

The hat matrix for the LM is defined as

$$H = X\left(X^T X\right)^{-1} X^T. \tag{2.7}$$

The hat matrix H is symmetric and idempotent as $H^2 = H$. The leverages in the LM are used to identify the influential observations. These for the LM are formulated as

$$h_{ii} = diag\left(H\right). \tag{2.8}$$

where, $0 < h_{ii} < 1$ or $\frac{1}{n} < h_{ii} < 1$. Note that $h_{ii} = \sum_{j=1}^{n} h_{ij}^2$ and also $h_{ii}(1 - h_{ii}) \leq 0$. The term leverage is directly related to the independent variables, indicated that high leverage point in the X space results as influential observation. h_{ii} measure the influence on the fitted values.

The residual defined in Eq.(2.6) is also known as raw or fitted residual. Other forms of residuals include standardized and studentized residuals. Some authors use the term studentized residuals instead of standardized residuals. These residuals are further used in the LM diagnostics which includes outlier, influential observations and heteroscedasticity. For the LM, the standardized residuals are defined by

$$r_i = \frac{\hat{\in}_i}{\sqrt{1 - h_{ii}}}, \tag{2.9}$$

where $m_{ii} = 1 - h_{ii};\ 0 \leq m_{ii} \leq 1, m_{ii}$ has the similar properties as h_{ii} has. Since X is full rank matrix so

$$\sum h_{ii} = p' \ and \ \sum m_{ii} = n - p',$$

where $p' = p + 1$. Similarly, the studentized residuals for the LM are computed by the relations as

$$t_i = \frac{\hat{\in}_i}{\hat{\sigma}\sqrt{1 - h_{ii}}}, \tag{2.10}$$

where $\hat{\sigma}^2 = \frac{\in^T \in}{n-p-1}$ is an unbiased estimator of the variance. t_i is used to test the outlier and influential observation in the LM.

2.4 The LM Influence Diagnostics

In this section, we present some influence diagnostics which are proposed for the LM.

2.4.1 Summary of Influence Diagnostic Methods

Since the influential observation effect the slope coefficients and standard error of the estimates. Various authors proposed the influence diagnostics in the LM. For further details see Belsley *et al.* (2004). These influence diagnostics are summarized in Table 2.1.

Table 2.1: Summary of the LM influence diagnostics

Influence Diagnostics	Statistic/Formula	Cut Point	Reference
Leverage	$h_{ii} = x_i{}^T \left(X^T X\right)^{-1} x_i$	$h_{ii} \geqslant \frac{2p'}{n}$	Belsley *et al.* (1980)
Cooks Distance	$CD_i = \frac{t_i^2}{p'} \left(\frac{h_{ii}}{m_{ii}}\right)$	$CD_i > F_{\alpha(p,n-p)}$	Cook (1977)
Modified Cook Distance	$MCD_i = \left[\frac{n-p'}{p'} \frac{h_{ii}}{m_{ii}}\right]^{\frac{1}{2}} \lvert t_i \rvert$	$MCD_i \geqslant 2\sqrt{\frac{n-p'}{n}}$	Belsley *et al.* (1980)
Andrew Pregibone	$AP_i = m_{ii} - \frac{\in_i^2}{\sum \in_i^2}$	$(1 - AP_i) > \frac{2p'}{n}$	Andrews and Pregibon (1978)
Covariance Ratio	$CVR_i = \frac{\left[\frac{\left(n-p-\in_i^2\right)}{n-p'}\right]^{p'}}{m_{ii}}$	$\lvert CVR_i - 1 \rvert > \frac{3p}{n}$	Ullah and Pasha (2009)
DFFITS	$DFFITS_i = \lvert t_i \rvert \sqrt{\frac{h_{ii}}{m_{ii}}}$	$\lvert DFFITS_i \rvert > 2\sqrt{\frac{p'}{n}}$	Ullah and Pasha (2009)
Welsch Distance	$WD_i = DFFITS_i \sqrt{\frac{n-1}{m_{ii}}}$	$WD_i > 3\sqrt{p'}$	Welsch. (1982)
Hadi Measure	$H_{di} = \frac{p'}{m_{ii}} \frac{d_i^2}{1-d_i^2} + \frac{h_{ii}}{m_{ii}},$	$H_{di} > Mean\left(H_{di}\right) + c\sqrt{Var\left(H_{di}\right)}$	Hadi (1992)

where $d_i = \frac{\in_i^2}{\sum \in_i^2}$, and $t_i = \frac{\in_i}{\hat{\sigma}\sqrt{m_{ii}}}$

2.4.2 Comparison of Influence Diagnostics in the LM

In the literature, some researchers develop new influence diagnostics and some researchers compare these influence diagnostics in the LM for the detection of influential observations. Chatterjee and Yilmaz (1992) say that there are 25 or more influence diagnostic measures are available in the literature. Of these diagnostics, some of them gives identical results and some give different results. Chatterjee and Hadi (2012) say that Cooks distance and DFFITS methods perform equally in identifying the influential observations because both diagnostics are the functions of leverages and residuals. So, they recommended that use only of these for the influence analysis. Ozkale (2013) say that if researcher's objective is to determine the influence on slope coefficients then Cooks distance is preferable. He also says further that DFFITS is used, if researcher determine the influence of an observation on both the slope coefficients and their variances.

2.5 Illustrative Examples

In this section, we discuss two real data sets which are used for the illustration of available influence diagnostics. In this chapter, these data sets are analyzed initially for the LM influence diagnostics. Later on, these data sets are used for our proposed influence diagnostics to show how our proposed diagnostic methods are significant from the LM influence diagnostics.

2.5.1 Example 1: Reaction Rate data

The reaction rate data set is taken from Huet *et al.* (2004). The main objective of this data set is to determine the reaction rate of the catalytic isomerization of n-pentane to iso-pentane based on partial pressure of different independent variables (factors). These independent variables are used to speed up the reaction rate. This data set consists of $n = 24$ experimental data values with one dependent variable i.e. reaction rate (y) and $p = 3$ independent variables i.e. partial pressure of hydrogen (x_1), partial pressure of n-pentane (x_2) and partial pressure of iso-pentane (x_3).

Table 2.2: The LM Influence Assessment of the Reaction Rate Data

Obs.No	$\hat{\in}_i$	$\hat{t}_i$	$\hat{h}_i$	CD_i
1	-0.5525	-0.7144	0.1928	0.0312
2	-0.0223	-0.0283	0.1822	0.0000
3	0.1096	0.1313	0.0836	0.0004
4	-0.7625	-0.995	0.188	0.0573
5	-0.3342	-0.429	0.1948	0.0116
6	1.0747	1.4201	0.1675	0.0965
7	-0.9051	-1.197	0.192	0.0833
8	0.2921	0.3853	0.2391	0.0122
9	-1.3762	-1.9359	0.2048	**0.2122**
10	0.3142	0.4073	0.2112	0.0116
11	0.1741	0.215	0.1362	0.0019
12	-0.0775	-0.0947	0.1197	0.0003
13	-0.4511	-0.5684	0.158	0.0157
14	0.7085	0.9415	0.221	0.0632
15	-0.4898	-0.58	0.0461	0.0042
16	0.2167	0.2551	0.0485	0.0009
17	-0.0003	-0.0003	0.0471	0.0000
18	0.2162	0.2543	0.0468	0.0008
19	-0.5663	-0.696	0.1075	0.015
20	**2.2699**	**4.0437**	0.2295	**0.6888**
21	-0.1165	-0.1508	0.215	0.0016
22	-1.1861	-1.7813	**0.3201**	**0.3368**
23	1.1524	1.5087	0.1414	0.0881
24	0.3121	0.432	**0.307**	0.0215

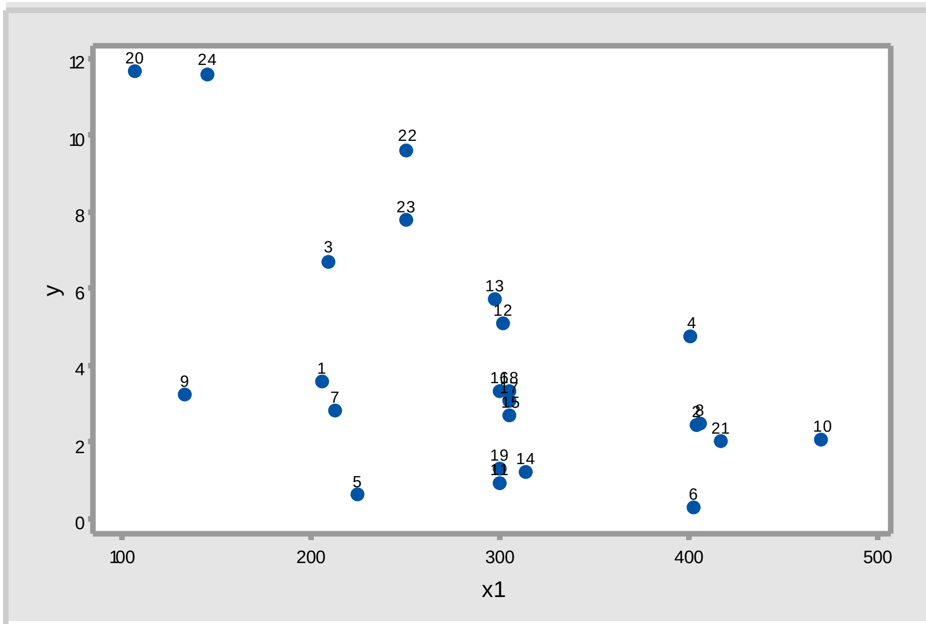


Figure 2.1: Scatter Plot y vs x_1 for Reaction Rate Data

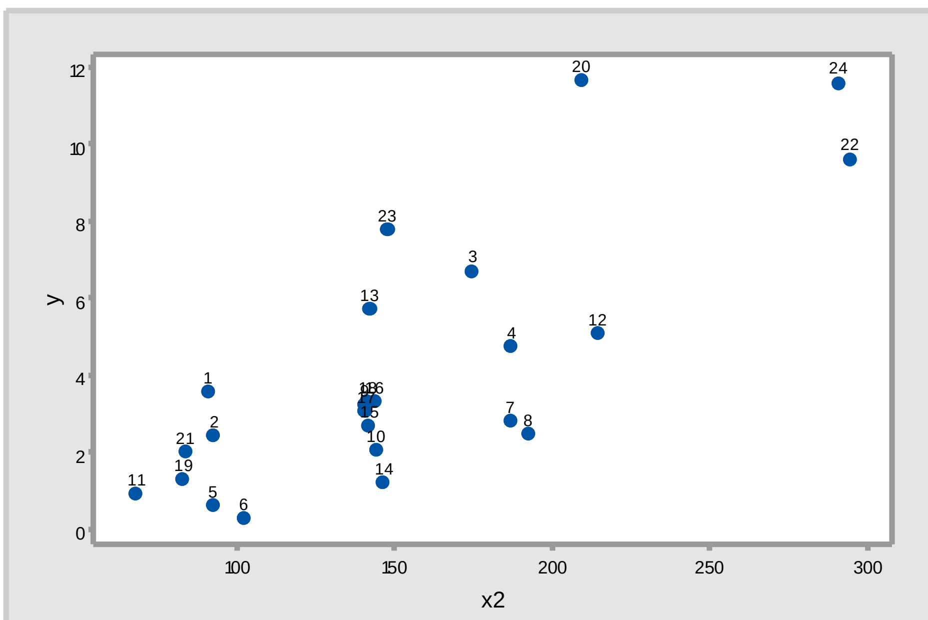


Figure 2.2: Scatter Plot y vs x_2 for Reaction Rate Data

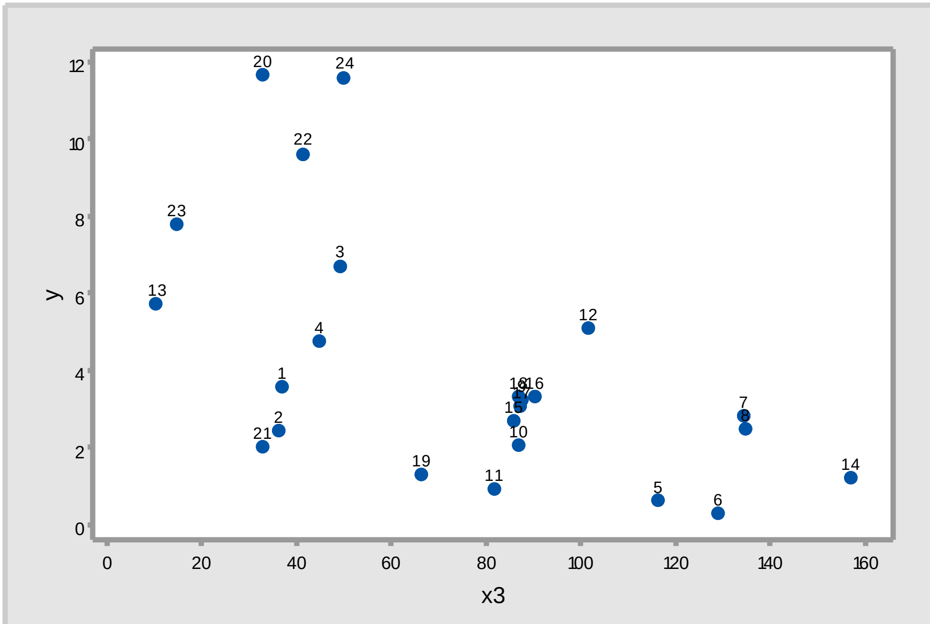


Figure 2.3: Scatter Plot y vs x_3 for Reaction Rate Data

From Fig. (2.1-2.3), we observe that 20^{th}, 22^{nd}, and 24^{th} observations are seemed to be influential observations. While from Table 2.2, we find that the LM diagnostic (CD) declare that the 9^{th}, 20^{th}, 22^{nd} and 24^{th} observations are the influential observations. In Table 2.3, we study the effect of influential observations on the LM estimates, standard errors of the estimates and their significance. We find that the standard errors of the estimates are reduced from 15 % to 27%.

Table 2.3: Effect of Influential Observations on the LM Estimates for Reaction Rate Data

	Full Data				Without Influential Observations			
	Estimate	S.E	T	p-value	Estimate	S.E	T	p-value
Constant	4.1168	0.9212	4.4688	0.0002	3.5346	0.7262	4.8676	0.0002
x_1	-0.0089	0.002	-4.4399	0.0003	-0.0072	0.0017	-4.2142	0.0007
x_2	0.0357	0.0032	11.2728	0.0000	0.0346	0.0032	10.6496	0.0000
x_3	-0.0386	0.0044	-8.7007	0.0000	-0.0361	0.0032	-11.1365	0.0000

Table 2.4: Effect of Influential Observations on the LM Goodness of Fit Criterias for Reaction Rate Data

Model Criterias	Full Data	Without Influential Observations
AIC	65.94	40.35
BIC	71.83	45.33
R^2	0.94	0.93
$\hat{\sigma}^2$	0.72	0.33
F-test	7.88	217.13
P-value	0.0000	0.0000

From Table 2.4, the influential observations are substantially affects the model goodness of fit criterias. The results shows that after removing the influential observations, model dispersion is reduced to 50%. The correlation matrix and VIF are used to diagnose the multicollinearity. From Table 2.5, we observe that there is no indication of multicollinearity among independent variables and same results are also observed from VIF after fitting the LM.

Table 2.5: Correlation Matrix and the VIF for Reaction Rate Data

	x_1	x_2	x_3	VIF
x_1	1.00	0.34	-0.16	1.17
x_2	0.34	1.00	0.01	1.14
x_3	-0.16	0.01	1.00	1.03

2.5.2 Example 2: Stack Loss Data

Brownlee (1965), first used the stack loss data for the model inferences. This data set consists of $n = 21$ operations of a plant for the oxidation of ammonia to nitric acid with dependent (stack loss) and $p = 3$ independent variables (air flow, cooling water inlet temperature and acid concentration). The representation of these variables are as given below.

y = the percent of the ingoing ammonia that is lost by escaping in the unabsorbed nitric oxides;

x_1 = air flow (which reflects the rate of operation of the plant);

x_2 = temperature of the cooling water in the coils of the absorbing tower for the nitric oxides;

x_3 = concentration of nitric acid in the absorbing liquid.

This data has great attention in the literature (Balassoriya *et al.*, 1987; Hossain and Naik, 1991; Li *et al.*, 2001; Meloun and Militky, 2001; Hoaglin *et al.*, 2006 and Nurunnabi *et al.*, 2014). All these researchers used this data sets for influence diagnostic purposes in the LM. Thats why we consider this data set for the influence diagnostics purposes.

From Fig. (2.4-2.6), we find that 1st, 2nd, 3rd, 4th, 9th, 17th and 21st observations are considering special attentions. While from Table 2.2, we find that the LM diagnostic declare the 1st, 2nd, 17th and 21st observations are with high leverages. While 1st, 4th and 21st observations are found to be influential observations. From Table 2.3, we study the effect of influential observations on the LM estimates, standard errors of the estimates and their significance. We find that the standard errors of the estimates are reduced by 21% to 23%.

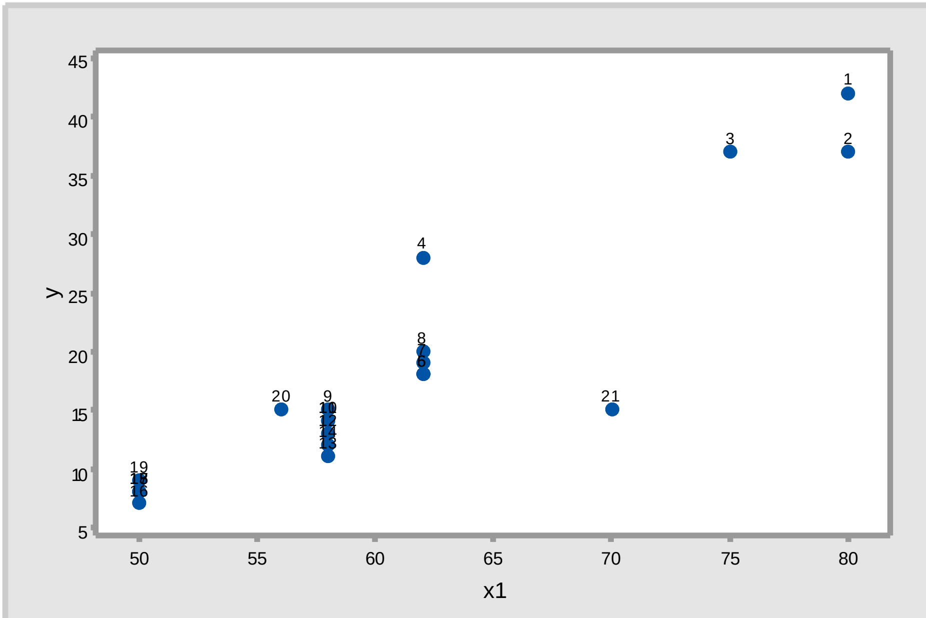


Figure 2.4: Scatter Plot y vs x_1 for Stack Loss Data

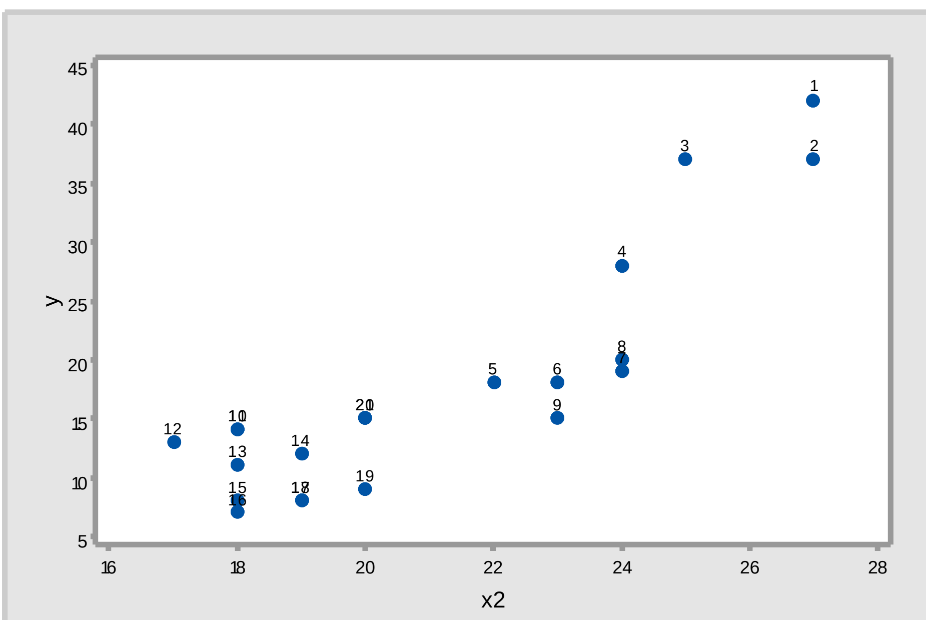


Figure 2.5: Scatter Plot y vs x_2 for Stack Loss Data

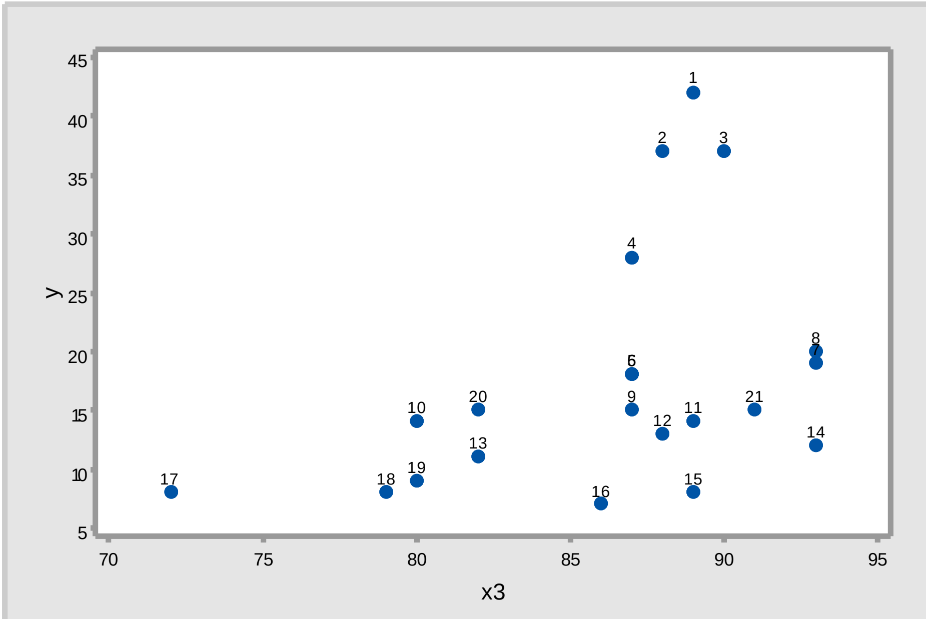


Figure 2.6: Scatter Plot y vs x_3 for Stack Loss Data

Table 2.7: Effect of Influential Observations on the LM Estimates for Stack Loss Data

	Full Data				Without Influential Observations			
	Estimate	S.E	T	p-value	Estimate	S.E	T	p-value
Constant	-39.9197	11.896	-3.3557	0.0038	-40.8578	9.3473	-4.3711	0.0011
x_1	0.7156	0.1349	5.3066	0.0001	0.8662	0.1055	8.2088	0.0000
x_2	1.2953	0.368	3.5196	0.0026	0.6922	0.2928	2.3638	0.0376
x_3	-0.1521	0.1563	-0.9733	0.344	-0.0969	0.1197	-0.8097	0.4353

From Table 2.8, the influential observations also affect the model goodness of fit criterias. The results indicate that after deleting the influential observations, model variance is reduced to 63%. The correlation matrix and VIF are used to identify the multicollinearity in stack loss data. From Table 2.9, we observe that there is an indication of multicollinearity between x_1 and x_2 while VIF shows that there is no indication of multicollinearity on fitting the LM. Now we turn to the influence diagnostics in the GLM by extending the LM influence diagnostics. Before discussing

Table 2.6: The LM Influence Assessment of the Stack Loss Data

Obs.No	$\hat{\epsilon}_i$	$\hat{t}_i$	$\hat{h}_i$	CD_i
1	**3.2346**	1.2095	**0.3016**	**0.1537**
2	-1.9175	-0.7051	**0.3178**	0.0597
3	**4.5555**	1.6179	0.1746	0.1264
4	**5.6978**	2.0518	0.1285	**0.1305**
5	-1.7117	-0.5305	0.0522	0.004
6	**-3.0069**	-0.9632	0.0775	0.0196
7	-2.3895	-0.8259	0.2192	0.0488
8	-1.3895	-0.4737	0.2192	0.0165
9	**-3.1444**	-1.0486	0.1402	0.0446
10	1.2672	0.4262	0.2000	0.0119
11	2.6363	0.8783	0.155	0.0359
12	2.7795	0.9667	0.2172	0.0651
13	-1.4286	-0.4687	0.1575	0.0108
14	-0.0505	-0.017	0.2058	0.0000
15	2.3614	0.8006	0.1905	0.0385
16	0.9051	0.2912	0.1311	0.0034
17	-1.5200	-0.5996	**0.4121**	0.0655
18	-0.4551	-0.1487	0.1606	0.0011
19	-0.5983	-0.1972	0.1745	0.0022
20	1.4121	0.4431	0.0802	0.0045
21	**-7.2377**	**-3.3305**	**0.2845**	**0.692**

Table 2.8: Effect of Influential Observations on the LM Goodness of Fit Criterias for Stack Loss Data

Model Criterias	Full Data	Without Influential Observations
AIC	114.58	68.37
BIC	119.8	71.91
R^2	0.91	0.96
$\hat{\sigma}^2$	10.52	3.91
F-test	179.71	302.91
P-value	0.0000	0.0000

Table 2.9: Correlation Matrix and the VIF for Stack Loss Data

	x_1	x_2	x_3	VIF
x_1	1.00	-0.73	-0.34	2.91
x_1	-0.73	1.00	0.01	2.57
x_1	-0.34	0.01	1.00	1.33

and developing these diagnostics for the GLM, we first provide the overview and estimation methods of the GLM.

2.6 The Generalized Linear Models

The theory of the GLM was introduced by Nelder and Wedderburn (1972), is an extension of the LM or normal linear model under the assumption that the probability distribution of the model error follows some exponentially family of distribution like binomial, Poisson, gamma etc. Later on McCullagh and Nelder (1980) given the extensive formulation for the GLM with different exponential family of distributions. Mathematical form of the GLM is given by

$$y = g(\mu) + \in, \tag{2.11}$$

where y be the response vector, $g(\mu)$ is the mean function vector and $\in$ is the error vector which follows some exponential family of distributions. Generally speaking, the GLM has three components i.e. random component, systematic component and a link function.

Random Component: The random component means that the probability distribution of the dependent variable y vector belongs to an exponential family of distributions with mean vector $\mu = E(y)$

Systematic Component: The systematic component is the linear combination of independent variable and slope coefficients are given as

$$\eta = X\beta,$$

where $\eta = (\eta_1, ..., \eta_n)^T$, $\beta = (\beta_1, ..., \beta_p)^T$ is the $(p' \times 1)$ column vector of unknown parameters and X is the $(n \times p)$ design or model matrix of independent variables of full rank $p < n$.

Link Function: The link function connects the random and systematic component as

$$\eta = g(\mu) \ \ or \ \ \mu = g^{-1}(\eta).$$

The link function g transform to natural parameter is known as canonical link function. The link function actually relates the mean function of the dependent variable y to the linear combinations of the design matrix X.

2.6.1 Assumptions of the Generalized Linear Model

Hardin and Hilbe (2012) had mentioned the following assumptions of the GLM.

- The observations are statistically independent
- The variance function is correctly specified
- The dispersion parameter is correctly specified
- The link function is correctly specified
- The independent variables are in correct form

- There is no undue influence of the individual observation on the fit.

Failure to any of these assumptions may lead to wrong inference to the fitted GLM. Validating these assumptions may be safeguard of false conclusion of the fitted GLM. So, we initially deal with diagnostics of these assumptions and then gave the GLM interpretations about the fitted model.

2.6.2 Formulation of the Generalized Linear Model

One of the important concept for the construction of the GLM theory is the exponential family of distributions. This means that the dependent variable y assumes to follow one of the exponential family of distributions as already mentioned. The exponential form of the exponential family of distributions can be written as

$$f\left(y_i;\theta_i,\phi\right)=exp\left[\frac{y_i\theta_i-b\left(\theta_i\right)}{\phi}+c\left(y_i,\phi\right)\right],\ i=1,...,n, \tag{2.12}$$

where $b\left(\cdot\right)$ and $c\left(\cdot\right)$ are the specific functions for each exponential family of distribution. The detail may be seen in McCullagh and Nelder (1980). The parameter θ is the natural/canonical or location parameter and ϕ is the dispersion parameter and is equivalent to σ^2. The behavior of canonical parameter θ is not similar as of dispersion parameter ϕ because ϕ does not affect the mean function in which we are interested. That is why these exponential family of distributions are known as one-parameter exponential family of distribution. Assume that this dispersion parameter is fixed and unknown. When ϕ in Eq.(2.12) is unknown, then density may or may not be one parameter exponential family. In the GLM, researcher wants to estimate the mean function which needs the parameter θ. The estimation of θ is now discussed in subsequent section.

2.7 Estimation of the Generalized Linear Models

There are various methods for the estimation of unknown parameters in the GLM. Generally, Fisher scoring with Newton-Raphson methods are used to estimate the GLM for the exponential family with single parameter (McCullagh and Nelder, 1980; Hardin and Hilbe, 2012). The simplified form of these methods now widely used method is the iterative reweighted least squares (IRLS) estimation method.

Assume that the dependent variable y measured independently and identically from exponential family of distribution. For the GLM cases, these exponential families include binomial, Poisson, gamma, inverse Gaussian, etc.

The log likelihood of Eq.(2.12) for the ith observation may be written as

$$l_i = \sum_{i=1}^{n} \left\{ \frac{y_i \theta_i - b(\theta_i)}{\phi} + c(y_i, \phi) \right\}. \tag{2.13}$$

Here our interest in the estimation of θ_i. For this purpose, differentiate Eq.(2.13) with respect to θ_i and equating zero, then the expectation may result as

$$\frac{\partial l_i}{\partial \theta_i} = 0$$

.

$$\frac{E(y_i) - b'(\theta_i)}{\phi} = 0. \tag{2.14}$$

From Eq. (2.14), the mean function of the dependent variable simplified as

$$\mu_i = E\left(y_i\right) = b'\left(\theta_i\right) \ \ \& \ \phi \neq 0. \tag{2.15}$$

Here prime indicated the derivatives. Similarly, the variance of the dependent variable can be computed as

$$E\left\{\frac{\partial^2 l_i}{\partial \theta_i^2} + \left(\frac{\partial l_i}{\partial \theta_i}\right)^2\right\} = 0. \tag{2.16}$$

Eq.(2.16) can be written as

$$-\frac{b''\left(\theta_i\right)}{\phi} + \frac{1}{\phi^2} E\{y_i - b'\left(\theta_i\right)\}^2 = 0.$$

or

$$-\frac{b''\left(\theta_i\right)}{\phi} + \frac{1}{\phi^2} V\left(y_i\right) = 0.$$

or

$$V\left(y_i\right) = \phi b''\left(\theta_i\right) = \phi \, diag\left(V\left(\mu_i\right)\right). \tag{2.17}$$

This indicates that variance of the dependent variable is the combination of two functions, i.e. the variance function and the dispersion parameter. The GLMs may be with underdispersion or with overdispersion model because of the GLM dispersion parameter . Mostly the GLM are underdispersion models. While overdispersion arises, when deviance residuals are larger than its degrees of freedom.

2.7.1 Link Functions

In the GLM formulation, link function plays the central role in identifying the GLM structure to define actually what type of the GLM for which exponential family of probability distribution is?. The link function also provides the GLM linearization for specific exponential family of distribution. Linearization in the GLM means that the transformation of the dependent variable in such a way that the resultant model becomes linear. Generally, the link function is given as

$$g(\mu) = \eta = X\beta,$$

where $g(\mu)$ denotes the link function. The link function assumes to be differentiable and monotonic. It is important to note that transformation for linearity is made in the mean of the dependent variable to the independent variables but not transforming the original dependent variable (Gbur *et al.*, 2012). Generally, transforming the dependent variable is to achieve the assumptions of normality and variance stabilization. While in the GLM with link function, it is not necessarily provide the transformation to approach normality and constant variance. Jiao and Chen (2004) discussed that the choice of link functions does not affect the distributional assumption of y. Misspecification of the link function creates bias in the GLM estimates and in predicted variable(s). These link functions include identity, log, reciprocal and power (Czado and Raftery, 2006). The structure of these link functions are given in Table 2.10. Here, we focused on the inverse Gaussian and the gamma cases.

Table 2.10: Types of the GLMs link functions with mean and variance functions

Sr.No	Name	Link Function	Mean Function	Variance Function
1	Linear	$\eta = \mu$	$\mu = \eta$	$V(\mu) = 1$
2	Logrithm	$\eta = \log \mu$	$\mu = e^{\eta}$	$V(\mu) = \mu$
3	Logit	$\eta = \frac{\log \mu}{1-\mu}$	$\mu = \frac{e^{\eta}}{1+e^{\eta}}$	$V(\mu) = \mu(1-\mu)$
4	Reciprocal	$\eta = \frac{1}{\mu}$	$\mu = \frac{1}{\eta}$	$V(\mu) = \mu^2$
5	Squared Reciprocal	$\eta = \frac{1}{\mu^2}$	$\mu = \frac{1}{\sqrt{\eta}}$	$V(\mu) = \mu^3$

For the gamma case, the link function may either identity, log or reciprocal (Halekoh and Hjsgaard, 2007).

2.7.2 The MLE and IRLS

As our interest is in estimating the parameter vector $\beta = (\beta_1, ..., \beta_p)^T$. As before we have seen, according to chain rule the estimation of β is made by taking the first derivative of the log likelihood Eq.(2.12) with respect to β and then equating to zero as

$$U(\beta_j) = \frac{\partial l_i}{\partial \beta_j} = \sum_{i=1}^{n} \left(\frac{\partial l_i}{\partial \theta_i}\right) \left(\frac{\partial \theta_i}{\partial \mu_i}\right) \left(\frac{\partial \mu_i}{\partial \eta_i}\right) \left(\frac{\partial \eta_i}{\partial \beta_j}\right), \; i = 1, ..., n; \; j = 1, ..., p'. \quad (2.18)$$

Eq. (2.18) further be simplified as

$$\frac{\partial l_i}{\partial \beta_j} = \sum_{i=1}^{n} \left(\frac{y_i - b'(\theta_i)}{\phi}\right) \left(\frac{1}{V(\mu_i)}\right) \left(\frac{\partial \mu_i}{\partial \eta_i}\right) x_{ji}, \; . \quad (2.19)$$

Let V represents the diagonal matrix of the observations variances, D represents the diagonal matrix with elements $\frac{\partial \mu_i}{\partial \eta_i}, i = 1, 2, ..., n$, and then likelihood Eq. (2.19) may be written as

$$U(\beta) = X^T D V^{-1} (X - \mu) = 0, \tag{2.20}$$

where

$$V = \begin{bmatrix} v(\mu_1) & 0 & .. & 0 \\ 0 & v(\mu_2) & .. & 0 \\ 0 & 0 & . & 0 \\ 0 & 0 & .. & v(\mu_n) \end{bmatrix} = diag\left(v\left(\mu_i\right)\right).$$

$$D = \begin{bmatrix} \frac{\partial\mu_1}{\partial\eta_1} & 0 & .. & 0 \\ 0 & \frac{\partial\mu_2}{\partial\eta_2} & .. & 0 \\ 0 & 0 & . & 0 \\ 0 & 0 & .. & \frac{\partial\mu_n}{\partial\eta_n} \end{bmatrix} = diag\left(\frac{\partial\mu_i}{\partial\eta_i}\right).$$

As β does not involve in Eq. (2.20) while $\mu = g^{-1}(X\beta)$. Since, these equations are non-linear in β, so we need the iterative methods for the solution of Eq. (2.20) to estimate the unknown parameter β. These methods need starting values to run the algorithm and mostly the started values are set at β_0.

One of the important feature of the GLM is that MLE and IRLS having same algorithm for the estimation of unknown parameter. Now we discuss the algorithm for the estimation process. For this algorithm, we need starting values we called here the trial estimate. Green (1984) discussed that the starting values of the parameter estimates are not generally critical. Different authors used some formulae for starting values for the estimates according to the different GLM structure. For linear regression, with identity link function, starting values can be obtained by simply regressing y on X. Jorgensen (1983), suggests that some modifications are needed for the non-linear models to determine the starting values of the parameter estimates. Using the trial estimated $\hat{\beta}$, we then compute the estimated linear predictor vector as $\hat{\eta} = X\hat{\beta}$. Using this linear estimated predictor vector, the estimated predicted vector is defined as $\hat{\mu} = g^{-1}(\hat{\eta})$. Using these estimated vectors, the adjusted (working) dependent variable can be computed using the following relation

$$y^* = \hat{\eta} + (y - \hat{\mu})\frac{d\eta}{d\mu}, \tag{2.21}$$

where d is the derivative of the link function which is evaluated at trial estimate β_0. Then the iterative weight can be estimated by using the formula as given by

$$W = diag\left[\frac{1}{b''(\theta)}\left(\frac{d\mu}{d\eta}\right)^2\right], \tag{2.22}$$

where $b''(\theta)$ is the second derivative of $b(\theta)$ and also evaluated at the trial estimate. This weight has inverse relation to the variance of the adjusted dependent variable y^* given that the current estimate, with proportionality factor ϕ. Note the functional form of $b(\theta)$, η, $\mu\, and\, V(\mu)$ are different due to exponential family of distributions as given in Table 2.1. Let $\hat{\beta}$ be the estimated value of β as obtained by methods as we discussed above. Finally, the improved estimate by regressing the y^* on X can be obtained by weighted least square as

$$\hat{\beta} = \left(X^T\hat{W}X\right)^{-1}X^T\hat{W}y^*. \tag{2.23}$$

This procedure is repeated until the improved estimates change by less than specific small quantity. Some name the later sentence as this procedure is repeated until the convergence theory achieved or the process is repeated until the difference between trial and improved estimates is smaller than tolerance. McCullagh and Nelder (1989) had shown that this algorithm is similar to Fisher scoring by MLE method. Also, it is important to note that the vector y^* and the matrix W which are used in Eq. (2.23) are actually the values which are produced at convergence. As in this algorithm, the weights are recalculated at each step of the iteration so is called the IRLS algorithm. The sampling distribution of $\hat{\beta}$ is asymptotically multivariate normal with mean β and variance-covariance matrix I^{-1}, where $I = \frac{1}{\phi} X^T \hat{W} X$. For some of the exponential family of distributions, the dispersion parameter is known i.e. binomial and Poisson while for some others, the dispersion parameter is to be estimated from data i.e. gamma, inverse Gaussian etc.

$$Var\left(\hat{\beta}\right) = \hat{\phi}\left(X^T \hat{W} X\right)^{-1}, \tag{2.24}$$

where $\hat{\phi}$ is the estimated dispersion parameter and we discuss this in detail in subsequent section.

2.7.3 Estimation of Dispersion Parameter

The dispersion parameter is very useful in model inferences and diagnostics. Generally, for some of the GLM cases with continuous distributions, the dispersion parameter is unknown and estimated through sample information. McCullagh and

Nelder (1989) discussed two methods to estimate the dispersion in the GLM cases, one is with deviance function and second is with Pearson function. With deviance function, the dispersion parameter is computed as

$$\hat{\phi}_d = \frac{d\left(y; \hat{\mu}\right)}{n - p'}, \tag{2.25}$$

where $d\left(y; \hat{\mu}\right)$ is the deviance function for certain exponential family of distribution. Generally, deviance function is defined by

$$d\left(y; \mu\right) = -2 \int_{y}^{\mu} \frac{y - \mu}{V\left(\mu\right)} d\mu.$$

In the similar way with Pearson function, the estimated dispersion parameter is computed as

$$\hat{\phi}_p = \hat{\phi} = \frac{1}{n - p'} \sum_{i=1}^{n} \frac{\left(y_i - \hat{\mu}_i\right)^2}{V\left(\hat{\mu}_i\right)} = \frac{\chi^2}{n - p'}, \tag{2.26}$$

where $\chi^2 = \sum_{i=1}^{n} \frac{(y_i - \hat{\mu}_i)^2}{V(\hat{\mu}_i)}$ be the chi-square statistic. In the literature, the second choice for the estimation of dispersion is mostly preferred. Because Pearson based computed dispersion estimator has good asymptotic properties and more robust than other estimators. Another advantage of this estimator is that it is less sensitive to extreme observation because of estimated values (Ricci and Martinez, 2008). So, we use the Pearson method for the estimation of dispersion parameter. If the value of dispersion parameter is larger than anticipated for specific exponential family of distribution, it is referred as the GLM with over dispersion and also under dispersion for smaller than anticipated (Cordeiro *et al.*, 2006).

2.8 The Hat matrix and Leverages in the Generalized Linear Model

The hat matrix in GLM is a quantity which is used in influence diagnostics has similar pattern as the LM hat matrix except the formulation. Because the Hat matrix in the GLM is actually based on the final iteration of IRLS method. Another significance of the GLM hat matrix as the LM, hat matrix is that in the LM hat matrix depends only on the design matrix of the independent variables while in the GLM, hat matrix depends on the independent as well as dependent variable. Because the GLM hat matrix is based on IRLS the iterative process. As the GLM Hat matrix is defined as

$$H = \hat{W}^{\frac{1}{2}} X \left(X^T \hat{W} X \right)^{-1} X^T \hat{W}^{\frac{1}{2}}, \tag{2.27}$$

where $\hat{W}$ is the estimated weight of the GLM at the final iteration. The order of H is $n \times n$ and is also called projection matrix. H is idempotent and symmetric. The leverages in the GLM are defined by

$$h_{ii} = \hat{w}_i x_i^T \left(X^T \hat{W} X \right)^{-1} x_i = diag\left(H\right) = \frac{\partial \hat{\mu}_i}{\partial y_i}. \tag{2.28}$$

Here $\sum_{i=1}^{n} h_{ii} = p'$ and $0 \leqslant h_{ii} \leqslant 1$. Any observation with leverage value close to zero indicated that the observation is not influential on the fitted value. While any observation with leverage value close to one declared the observation is influential on the fitted value. Observation with large hat values apart from the rest values may or may not be influential such observation sometimes known as high leverage point. Leverages are also used as a influence diagnostic tool in the GLM. For the ith case, the observation with hat value larger than $\frac{2p'}{n}$ declared as the influential observation. The index plot of the leverages are also help full in diagnosing of influential observation. Collet (1991) also suggests that plot absolute GLM standardized residuals against the leverages give more informative results about outlier as well as influential observation. Note that extreme values in the X does not necessary have larger influence due to small weight because weight involve in the estimation of hat matrix (McCullagh and Nelder, 1989).

2.9 The GLM Residuals

In the LM, raw residuals and its other forms are used for testing regression model diagnostics. While in the GLM, these raw residuals are inappropriate to study the GLM diagnostics because $V(y_i)$ is not constant (Myers *et al.*, 2010). The GLM residuals are in various forms, which are available in the literature. These forms of the GLM residuals include the Pearson residuals, deviance residuals, likelihood residuals, Anscombe residuals and working residuals. In this section, we also discuss the standardized residuals and in Section 2.10 and in Chapter 3 and Chapter 4, we discuss adjusted forms of the GLM residuals.

2.9.1 The Pearson Residuals in the GLM

The Pearson residuals in the GLM are defined as

$$r_{Pi} = \frac{y_i - \hat{\mu}_i}{\sqrt{Var\left(\hat{\mu}_i\right)}}, \tag{2.29}$$

where $\hat{\mu}_i$ is mean function which has different functional form according to exponential family of distributions as indicated in Table 2.1. For large samples, r_{Pi} are distributed as approximately normal with constant variance (Olsson, 2002). Cordeiro (2004) has pointed out that r_{Pi} is proportional to $\sqrt{\phi}$. These residuals are used to measure the quality of well fitted model. Similarly the standardized Pearson residuals are computed as

$$sr_{Pi} = \frac{r_{Pi}}{\sqrt{\hat{\phi}\left(m_{ii}\right)}}, \tag{2.30}$$

where h_{ii} is the i-th diagonal elements of the hat matrix as indicated in Eq. (2.28). The GLM residuals can give misleading results due to skewed behaviors (Gill, 2001). Large values of these residuals indicating model failures. Outliers are also identified by plotting these residuals verses by observation numbers (Hardin and Hilbe, 2012). These residuals are also used in testing the homogeneity of fitted model errors (Jiao and Chen, 2004).

2.9.2 The Deviance Residuals

The deviance residuals for the GLM are defined by

$$r_{di} = sign\left(y_i - \hat{\mu}_i\right)\sqrt{|d_i|}, \tag{2.31}$$

where $d_i = 2\left[l\left(\tilde{\eta}, y\right) - l\left(\eta, y\right)\right]$, where also $l\left(\tilde{\eta}, y\right)$ be the likelihood with saturated model and $l\left(\eta, y\right)$ be the likelihood of unsaturated model. For each family of exponential distribution, the function d_i is different and computational form may be seen in Hardin and Hilbe (2012). The standardized deviance residuals are defined by

$$sr_{di} = \frac{r_{di}}{\sqrt{\hat{\phi}\left(m_{ii}\right)}}. \tag{2.32}$$

These residuals are latterly used for the influence diagnostics. Lee (1987) says that the influential observation may not be detected by the deviance residuals due to two reasons. One is there is situation that observation with small residual and second is that large residual may having little effect on the fit.

2.9.3 The Likelihood Residuals

The likelihood residuals are the weighted average of the standardized deviance residuals and the standardized Pearson residuals as we defined in Eq. (2.30) and Eq. (2.32). According to Fox (2002), these residuals mathematically computed as

$$r_{li} = sign\left(y_i - \hat{\mu}_i\right)\sqrt{h_{ii}(sr_{pi})^2 + \left(m_{ii}\right)\left(sr_{di}\right)^2}, \tag{2.33}$$

where sr_{pi} and sr_{di} are the standardized Pearson and the standardized deviance residuals which are defined in Eq. (2.30) and Eq. (2.32) respectively. The detail of theses residuals may be seen in McCullagh and Nelder (1989). When h_{ii} becomes smaller r_{li} and sr_{di} give approximately identical results (Fox, 2002).

2.9.4 The Anscombe Residuals

All the other GLM residuals are not closely related to normal distribution. An alternative form of these GLM residuals, we have which is known to be Anscombe residuals and is given by Anscombe (1953) and are defined as

$$r_{ai} = \frac{A\left(y_i\right) - A\left(\hat{\mu}_i\right)}{A'\left(\hat{\mu}_i\right)\sqrt{Var\left(\hat{\mu}_i\right)}}, \tag{2.34}$$

where the function $A\left(.\right)$ is selected on the basis of exponential family of distributions. Ancombe residuals are used to test the normality of the model (Jiao and Chen, 2004).

2.9.5 The Working Residuals

This type of the GLM residuals by Hardin and Hilbe (2012) are defined as

$$r_{wi} = \left(y_i - \hat{\mu}_i\right)\frac{\partial \eta}{\partial \mu}. \tag{2.35}$$

This form of the GLM residuals are used in convergence theory and in well fitting of the statistical model (Gill, 2001). Similarly the standardized working residuals are defined by

$$sr_{Wi} = \frac{r_{Wi}}{\sqrt{\hat{\phi} m_{ii}}}. \tag{2.36}$$

All these forms of GLM residuals will be compared later with respect to their performance.

2.10 Comparison of the GLM Residuals

McCullagh and Nelder (1983) have shown that Anscombe and adjusted deviance residuals are in similar pattern in terms of approximate normality. Pierce and Schafer (1986) have shown that deviance residual are better than Pearson for testing the goodness of fit test and model diagnostics because of distributional properties. The GLM residuals analysis are mostly based on Pearson and deviance residuals (McCullagh and Nelder, 1989). Collett (1991) argued that either deviance or likelihood residuals are used for model assessment. Cameron and Trivedi (1998) compare Pearson, deviance and anscombe residuals in Poisson regression and found that deviance and anscombe residuals are approximately equal while Pearson residuals are significant as other these two. Olsson (2002) has shown that deviance residuals may be preferred over the other GLM residuals for model diagnostics. Cordeiro (2004) has proved that adjusted Pearson residuals are good because this approaches normality with approximately zero mean and unit variance. Simas and Cordeiro (2009) argued that for exponential family of non-linear models, adjusted Pearson residuals are better than unadjusted Pearson residuals.

2.11 Influence Diagnostics in the GLM

In this section, we explored how the ith observation influenced the GLM parameter estimates and model inferences. How do we diagnose these influential observations in the GLM cases?. Various measures are available in the literature of the LM. Some of these measures are extended to the GLM cases for example William (1987). Jearpaporn *et al.* (2004) say that GLMs are still to be sensitive to the influential observations. Here we extend some more of these measures to the GLM cases for the identification influential observations. These influence measures include leverages, Cooks distance, Modified Cooks distance, Adrews Pregibon, Welsch distance, Difference of Fit, Covariance ratio and Hadis measure. According to the different GLM residual structures, we use all these residuals in all mentioned influence measures. Because Cordeiro (2004) says that there is a need to study the asymptotic properties of the GLM residuals for diagnostic purposes. Thats our one research objective is to compare the GLM residuals in influence diagnostics.

2.11.1 Cook's Distance

Cook's distance (CD_i) was first proposed by Cook (1977) for the LM to diagnose the influential observation. CD_i measures the overall change in the fitted model, when ith observation is deleted from the model. For the GLM case, CD_i formulated as

$$CD_i = \frac{\left(\hat{\beta} - \hat{\beta}_{(i)}\right)^T \left(V\left(\hat{\beta}\right)\right)^{-1} \left(\hat{\beta} - \hat{\beta}_{(i)}\right)}{p'} = \frac{\left(\hat{\beta} - \hat{\beta}_{(i)}\right)^T X^T \hat{W} X \left(\hat{\beta} - \hat{\beta}_{(i)}\right)}{p'\hat{\phi}}, \tag{2.37}$$

where the subscript (i) indicate the results without or after deleting the ith observation. After simplification, Eq. (2.37) now becomes

$$CD_i = \frac{r_{Pi}^2}{p'\hat{\phi}} \frac{h_{ii}}{(1 - h_{ii})^2}.$$

This relation further simplified to be as

$$CD_i = \frac{sr_{Pi}^2}{p'} \frac{h_{ii}}{(1 - h_{ii})}, \quad i = 1, 2, ..., n, \tag{2.38}$$

where sr_{Pi} are the standardized Pearson residuals as defined in Eq.(2.30). This formulation of Cook's statistic depends on three quantities, standardized Pearson residuals, leverages and the number of unknown parameters. This diagnostic is used to measure the effect of influential observation on $\hat{\beta}$ only (Ullah and Pasha, 2009). A large value of CD_i indicates the i^{th} observation is influential. Cook (1977) suggests that the influential observation is also detected by using the cut point i.e. $CD_i \geqslant F_\alpha, (p', n - p')$. This cut point is unable to detect influential observation in the GLM cases. Alternative cut point for the detection of influential observation in the GLM is $\frac{4}{n-1}$ given by Hardin and Hilbe (2012).

2.11.2 Modified Cook's Distance

CD_i challenged by Welsch and Kuh (1977), so according to Belsley *et al.* (2004) modified Cook's distance to diagnose the influential observation more sharply and is computed as for the GLM as

$$MCD_i = \left[\frac{\left(\hat{\beta} - \hat{\beta}_{(i)}\right)^T X^T W X \left(\hat{\beta} - \hat{\beta}_{(i)}\right)}{\frac{p'}{n-p'} \hat{\phi}_{(i)}} \right]^{\frac{1}{2}}. \tag{2.39}$$

Eq.(2.39) can be further simplified as

$$MCD_i = \left[\frac{n-p'}{p'} \frac{h_{ii}}{(1-h_{ii})^2} \frac{r_{Pi}^2}{\hat{\phi}_{(i)}} \right]^{\frac{1}{2}}, \tag{2.40}$$

where $\hat{\phi}_{(i)} = \left(\frac{n-p-sr_{Pi}^2}{n-p'}\right)\hat{\phi}$, According to SMW theorem, Eq. (2.40) can also be written as in terms of standardized Pearson residuals

$$MCD_i = \left[\frac{n-p'}{p'} \frac{h_{ii}}{1-h_{ii}} \right]^{\frac{1}{2}} |t_i|, \tag{2.41}$$

where $t_i = sr_{Pi}\sqrt{\frac{n-p'}{n-p-sr_{Pi}^2}}$. The value of t_i is also known as deleted studentized residuals or externally studentized residuals or standardized residuals (Fox, 1999). The observation is declared as influential observation, if $MCD_i \geqslant 2\sqrt{\frac{n-p'}{n}}$.

2.11.3 Andrew's Pregibon Measure

Another influence measure for the LM, which was given by Andrews and Pregibon (1978) and for the GLM, it is reformulated as

$$AP_i = 1 - h_{ii} - \frac{r_{Pi}^2}{\sum r_{Pi}^2} or AP_i = m_{ii} - \frac{r_{Pi}^2}{\sum r_{Pi}^2}. \tag{2.42}$$

AP_i for each case is lies between zero and one inclusive and small values of AP_i indicated that the observation is substantial and this observation may or may not be influential (Gray, 1989). An observation is declared as the influential observation, if $(1 - AP_i)$ crosses $\frac{2p'}{n}$. Note that this statistic provides only significant results, if we have less than 30 data values (Belsley *et al.*, 2004).

2.11.4 Covariance Ratio

DFFITS is used to measure the effect of ith influential observation on the estimated value and fitted values. This statistic have no information regarding the overall precision of the estimation. To measure such influence, we have another technique measuring the effect of ith influential observation on the overall precision of the estimation, this technique is known as covariance ratio (CVR_i). CVR_i is used to measure the ith observation influence on the variances of the estimates (Ullah and Pasha, 2009) and is defined for the GLM as

$$CVR_i = \frac{\left| MSE_{(i)} \left(X^T{}_{(i)} W_{(ii)} X_{(i)} \right)^{-1} \right|}{\left| (X^T W X)^{-1} \right|}. \tag{2.43}$$

According to SMW theorem, CVR_i by Eq.(2.43), can be simplified to be as

$$CVR_i = \frac{\left[\frac{\left(n - p - sr^2_{Pi} \right)}{n - p'} \right]^{p'}}{m_{ii}}. \tag{2.44}$$

If $CVR_i > 1$, this indicated that the ith observation improve the precision otherwise the inclusion of ith observation degrades the precision of estimation. Belsley *et al.*

(2004) has pointed out that the ith observation is influential if $CVR_i > 1 + 3\frac{p'}{n}$ or if $CVR_i < 1 - 3\frac{p'}{n}$.

2.11.5 Difference of Fit Test (DFFITS)

DFFITS is also a measure of influence diagnostic and is defined as the scaled difference between the fitted value of complete data and fitted value after deleting ith observation. For the GLM, DFFITS is mathematically defined by the following relation

$$DFFITS_i = \frac{\hat{\mu}_i - \hat{\mu}_{i(i)}}{\sqrt{\hat{\phi}_{(i)} h_{ii}}}, \tag{2.45}$$

where $\hat{\mu}$ is the predicted dependent variable and $\hat{\mu}_{(i)}$ is the predicted dependent variable after deleting ith observation. Eq.(2.45) can also be written as

$$DFFITS_i = \frac{W_{ii}^{\frac{1}{2}} x_i^T \left(\hat{\beta} - \hat{\beta}_{(i)}\right)}{\sqrt{\hat{\phi}_{(i)} h_{ii}}}. \tag{2.46}$$

By using SMW theorem, Eq. (2.46) is retransformed as

$$DFFITS_i = |t_i| \sqrt{\frac{h_{ii}}{m_{ii}}}, \tag{2.47}$$

where $t_i = sr_{Pi}\sqrt{\frac{n-p'}{n-p-sr_{Pi}^2}}$ and $sr_{Pi} = \frac{r_{Pi}}{\sqrt{\hat{\phi}(m_{ii})}}$. DFFITS value larger than one indicated that the observation is influential under the assumption, if the data is small (Chatterjee and Hadi, 1986). For large data sets, if the value of DFFITS crossing $2\sqrt{\frac{p'}{n}}$ shows the observation is influential (Belsely *et al.*, 2004).

2.11.6 Welsch's Distance

Welsch (1982) modify the DFFITS (2.47) to welsch distance (WD) as

$$WD_i = DFFITS_i \sqrt{\frac{n-1}{m_{ii}}}. \tag{2.48}$$

For the GLM, case we modify WD by Eq. (2.48) as

$$WD_i = (n-1)\, t_i^2 \frac{h_{ii}}{m_{ii}^2}. \tag{2.49}$$

The observation is declared as the influential observation, if $WD_i > 3\sqrt{p'}$.

2.11.7 Hadi's Measure

Hadi (1992) proposed a new measure for the detection of influential observation in the LM as

$$H_{di} = \frac{p+1}{m_{ii}} \frac{d_i^2}{1-d_i^2} + \frac{h_{ii}}{m_{ii}},\ i = 1, 2, ..., n.. \tag{2.50}$$

where $d_i = \frac{e_i^2}{\sum e_i^2}$. For the GLM with the Pearson residuals in Eq.(2.50), replace d_i as $d_i = \frac{sr_{Pi}^2}{\sum sr_{Pi}^2}$. On the basis of H_d measure, Hadi (1992) has pointed out that the ith observation is influential observation, if $H_{di} > Mean\,(H_{di}) + c\sqrt{Var\,(H_{di})}$, where c may be 2 or 3.

2.12 Influence Measures Yard Sticks

Different authors give various yard sticks or cut points of various influence measures for the detection of influential observation. Here we summarize these cut-off values for our proposed influence diagnostics in the GLM.

Leverage: Hoaglin and Welsch (1978) proposed the cut-off value for h_{ii} that is $h_{ii} > \frac{2p'}{n}$. Velleman and Welsh (1981) suggests that the cut-off point for the leverage is $h_{ii} > \frac{3p'}{n}$, when $p' > 6$ and $(n - p') > 12$. Huber (1981) proposed that the cut-off value for the leverages is $h_{ii} > 0.2$.

CD: Cook (1977) suggests that the value of CD for any case greater than 50% percentile of the F-distribution with p' and $n - p'$ degree of freedom, declared the observation as influential. In most of the regression cases, this cutoff point failed to detect influential observations (Hossain and Naik, 1991). Fox (2002) says for the simple regression model that the ith case is influential, if the value of CD crosses $\frac{4}{n-2}$. Weisberg (2005) says that for the ith observation, the value of Cook's distance greater than one indicated as influential observation. Hardin and Hilbe (2012) recommended that the yard stick of CD in the GLM influence diagnostic is to be $\frac{4}{n}$.

MCD: It is suggested that, if n is large, then the cutoff value for the identification of influential observation by MCD method is 2. An other cutoff point for MCD for detection of influential observation is that the ith observation is influential observation, if $MCD_i \geqslant 2\sqrt{\frac{n-p'}{n}}$.

AP: The cutoff point to detect the influential observation in the literature is that the ith observation is influential observation, if $(1 - AP_i) > \frac{2p'}{n}$. But for the GLM cases, we propose that the ith observation is declared as influential observation, if $|1 - AP_i| > \frac{2p'}{n}$.

CVR: Several researchers say that for the ith observation, if the CVR crosses one then observation has special attention. Belsley *et al.* (1980) recommended that the cutoff point of the CVR method is, if $|CVR_i - 1| \geqslant \frac{3p'}{n}$, then the observation is the influential observation.

DFFITS: General yard stick for the identification of influential observation of the DFFITS method is 2. While Chatterjee and Hadi (2012) say that the yard stick for the DFFITS is, if $|DFFITS_i| > 2\sqrt{\frac{p'}{n}}$, then the ith observation is the influential observation.

WD: Welsch (1982) proposed the yard stick for the WD that is the observation is influential observation, if $WD_i \geqslant 3\sqrt{p'}$.

Hd: Hadi (1992) proposed the cutoff point for his new developed diagnostic. He said that the observation has great attention, if for the$i\hat{}th$ observation, $H_d > Mean\,(H_d) + c\sqrt{Var\,(H_d)}$.

2.13 Algorithm (Steps)

Step 1: Start $\beta_0 = \bar{y}$ from the mean model $y_i = g\,(\mu_i) + \in_i$.

Step 2: Using β_0 to compute n residuals.

Step 3: Use sum of square of residuals to fit the variance model with independent variables and link function through IRLS.

Step 4: Use estimates from step 3 to form weight matrix $W = diag\,[\hat{\mu}_1, \hat{\mu}_2, ..., \hat{\mu}_n]$.

Step 5: Use W with IRLS to update β_0 to β.

Step 6: Go to step 2 with β replacing β_0.

Step 7: Repeat step 2 to 6 until convergence.

Step 8: If Convergence approach, then compute the all GLM residuals and influence diagnostics.

Chapter 3

Influence Diagnostics in Gamma Ridge Regression Model

3.1 Introduction

The Gamma regression model (GRM) is generally applied for positively skewed continuous dependent variable. With the addition that variance of the dependent variable is directly related to the square expectation of the dependent variable. The GRM has given a great attention in the literature i.e. survival models (Myers *et al.*, 2010), Hydrology (Das and Kim, 2012) etc.

When GRM is applied to multicolliear data, then alternative to IRLS are the biased estimation methods and the most popular is the ridge estimator. The problems of multicollinearity and influential observation may occur simultaneously and lead the GRM estimation and model inferences. The influence diagnostics with bias estimator has great attention in the literature examples include (Walker and Birch, 1988; Billor, 1999; Jahufer and Jianbao, 2009; Jahufer and Chen, 2011, 2012; Jahufer, 2013; Ullah *et al.*, 2013). All these researchers study the influence diagnostics with biased estimators related to the LM only. There is no work available in the literature related to the GRM influence diagnostics with ridge estimator.

3.2 The Gamma Regression Model

When the response variable is positively skewed with constant coefficient of variation and mean is proportional to variance, then the GRM is applied for modeling such type of response variable. The applications of such models include in reliability and survival models (Lawless, 2003) and in many industries like semiconductor and chemical industries where production data follow gamma distribution (Jearkpaporn *et al.*, 2005). The GRM is used for predicting the amount of rain fall (Segond *et al.*, 2006) and in hydrogeology to determine the characteristics of ground water quality (Das and Kim, 2012). Different authors study the GRM in different dimensions with different link functions as we have discussed in chapter 2. Ortega *et al.* (2008) proposed deviance residuals for the log-gamma regression model with censored survival time. Ortega *et al.* (2009) also developed the generalized log-gamma regression model with cure fraction. Ortega *et al.* (2012) introduced the log-gamma regression model with censored observations. The gamma based model has some similar and adverse results with other GLMs. Some authors (Atkinson, 1982; Firth, 1988; McCullagh and Nelder, 1989; Das, 2014) say that for positively skewed data gamma and log-normal distribution gives identical analysis. They also show that with non-constant variance, the gamma distribution and log-normal distribution regression estimates may be different. Jiao and Chen (2004) find that for sequential population, the gamma model estimates are better than log-normal model estimates. The GRM is also used to model in production monitoring process (Jearkpaporn *et al.*, 2005). Minoda and Yanagimito (2009) used the GRM with multiple strata to estimate the common slope.

3.3 Derivation of the Gamma Regression Model

Let y_i be the dependent variable which comes from the gamma distribution i.e. $G(\mu, \phi)$ and the probability density function of gamma distribution is defined as (Hardin and Hilbe 2012)

$$f(y/\mu, \phi) = \frac{1}{y\Gamma\left(\frac{1}{\phi}\right)}\left(\frac{y}{\mu\phi}\right)^{\frac{1}{\phi}} e^{-\frac{y}{\mu\phi}} \, ; \, y \geqslant 0, \tag{3.1}$$

where $\mu > 0$ and $\phi > 0$. The mean and variance of y are given by $E(y) = \mu \, and \, Var(y) = \phi\mu^2$.

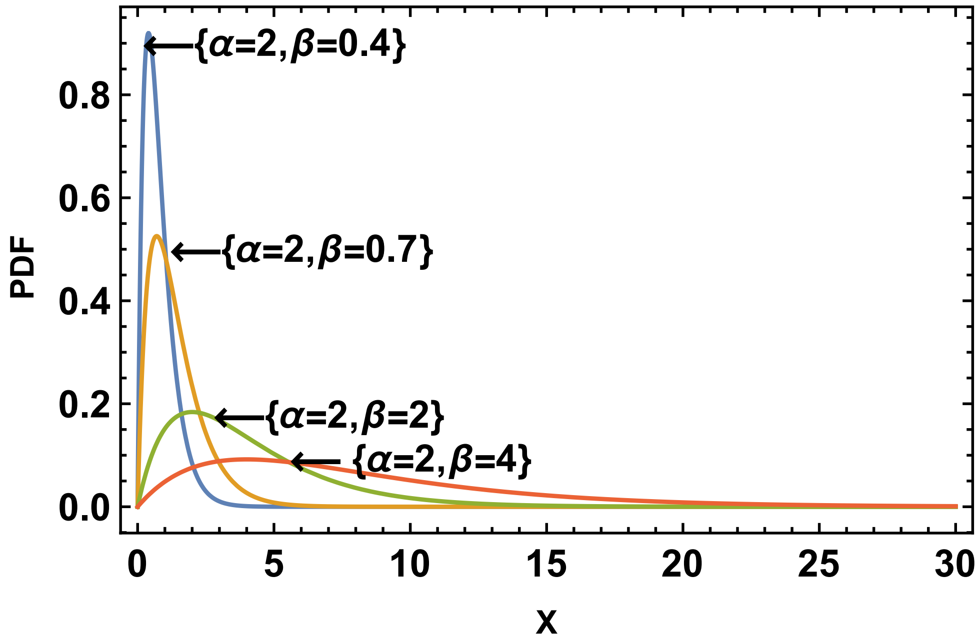


Figure 3.1: Gamma density Curves with several parametric values.

With fixed shape parameter, the variance of the gamma random variable is proportional to the square of its expectation as similar to log-normal and weibull distributions (Tsou, 2011). Eq. (3.1) is the reparametrization of the gamma density with parameters α & β is obtained by setting $\mu = \alpha\beta$ and $\phi = \frac{1}{\beta}$. The parameter β determines the shape of the gamma density. For $0 < \beta < 1$, the gamma density decreases monotonically. If $\beta = 1$, then gamma density is known as the standard gamma distribution. If $\alpha = 1$, then gamma density transformed to exponential density with parameter $\lambda = \frac{1}{\beta}$. Fahrmeir and Tutz (2001) have shown that for different values β the gamma distribution is positively skewed. These relations are also shown in Fig. 3.1. The Gamma distribution transformed to normal distribution as $\beta \to \infty$. The GRM is modeled with reciprocal link function as $\mu_i^{-1} = x_i^T\beta,\ i = 1, ..., n.$ where X is design matrix and β is the vector of slope coefficients including intercept.

3.4 Estimation of the Gamma Regression Model

The GRM is used to model, when the response variable y_i is positively skewed and is distributed as the gamma with parameter μ and ϕ represented as $G(\mu, \phi)$, and X is the $n \times p$ data matrix of standardized p explanatory variables so that $\sum_{i=1}^{n} x_{ij} = 0$ and $\sum_{i=1}^{n} x_{ij}^2 = 1$ β and is vector of slope coefficients.

Let $Z = (1, X)$ be the $n \times p'$ design matrix of full rank $p' < n$ and $E(y_i) = \mu_i$ be the mean, $g(\mu) = \frac{1}{\mu} = \eta$ be the link or canonical function, here $\eta = (\eta_1, ..., \eta_n)^T = z_i^T \beta^* = \sum_{j=1}^{p'} z_{ij} \beta_j^*$ be the linear predictor and β^* are the $n \times p'$ vectors of unknown parameters. The link function of the GRM gives the sufficient statistic with imposed restriction on β^* that $\mu_i > 0, i = 1, ..., n$. The link function which perform the functional relationship between μ and η. The GRM model is estimated by ML estimation method so the log likelihood of the gamma density (3.1) may be written as

$$l(y; \beta^*, \phi) = \frac{y z_i^T \beta^* - \ln\left(z_i^T \beta^*\right)}{-\phi} + \left(\frac{1-\phi}{\phi}\right) \ln y - \frac{1}{\phi} \ln \frac{1}{\phi} - \ln \Gamma \frac{1}{\phi}, \tag{3.2}$$

where ϕ is the dispersion parameter which is assumed to be fixed and unknown and for the GRM $\phi = \frac{1}{\beta}$. The estimates of β^* by ML method equal solution of the system of equations by setting the first derivative of Eq. (3.2) equals to zero, then we have

$$U\left({\beta_j}^*\right) = \frac{\partial l}{\partial {\beta_j}^*} = \frac{1}{\phi} \left(y_i - \frac{1}{z_i^T \beta^*} \right) z_{ij} = 0, \tag{3.3}$$

where $U(\beta^*)$ is the score vector with dimension $p' \times 1$. Since the solution of the system of Equations in Eq. (3.3) is non-linear, so the Newton-Raphson iterative procedures are used to estimate the unknown parameter. For the iterative procedure of the GRM, initial values and full algorithm for the estimation of unknown parameter can be found in Hardin and Hilbe (2012). Let $\beta^{*(m)}$ be the approximated ML value of β^* and by Green (1984) the iterative method gives the relation as

$$\beta^{*(m+1)} = \beta^{*(m)} + \left\{ I\left(\beta^{*(m)}\right) \right\}^{-1} U\left(\beta^{*(m)}\right), \tag{3.4}$$

where $I\left(\beta^{*(m)}\right)$ is the $p' \times p'$ information matrix and both information and score vectors are evaluated at $\beta^{*(m)}$. At convergence Eq.(3.4), the unknown parameter can be evaluated as

$$\hat{\beta}^* = \left(Z^T \hat{W} Z\right)^{-1} Z^T \hat{W} y^*, \tag{3.5}$$

where $y_i^* = \hat{\eta}_i + \frac{y_i - \hat{\mu}_i}{\hat{\mu}_i^2}$ is the adjusted response variable, $\hat{W} = diag\left(\hat{\mu}_1^2, ..., \hat{\mu}_n^2\right)$ is the weighted matrix and $\hat{\mu}_i = \frac{1}{z_i^T \hat{\beta}^*}$ with reciprocal link function. The estimated mean function $\hat{\mu}_i = e^{z_i' \hat{\beta}^*}$ with log link function may also be used for the GRM but our concerns with the reciprocal link. The vector $\hat{\beta}^*$ is estimated by minimizing the weighted sum of square error under the assumption that there is no multicollinearity in the cross-product of $\left(X^T \hat{W} X\right)$ terms. Both y^* and $\hat{W}$ are found by the iterative methods and for the detail derivations and procedure, reader are referred to Hardin and Hilbe (2012). Green (1984) shows that $X^T y$ is sufficient for β due to this for the GRM, Newton Raphson method for the estimation of β will work well with mean function $\mu_i^{-1} = z_i^T \hat{\beta}^*, \ i = 1, 2, ..., n.$

3.5 The GRM Dispersion Parameter

As we have already mentioned in chapter 2, the dispersion parameter is actually computed in two different ways, one is with deviance function and second one is with Pearson statistic. Following Eq. (2.25), the dispersion parameter for the GRM is given by

$$\hat{\phi}_d = \frac{d\left(y;\hat{\mu}\right)}{n-p'}, \tag{3.6}$$

where $d\left(y;\hat{\mu}\right) = 2\sum_{i=1}^{n}\left\{\log\left(\frac{\hat{\mu}_i}{y_i}\right) + \frac{y_i}{\hat{\mu}_i} - 1\right\}$. The dispersion parameter using Pearson chi-square statistic for the GRM is similar as defined in Eq. (2.26) with different functional form of the mean function.

3.6 Hat matrix and Leverages in the GRM

Hat matrix for the GRM is defined by

$$H = \hat{W}^{\frac{1}{2}} Z \left(Z^T \hat{W} Z\right)^{-1} Z^T \hat{W}^{\frac{1}{2}}, \tag{3.7}$$

where $\hat{W} = diag\left(\hat{\mu}_1^2, ..., \hat{\mu}_n^2\right)$ is the weight matrix for the GRM. The leverages are the i-th diagonal elements of the hat matrix H, given in Eq. (3.7), $h_{ii} = diag\left(H\right)$. The usefulness of leverages in the GRM is superior to the LM in detection of influence on the parameter estimates due to the combination of covariates and the response variable. Observations with leverages greater than $\frac{2p'}{n}$ are declared as influential observation. The $h_{ii} \to 0$, indicated that the ith observation has no influence while $h_{ii} \to 1$ shows ith observation has influence on the predicted values.

3.7 The GRM Residuals

Influential diagnostics actually are based on the GLM residuals. Residuals analysis has key role in formulating theories and their validation in regression models. In this section, we are going to present the formulation of all the GRM types of residuals with their standardized form. Latterly, these forms of the GRM residuals are used for studying the influence diagnostics.

3.7.1 The Pearson Residuals for the GRM

It follows from the Eq. (2.29), the Pearson residuals of the GRM are defined by

$$r_{Pi} = \frac{y_i - \mu_i}{\sqrt{V(\mu_i)}} = \frac{y_i - \mu_i}{\mu_i}. \tag{3.8}$$

The standardized Pearson residuals are obtained by dividing Eq. (3.8) by the component $\sqrt{\hat{\phi} m_{ii}}$ as

$$sr_{Pi} = \frac{r_{Pi}}{\sqrt{\hat{\phi} m_{ii}}}. \tag{3.9}$$

3.7.2 The Deviance Residuals for the GRM

The GRM deviance residuals, according to Eq. (2.31), are defined by the relation

$$r_{d_i} = sign\,(y_i - \hat{\mu}_i)\,\sqrt{|d_i|}, \tag{3.10}$$

where $d_i = -2\left\{\ln\left(\frac{y_i}{\hat{\mu}_i}\right) - \frac{y_i - \hat{\mu}_i}{\hat{\mu}_i}\right\} = -2\left\{\ln\left(y_i z_i^T \hat{\beta}^*\right) - \left(y_i z_i^T \hat{\beta}^* - 1\right)\right\}$. The standardized deviance residuals are defined by

$$sr_{di} = \frac{r_{di}}{\sqrt{\hat{\phi} m_{ii}}}. \tag{3.11}$$

These standardized deviance residuals are used in detecting the mean shift signal (Jearkpaporn *et al.*, 2005). They also pointed that when deviance residuals are in control states, then the GRM standardized deviance residuals follow normal distribution. Further these forms of the GRM deviance residuals can be used in measuring the influence diagnostics and other assumptions of the GRM.

3.7.3 The Likelihood Residuals for the GRM

Likelihood residuals in the GRM by following Eq. (2.33) are defined by the relation

$$r_{li} = sign\,(y_i - \hat{\mu}_i)\,\sqrt{h_{ii}(sr_{pi})^2 + (m_{ii})\,(sr_{di})^2}, \tag{3.12}$$

where sr_{pi} and sr_{di} are the standardized Pearson and the standardized deviance residuals as defined in Eq. (3.9) and Eq. (3.11). We use these residuals for the computation of influence diagnostics in the GRM for the detection of influential observations.

3.7.4 The GRM Anscombe Residuals

As we defined the Anscombe residuals for the GLM by Eq. (2.34)

$$r_{ai} = \frac{A(y_i) - A(\hat{\mu}_i)}{A'(\hat{\mu}_i)\sqrt{Var(\hat{\mu}_i)}}. \tag{3.13}$$

For the GRM Eq. (3.13) transformed to as

$$r_{ai} = \frac{3\left(y_i^{\frac{1}{3}} - \hat{\mu}_i^{\frac{1}{3}}\right)}{\hat{\mu}_i^2}, \tag{3.14}$$

where $A(y_i) = 3y_i^{\frac{1}{3}},\ A(\hat{\mu}_i) = 3\hat{\mu}_i^{\frac{1}{3}},\ A'(\hat{\mu}_i) = \hat{\mu}_i^{-\frac{2}{3}}\ \&\ Var(\hat{\mu}_i) = \hat{\mu}_i^2$ and also $\hat{\mu}_i = \frac{1}{z_i^T\hat{\beta}^*}$. The standardized Anscombe residuals now defined by

$$sr_{ai} = \frac{r_{ai}}{\sqrt{\hat{\phi}m_{ii}}}. \tag{3.15}$$

3.7.5 The GRM Working Residuals

This type of the GLM residuals by Eq. (2.35) are defined as

$$r_{Wi} = (y_i - \hat{\mu}_i)\left(\frac{\partial\eta_i}{\partial\mu_i}\right), \tag{3.16}$$

where for the GRM, $\frac{\partial \eta}{\partial \mu} = -\frac{2}{\mu^3}$. So the estimated working residuals for the GRM, Eq.(3.16) becomes

$$r_{Wi} = (y_i - \hat{\mu}_i)\left(-\frac{2}{\hat{\mu}_i^3}\right). \tag{3.17}$$

Similarly, the standardized working residuals are defined by

$$sr_{Wi} = \frac{r_{wi}}{\sqrt{\hat{\phi} m_{ii}}}. \tag{3.18}$$

3.8 The GRM Adjusted Residuals

The adjustment in the GRM residuals are needed to approach the standardized theory with zero mean and unit variance. Cox and Snell (1968) give basic theoretical results for the derivation of the adjusted residuals. Pierce and Schafer (1986) had given the adjusted deviance residuals. Following Cox and Snell's (1968) method, some researchers provide adjusted residuals for the GLM. Cordeiro (2004) give the adjusted Pearson residual, Urbano *et al.* (2012) give the adjusted Wald residuals and Anholeto *et al.* (2014) give the adjusted Pearson residuals for the beta regression models. They all conclude that adjusted residuals are better than the other form of the residuals. In this section, we provide the adjusted residuals for the GRM on the basis of Cox and Snell (1968) theory. Also in this section, we proposed other form of adjustment in the GRM residuals.

3.8.1 The Adjusted Pearson Residuals

Following Cordeiro (2004), we consider the following notations $\mu = \frac{1}{\eta}$, $\mu' = \frac{\partial \mu}{\partial \eta} = -\mu^2$, $\mu'' = \frac{\partial^2 \mu}{\partial \eta^2} = 2\mu^3$, $\mu''' = \frac{\partial^3 \mu}{\partial \eta^3} = -6\mu^4$ and $V = Var\left(\mu\right) = \mu^2$, $V^{(1)} = \frac{\partial V}{\partial \mu} = 2\mu$, $V^{(2)} = \frac{\partial^2 V}{\partial \mu^2} = 2$.

The true standardized Pearson residuals of the GRM are given by

$$R_{Pi} = \phi^{-\frac{1}{2}} \left(Y_i - \mu_i\right) V_i^{-\frac{1}{2}}. \tag{3.19}$$

We define $H_r^{(i)} = \frac{\partial R_{Pi}}{\partial \beta_r^*}$ *and* $H_{rs}^{(i)} = \frac{\partial^2 R_{Pi}}{\partial \beta_r^* \partial \beta_s^*}$, $r = s = 0, 1, ...p$.

Now by following Cox and Snell's (1968) method, the first two moments of the Pearson residuals to order $O\left(n^{-1}\right)$ are

$$E\left(R_{Pi}\right) = \sum B\left(\hat{\beta}_r^*\right) E\left(H_r^{(i)}\right) - \sum K^{rs} E\left(H_r^{(i)} U_s^{(i)} + \frac{1}{2} H_{rs}^{(i)}\right). \tag{3.20}$$

$$V\left(R_{Pi}\right) = 1 + 2\sum B\left(\hat{\beta}_r^*\right) E\left(R_{Pi} H_r^{(i)}\right) - \sum K^{rs} E\left(2R_{Pi} H_r^{(i)} U_s^{(i)} + H_r^{(i)} H_s^{(i)} + R_{Pi} H_{rs}^{(i)}\right). \tag{3.21}$$

$$Cov\left(R_{Pi}, R_{Pj}\right) = -\sum K^{rs} E\left(2R_{Pi} H_r^{(j)} U_s^{(i)} + R_{Pj} H_r^{(i)} U_s^{(j)} + H_r^{(i)} U_s^{(j)}\right), \tag{3.22}$$

where K^{rs} is the (r,s)th elements of the inverse expected information matrix and $B\left(\hat{\beta}_r^*\right)$ is the bias of $\hat{\beta}_r^*$ to order $O\left(n^{-1}\right)$

For the Pearson residuals moments, we obtain $E\left(H_r^{(i)}\right) = -\phi^{-\frac{1}{2}} V_i^{-\frac{1}{2}} \mu_i' z_{ir} = z_{ir}$ and $E\left(H_{rs}^{(i)}\right) = \frac{1}{2}\phi^{-\frac{1}{2}} \left(2V_i^{-\frac{3}{2}} V_i^{(1)} \mu_i'^2 - 2V_i^{-\frac{1}{2}} {\mu''}_i\right) z_{ir} z_{is}$.

According to Cordeiro and McCullagh (1991), $B\left(\hat{\beta}^*\right)$ is defined by

$$B\left(\hat{\beta}^*\right) = -\frac{\phi}{2}\left(Z^TWZ\right)^{-1}Z^TA_dF\mathbf{1},$$

where $W = diag\left(w_i\right) = diag\left(V\left(\mu_i\right)\right) = diag\left(\mu_i^2\right)$, $\mathbf{1}$ is the vector of ones, $F = diag\left(\frac{1}{V_i}\mu'_i\mu''_i\right) = diag\left(-2\mu_i\right) = diag\left(-\frac{2}{z'_i\beta^*}\right)$, $A = \{a_{ij}\} = Z\left(Z^TWZ\right)Z^T$ and $A_d = diag\left(a_{ii}\right)$.

Next, we have

$$\sum_{r=1}^{p} B\left(\hat{\beta}_r\right)E\left(H_r^{(i)}\right) = -\phi^{-\frac{1}{2}}V^{-\frac{1}{2}}\mu_i'\gamma_i^TZB\left(\hat{\beta}\right) = \phi^{-\frac{1}{2}}\mu_i\gamma_i^TZB\left(\hat{\beta}\right),$$

where $\mu_i = \frac{1}{z_i^T\hat{\beta}^*}$ and γ_i is an $n \times 1$ vector with one in the ith position and zeros elsewhere.

Further, we can write

$$E\left(H_r^{(i)}U_s^{(i)}\right) = -\frac{1}{2}\phi^{-\frac{1}{2}}V_i^{-\frac{3}{2}}V_i^{(1)}\mu\prime_i^2 z_{ir}z_{is} = \phi^{-\frac{1}{2}}V_iz_{ir}z_{is}.$$

$$E\left(H_r^{(i)}U_s^{(i)} + \frac{1}{2}H_{rs}^{(i)}\right) = -\frac{1}{2}\phi^{-\frac{1}{2}}V_i^{-\frac{1}{2}}\mu\prime\prime_i z_{ir}z_{is} = \phi^{-\frac{1}{2}}\mu\prime_i z_{ir}z_{is}.$$

and

$$-\sum_{r,s=1}^{p'} K^{rs}E\left(H_r^{(i)}U_s^{(i)} + \frac{1}{2}H_{rs}^{(i)}\right) = -\frac{1}{2}\phi^{-\frac{1}{2}}V^{-\frac{1}{2}}\mu_i''\,z_{ir}z_{is} = \phi^{-\frac{1}{2}}\mu_i'z_{ir}z_{is}$$

. Let $r = (E(R_{Pi}))^T$; for $i = 1, ..., n$ are the expected Pearson residuals of order $O(n^{-1})$, then

$$r = -\frac{\sqrt{\phi}}{2}(I - H) Ja, \tag{3.23}$$

where $H = W^{\frac{1}{2}} X \left(X^T W X\right) X^T W^{\frac{1}{2}}$, $W = diag(w_i) = diag(V(\mu_i)) = diag(\mu_i^2)$, $J = diag\left(V^{-\frac{1}{2}}\mu''\right) = diag(2\mu^2)$ and $a = (a_{11}, a_{22}, ..., a_{nn})^T$ is an $n \times 1$ vector.

Let $v_i = (Var(R_{Pi}))^T$ be the vector of the variances of the Pearson residuals of order $O(n^{-1})$. Then, $v_i = (Var(R_{Pi}))^T$,

$$E\left(2R_{Pi}H_r^{(i)}U_s^{(i)}\right) = \left(-V_i^{-2}V_i^{(1)^2}\mu'^2_i - 2\phi^{-1}V_i^{-1}\mu'^2_i\right) z_{ir}z_{is} = -\left(2 + \phi^{-1}\right)\mu_i^2 z_{ir}z_{is},$$

$$E\left(H_r^{(i)}H_s^{(i)}\right) = \left(\phi^{-1} + \frac{V_i^{(1)^2}}{4V_i}\right) w_i z_{ir}z_{is} = \left(\phi^{-1} + 1\right) w_i z_{ir}z_{is}$$

and

$$E\left(R_{Pi}H_{rs}^{(i)}\right) = \frac{1}{2\phi}\left(V_i^{-\frac{1}{2}}\mu'_i\right) = -\frac{\mu_i}{2\phi}.$$

Then

$$-\sum_{r,s=1}^{p} K^{rs} E\left(2R_{Pi}H_r^{(i)}U_s^{(i)} + H_r^{(i)}H_s^{(i)} + R_{Pi}H_{rs}^{(i)}\right) = \left(-\phi^{-1}w_i - \frac{V_i^{(1)}\mu''_i}{2V_i} - \frac{V_i^{(2)}}{2}\right) a_{ii}\phi = \left(-\phi^{-1} - 3\right) w_i a_{ii}\phi. \tag{3.24}$$

We also have

$$2\sum_{r=1}^{p} B\left(\hat{\beta}_r\right) E\left(R_{Pi}H_r^{(i)}\right) = \left(\frac{\phi}{2}\frac{V_i^{(1)}\mu'_i}{V_i}\gamma_i^T A A_d F_1\right) = -\phi\mu_i\gamma_i^T A A_d F\mathbf{1}. \tag{3.25}$$

Thus from Eq. (3.21), Eq. (3.24) and Eq. (3.25), we have

$$\nu = 1 + \frac{\phi}{2}\left(QHJ - T\right)z, \tag{3.26}$$

where $Q = diag\left(V^{-\frac{1}{2}}V^{(1)}\right) = diag\left(2\right)$ and $T = diag\left(2\phi^{-1}w + wV^{(2)} + \frac{V^{(1)}\mu''}{V}\right) = diag\left(\left(2\phi^{-1} + 6\right)\mu^2\right)$. From Eq. (3.23) and Eq. (3.26), we can define the adjusted Pearson residuals as

$$Ar_{Pi} = \frac{R_{Pi} - \hat{E}\left(R_{Pi}\right)}{\sqrt{\hat{V}\left(R_{Pi}\right)}}, \tag{3.27}$$

where $\hat{E}\left(R_{Pi}\right)$ and $\hat{V}\left(R_{Pi}\right)$ are obtained by replacing μ_i with the estimated value $\hat{\mu}_i$. The adjusted Pearson residuals are approximately follow a normal distribution (Cordeiro, 2004). Further he suggests that a similar can be done to study the deviance residuals. These studies can be helpful for the researchers in determining which one residual is more suitable for model diagnostics.

3.8.2 The Adjusted Deviance Residuals

In this section, we developed the adjusted deviance residuals by following Cox and Snell (1968) and Cordeiro (2004). So some theoretical justification are given here as following, The true deviance residuals for the GRM are defined by

$$R_{di} = sign\left(y_i - \mu_i\right)\sqrt{\left|\frac{-2}{\phi}\left\{\log\left(\frac{\mu_i}{y_i}\right) + \left(\frac{y_i - \mu_i}{\mu_i}\right)\right\}\right|}, \tag{3.28}$$

where $sign\left(y_i-\mu_i\right)$ is the signum function and defined as

$$sign\left(y_i-\mu_i\right)=\begin{cases} +\ if\ y_i>\mu_i \\ 0\ if\, y_i=\mu_i \\ -\ if\ y_i>\mu_i \end{cases}$$

The above relation shows that sign function indicated the signs $(+,-)$ of the difference between observed response and their mean function. So for the adjusted deviance residuals, we omit the signum function and Eq. (3.28) now becomes

$$R_{di}=\left[\frac{2}{\phi}\left\{\log\left(\frac{\mu_i}{y_i}\right)+\left(\frac{y_i-\mu_i}{\mu_i}\right)\right\}\right]^{\frac{1}{2}}. \tag{3.29}$$

For the adjustment in deviance residuals defined in Eq. (3.29), we require the mean and the variance. For the computation of these moments, it follows from Cordeiro (2004) that

$$H_r^{(i)}=\frac{\partial R_{di}}{\partial\beta_r^*}\ \ \&\ \ H_{rs}^{(i)}=\frac{\partial^2 R_{di}}{\partial\beta_r^*\partial\beta_s^*}.$$

By following Cox and Snell's (1968) work, the first two moments of the deviance residuals to order $O\left(n^{-1}\right)$ are

$$E\left(R_{di}\right)=\sum B\left(\hat{\beta}_r\right)E\left(H_r^{(i)}\right)-\sum K^{rs}E\left(H_r^{(i)}U_s^{(i)}+\frac{1}{2}H_{rs}^{(i)}\right). \tag{3.30}$$

$$V\left(R_{di}\right)=1+2\sum B\left(\hat{\beta}_r\right)E\left(R_{di}H_r^{(i)}\right)-\sum K^{rs}E\left(2R_{di}H_r^{(i)}U_s^{(i)}+H_r^{(i)}H_s^{(i)}+R_{di}H_{rs}^{(i)}\right). \tag{3.31}$$

For computing moments of the deviance residuals, we introduced the following variables $t_{1i} = y_i \mu_i^{-1}$,$t_{2i} = \ln\left(\mu_i^{-1}\right)$, $t_{3i} = t_{1i} + t_{2i}$, $t_{4i} = t_{1i}\left(4 + t_{1i}\right) + 2\ln\left(t_{1i}\right) - 1$, $t_{5i} = t_{1i} + ln\left(t_{1i}\right) - 1$, $t_{6i} = \left(4 - t_{1i}\right) + 2\ln\left(t_{1i}\right)$.

We obtained $E\left(H_r^{(i)}\right) = \frac{t_{3i}}{\phi} z_{ir}$ and $E\left(H_{rs}^{(i)}\right) = \frac{1}{2\phi}\left(\frac{t_{4i}}{t_{5i}}\right)\mu_i' z_{ir} z_{is}$.

We obtain $\sum_{r=1}^{p} B\left(\hat{\beta}_r\right) E\left(H_r^{(i)}\right) = \frac{1}{\phi}\left(t_{3i}\right)\gamma_i^T ZB\left(\hat{\beta}\right)$.

Then, we have $E\left(H_r^{(i)} U_s^{(i)}\right) = \frac{t_{3i}^2}{\phi^2} z_{ir} z_{is}$, $E\left(H_r^{(i)} U_s^{(i)} + \frac{1}{2} H_{rs}^{(i)}\right) = \frac{1}{\phi}\left(\frac{t_{3i}^2}{\phi} - \frac{t_{4i}}{4t_{5i}}\mu'_i\right) z_{ir} z_{is}$ and $-\sum_{r,s=1}^{p} K^{rs} E\left(H_r^{(i)} U_s^{(i)} + \frac{1}{2} H_{rs}^{(i)}\right) = \left(\frac{t_{3i}^2}{\phi} - \frac{\phi^3 t_{4i}}{4\left(\phi - \psi'\left(\frac{1}{\phi}\right)\right)}\right)\mu_i' z_{ir} z_{is}$,

where $\psi\left(g\right) = \frac{\partial \log\Gamma(g)}{\partial g} = \frac{\Gamma'(g)}{\Gamma(g)}$, $\Gamma'\left(g\right) = \frac{\partial\Gamma(g)}{\partial g}$ is the digamma function, $g > 0$ and $\psi'\left(g\right) = \frac{\partial\psi(g)}{\partial g}$ is the trigamma function.

Let $r^* = \left(E\left(R_{di}\right)\right)^T$; for $i = 1, 2, ..., n$ be the vector of expected deviance residuals of order $O\left(n^{-1}\right)$. Then,

$$r^* = -\frac{\phi}{2}\left(I - H\right) Ja, \tag{3.32}$$

where $W = diag\left(w_i\right) = diag\left(V\left(\mu_i\right)\right) = diag\left(\mu_i^2\right)$, $J = diag\left(\frac{t_{4i}}{t_{5i}}\right)$ and $a = \left(a_{11}, a_{22}, ..., a_{nn}\right)^T$ is an $n \times 1$ vector.

Let $v^* = \left(Var\left(R_{di}\right)\right)^T$ be the vector of the variances of the deviance residuals of order $O\left(n^{-1}\right)$. Then,

$E\left(2R_{di} H_r^{(i)} U_s^{(i)}\right) = \frac{\sqrt{2}}{\phi}\left(\frac{y_i + \mu_i}{\sqrt{t_{5i}}}\right) z_{ir} z_{is}$ and $E\left(R_{di} H_{rs}^{(i)}\right) = \frac{\mu_i}{2\phi^2}\left(\frac{t_{4i} t_{6i}}{t_{5i}}\right) z_{ir} z_{is}$.

such that

$$-\sum_{r,s=1}^{p} K^{rs} E\left(2R_{di} H_r^{(i)} U_s^{(i)} + H_r^{(i)} H_s^{(i)} + R_{di} H_{rs}^{(i)}\right) = \left(\frac{3t_{4i}}{t_{5i}}\mu'''_i + \frac{\phi^{\frac{7}{2}}\left(y_i + \mu_i\right)^2\left(2 + t_{5i}\right)}{\sqrt{2}\left(\phi - \psi'\left(\frac{1}{\phi}\right)\right) t_{5i}}\right) z_{ir} z_{is}. \tag{3.33}$$

We also have

$$2\sum_{r=1}^{p} B\left(\hat{\beta}_r\right) E\left(R_{Pi}H_r^{(i)}\right) = \left(\frac{4t_{3i}}{\phi V_i^{(1)}}\left(t_{5i}-5\right)\gamma_i^T A A_d F\mathbf{1}\right). \tag{3.34}$$

Thus from Eq. (3.31), Eq. (3.33) and Eq. (3.34), we have

$$\nu^* = 1 + \frac{\phi}{2}\left(QHJ + T\right) A_d\, a, \tag{3.35}$$

where $Q = diag\left(\frac{t_{4i}t_{6i}}{t_{5i}}\mu_i\right)$ and $T = diag\left(\frac{t_{4i}}{t_{5i}}3\mu'''_i\right)$. From Eq. (3.32) and Eq. (3.35), the adjusted deviance residuals are defined by

$$Ar_{di} = \frac{R_{di} - \hat{E}\left(R_{di}\right)}{\sqrt{\hat{V}\left(R_{di}\right)}}, \tag{3.36}$$

where $\hat{E}\left(R_{di}\right)$ and $\hat{V}\left(R_{di}\right)$ are evaluated by replacing μ_i with the estimated value $\hat{\mu}_i$. In the similar way, we also compute the adjusted likelihood residuals, the adjusted working and the adjusted Anscombe residuals. Latterly, we used these residuals in computing influence diagnostics separately.

3.9 Influence Diagnostics in the GRM

Several studies have been conducted for influence diagnostics in various forms of the GRM. Brown studied the local influence diagnostic for the gamma model with inverse link function. Ortega *et al.* (2003) detected the influential observations by using local method for the generalized log-gamma regression model. Also the similar study is given by Hashimoto *et al.* (2013) for the estimation and diagnostic analysis in log-gamma regression model with censored time interval. Jearkpaporn *et al.* (2005) study the robust GRM using control charts with unusual observations.

3.9.1 Cook's Distance in the GRM

The Cook's distance for the GRM by following Eq. (2.37) defined by

$$CD_i = \frac{\left(\hat{\beta}^* - \hat{\beta}^*_{(i)}\right)^T Z^T \hat{W} Z \left(\hat{\beta}^* - \hat{\beta}^*_{(i)}\right)}{\hat{\phi} p'}. \tag{3.37}$$

After simplification, Eq. (3.37) becomes

$$CD_i = \frac{sr^2_{Pi}}{p'} \frac{h_{ii}}{m_{ii}}. \tag{3.38}$$

For the computation of CD with the other GRM residuals, we just replace the standardized Pearson residuals with other form of the GRM residuals. For example with the likelihood residuals defined in Eq. (3.12), now Eq. (3.38) become as

$$CD_i = \frac{r^2_{li}}{p'} \frac{h_{ii}}{m_{ii}}. \tag{3.39}$$

Similar is done for the other types of the GRM residuals. The cutoff values for Eq. (3.38) and Eq. (3.39) are same as we defined in Chapter 2, Section 2.12, page 49.

3.9.2 Modified Cook's Distance in the GRM

The MCD for the GRM using Eq. (2.41) defined by

$$MCD_i = \left[\frac{n-p'}{p'}\frac{h_{ii}}{m_{ii}}\right]^{\frac{1}{2}} |t_i| \tag{3.40}$$

where $t_i = sr_{Pi}\sqrt{\frac{n-p'}{n-p-sr_{Pi}^2}}$ and the observation is influential observation, if $MCD \geqslant 2\sqrt{\frac{n-p'}{n}}$. For the other form of the GRM residuals, we compute t_i with replacing sr_{Pi} by other GRM standardized residuals defined in Table 3.1. The value of t_i for other the GRM residual are given in Table 3.1. The cutoff point for the MCD with all forms of the GRM residuals is similar as we defined in Chapter 2, Section 2.12, page 49.

Table 3.1: The value of t_i for different GRM Residuals

GRRM Residuals	Standardized	Adjusted
Pearson	$t_i = sr_{Pi}\sqrt{\frac{n-p'}{n-p-sr_{Pi}^2}}$	$t_i = Ar_{Pi}\sqrt{\frac{n-p'}{n-p-Ar_{Pi}^2}}$
Deviance	$t_i = sr_{di}\sqrt{\frac{n-p'}{n-p-sr_{di}^2}}$	$t_i = Ar_{di}\sqrt{\frac{n-p'}{n-p-Ar_{di}^2}}$
Likelihood	$t_i = r_{li}\sqrt{\frac{n-p'}{n-p-r_{li}^2}}$	$t_i = Ar_{li}\sqrt{\frac{n-p'}{n-p-Ar_{li}^2}}$
Anscombe	$t_i = sr_{ai}\sqrt{\frac{n-p'}{n-p-sr_{ai}^2}}$	$t_i = Ar_{ai}\sqrt{\frac{n-p'}{n-p-Ar_{ai}^2}}$
Working	$t_i = sr_{wi}\sqrt{\frac{n-p'}{n-p-sr_{Wi}^2}}$	$t_i = Ar_{Wi}\sqrt{\frac{n-p'}{n-p-Ar_{Wi}^2}}$

3.9.3 Andrews Pregibon in the GRM

Another influence measure which we defined in Chapter 2 is the AP, so for the GRM it is defined by

$$AP_i = m_{ii} - \frac{sr_{Pi}^2}{\sum_{i=1}^{n} sr_{Pi}^2}. \tag{3.41}$$

Similarly the AP is computed with other forms of the GRM residuals by just replacing sr_{Pi} with other forms of the standardized GRM residuals. The influence diagnostic criteria of AP statistic is similar as we defined in Chapter 2.

3.9.4 Covariance Ratio in the GRM

CVR_i can also be used as a influence diagnostic statistic in the GRM. CVR_i for the GRM under the assumption that the independent variables are not multicollinear is defined by

$$CVR_i = \frac{\left[\frac{\left(n-p-sr_{Pi}^2\right)}{n-p'}\right]^{p'}}{m_{ii}}. \tag{3.42}$$

Chatterjee and Hadi (2012) indicated that CVR may exceed one, if standardized residuals are small and leverages are large and CVR may be smaller than one, if standardized residuals are large and leverages are small. While CVR approaches one if both standardized residuals and leverages are large (small) simultaneously. So these two factors may affect the influence diagnostic performance of the CVR method.

In the similar way, CVR with adjusted Pearson residuals (Arp) is obtained by replacing the srp with the Arp as

$$CVR_i = \frac{\left[\frac{\left(n-p-Ar_{Pi}^2\right)}{n-p'}\right]^{p'}}{m_{ii}}. \tag{3.43}$$

Similar computation of CVR is done with other types of the GRM residuals. The influence detection criteria is same as we mentioned in earlier chapter.

3.9.5 DFFITS in the GRM

DFFITS is also a measure of influence diagnostics and for the GRM according to Chatterjee and Hadi (1988) using the standardized Pearson residuals is defined by

$$DFFITS = |t_i| \sqrt{\frac{h_{ii}}{m_{ii}}}, \tag{3.44}$$

where t_i is already defined for all of the GRM residuals as given in Table 3.1. The criteria of cut-off point for the GRM DFFITS diagnostic is same as we discussed in Chapter 2.

3.9.6 Welsch Distance in the GRM

According to the Eq. (2.49), WD for the GRM defined as

$$WD_i = (n-1)\, t_i^2 \frac{h_{ii}}{m_{ii}^2}, \tag{3.45}$$

where t_i is given in Table 3.1. Similarly, WD is computed with other forms of the GRM residuals. The declared as influential observation by the WD method, if $WD_i > 3\sqrt{p'}$.

3.9.7 Hadi's Measure in the GRM

By following Hadi (1992), the influence can be extended to the GRM. According to Eq. (2.50), this diagnostic as

$$Hd_i = \frac{p'}{m_{ii}}\frac{d_i^2}{1-d_i^2} + \frac{h_{ii}}{m_{ii}}, \tag{3.46}$$

where $d_i = \frac{sr_{Pi}^2}{\sum sr_{Pi}^2}$ based on the standardized Pearson residuals in the GRM. On the basis of H_d measure, Hadi (1993) has pointed out that the observation is influential observation, if for the ith case, $H_{di} > Mean\,(H_{di}) + c\sqrt{Var\,(H_{di})}$.

where c may be 2 or 3. Similarly, the H_d statistic is computed with other forms of the GRM residuals and with similar cutoff point.

3.10 The Gamma Ridge Regression Model

As we have already noted that the ridge regression is used to handle multicollinearity in the independent variables for estimation purposes but not in predication (Seeber and Lee, 2003). There are the situations when the independent variables are correlated (ill-conditioned) but the data analyst does not diagnose and test the impact of the collinearity level on the parameter estimates. When the independent variables are highly correlated, then small change in the data may cause large changes in the estimated regression coefficients (Chatterjee and Hadi, 2012). These small changes are observed through ridge regression estimation process to control the ill-condition problem. The ill-conditioned data in the GLM has similar effects on the parameter estimates as in estimating the LM (Segerstedt, 1992). The ridge estimation in the GLM for the first time was studied by Schaefer *et al.* (1984) with reference to logistic regression.

3.11 Testing Multicollinearity in Gamma Regression Model

There are several methods for the testing of the multicollinearity which are available in literature regarding the LM. These methods include, variance inflation factor, examination of eigenvalues and eigenvectors of X^TX and condition index. Here our interest is the testing multicollinearity in the GRM that is the special case of the GLM. Multicollinearity in the GLM is tested through the condition number index for further detail see (Lesaffre and Marx, 1993). If $\lambda_1, \lambda_2, ..., \lambda_p$ are the eigenvalues of X^TWX , then condition index (CI) for the GLM cases is defined as

$$CI = \sqrt{\frac{\max(\lambda_j)}{\min(\lambda_j)}};\ j = 1, 2, ..., p. \tag{3.47}$$

CI examined the existence of multicollinearity but not explain the structure of dependencies among the columns of X. Roozbeh and Arashi (2014) had suggested that the existence and structure of multicollinearity examined in the better way with the help of eigenvalues and eigenvectors of X^TWX. Eigen values close to zero indicates the presence of multicollinearity, then *CI* is increased. Belseley *et al.* (1980) suggests that when *CI* is less than 10 not the serious problem but the serious problem of multicollinearity is occured when CI is larger than 30.

3.12 The Estimation of Gamma Ridge Regression

Before going to the estimation process, first we discuss why the centering and scaling is so important in statistical modeling with multicollinear independent variables. Maquardt (1980) says that scaling of the independent variables may reduce the unnecessary ill-conditioning problem. Seeber and Lee (2003) have shown that centering of independent variables do not affect the predicted values of the dependent variable. They also say that some authors recommended that does not use the centered and scaling the data but use in rescaled estimates. This does not lead to the scale invariant property unlike the OLS. Center and scaling in regression modeling is also helpful in the estimation process and due to this the multicollinearity in X^TX is reduced results X^TX is non-orthogonal. Center and scaling in the LM are made in such a way that $X^TX = r_{ij}$ the sample correlation matrix and $\sum\limits_{i=1}^{n} x_{ij}^2 = 1$. For the GRM case, we have to test the orthogonality in $X^T\hat{W}X$ subject to the conditions as mentioned earlier.

As the GRM estimation based on IRLS method, so multicollinearity affects the IRLS estimates and is adjusted at each iteration of the IRLS. As we know, for the GRM, $\hat{\beta}^* = \left(Z^T\hat{W}Z\right)^{-1} Z^T\hat{W}y^*$ and $V\left(\hat{\beta}^*\right) = \hat{\phi}\left(Z^T\hat{W}Z\right)^{-1}$. This can be seen that multicollinearity affect the parameter estimates and their variances. For the estimation of GRM with multicollinearity, the most popular method is the ridge estimation methods with the assumption that the MSE of ridge estimator less than the MSE of standard GLM. Some authors explored ridge regression for some of the GLMs. If the assumption of no multicollinearity is violated, then variances of $\hat{\beta}^*$ i.e. $V\left(\hat{\beta}^*\right) = \hat{\phi}\left(Z^T\hat{W}Z\right)^{-1}$ are so large (Mansson *et al.*, 2012). To overcome this problem, Schaeffer *et al.* (1984) introduce the ridge estimation method in logistic regression as in the GLM case, which is based on Horeld and Kennard (1970b) method. Following this, the GRM estimator under multicollinearity now becomes the GRRM. Before the calculation of GRRM estimator, it should be noted that Schaeffer *et al.* (1984) suggested that add small quantity k in the diagonal matrix of $Z^T\hat{W}Z$ that makes the biased estimator and so called K the ridge parameter. So the GRRM parameter vector with nearly singular of $Z^T\hat{W}Z$ is computed as

$$\hat{\beta}_R = \left(Z^T\hat{W}Z + KI\right)^{-1} Z^T\hat{W}Z\hat{\beta}^*, K \geqslant 0, \tag{3.48}$$

where K is the ridge parameter. In the GLM, it is estimated as $K = \frac{p'\hat{\phi}}{\hat{\beta}^{*\prime}\hat{\beta}^*}$ and $\hat{\varphi}$ is the estimated dispersion parameter $\hat{\varphi}$ and is computed as $\hat{\varphi} = \frac{1}{n-p'}\sum\frac{(y_i-\hat{\mu}_i)^2}{V(\hat{\mu}_i)}$, where $\hat{\mu}_i = \frac{1}{{z_i}^T\hat{\beta}^*}$ and $V\left(\hat{\mu}_i\right) = \left(\frac{1}{{z_i}^T\hat{\beta}^*}\right)^2$ are the mean and variance functions of the GRM and I is the identity matrix of order $p' \times p'$. The GRRM coincides with the GRM, if $K = 0$. The details of the GRRM estimations as given below

Let d be the distance from $\hat{\beta}_R$ to β^*, then the GRRM estimator $\hat{\beta}_R$ is obtained by minimizing $\beta_R^T\beta_R$ subject to condition that

$$\left(\hat{\beta}_R-\beta^*\right)^T Z^T\hat{W}Z\left(\hat{\beta}_R-\beta^*\right)<d^2. \tag{3.49}$$

Eq. (3.49) has the solution

$$\hat{\beta}_R=\left(Z^T\hat{W}Z+KI\right)^{-1}Z^T\hat{W}Z\hat{\beta}^*. \tag{3.50}$$

By following Schaefer *et al.* (1984), the GRM estimate is obtained by minimizing the weighted sum of square due to error (WSSE) as

$$WSSE=\left(y-g^{-1}\left(\eta\right)\right)^T W^{-1}\left(y-g^{-1}\left(\eta\right)\right). \tag{3.51}$$

So for the GRRM Eq. (3.49) has the form as

$$d^2=WSSE\left(\beta_R\right)-WSSE\left(\hat{\beta}^*\right).$$

$$d^2=\left(y-g^{-1}\left(\eta_R\right)\right)^T W^{-1}\left(y-g^{-1}\left(\eta_R\right)\right)+\left(y-g^{-1}\left(\eta\right)\right)^T W^{-1}\left(y-g^{-1}\left(\eta\right)\right),$$

where $\eta_R=z_i'\beta_R$ be the vector of linear predictors with ridge estimator. The above equation further may be written as

$$d^2=\left(y-\mu_R\right)^T W^{-1}\left(y-\mu_R\right)-\left(y-\hat{\mu}\right)^T W^{-1}\left(y-\hat{\mu}\right),$$

where $\mu \,\&\, \mu_R$ are the mean functions of the GRM and GRRM respectively.

$$d^2 = (\mu_R - \hat{\mu})^T W^{-1} (\mu_R - \hat{\mu}) + 2(\mu_R - \hat{\mu})^T W^{-1} (y - \hat{\mu}) \,.$$

By the Tylor series, $\mu \,\&\, \mu_R$ are approximated as

$$\mu_R = \hat{\mu} + \hat{W} Z \left(\hat{\beta}_R - \beta^*\right)$$

and

$$\hat{\mu} = \mu + \hat{W} Z \left(\hat{\beta}^* - \beta\right).$$

The above equations are further simplified to

$$d^2 = \left(\hat{\beta}_R - \beta^*\right)^T Z^T \hat{W} Z \left(\hat{\beta}_R - \beta^*\right) + \left(\hat{\beta}_R - \beta_R\right)^T Z \left(y - \hat{W} Z \hat{\beta}_R\right).$$

or

$$d^2 = \left(\hat{\beta}_R - \beta^*\right)^T Z^T \hat{W} Z \left(\hat{\beta}_R - \beta^*\right).$$

Now by the evaluation of Hoerl and Kennard (1970b), there exist a constant K subject to the condition that for large n and $K \geqslant 0$ as

$$MSE\left(\hat{\beta}_R\right) < MSE\left(\hat{\beta}^*\right),$$

where

$$MSE\left(\hat{\beta}^*\right) = E\left\{\left(\hat{\beta}^* - \beta^*\right)^T\left(\hat{\beta}^* - \beta^*\right)\right\} = tr\left(Z^T\hat{W}Z\right)^{-1} = \sum_{j=1}^{p}\lambda_j^{-1}. \tag{3.52}$$

and

$$\begin{aligned} MSE\left(\hat{\beta}_R\right) &= E\left\{\left(\hat{\beta}_R - \beta^*\right)^T\left(\hat{\beta}_R - \beta^*\right)\right\} \\ &= \sum_{j=1}^{p}\frac{\lambda_j}{\left(\lambda_j + K\right)^2} + K^2\sum_{j=1}^{p}\frac{\alpha_j}{\left(\lambda_j + K\right)^2} \\ &= tr\left(Z^T\hat{W}Z + KI\right)^{-1}, \end{aligned} \tag{3.53}$$

where λ_j *and* α_j are the eigenvalues and eigenvectors of $X^T\hat{W}X$. Smaller the distance from $\hat{\beta}_R$ to β^* produced smaller the MSE and larger the distance from $\hat{\beta}_R$ to β^* produced larger the MSE.

In Eq. (3.52), $MSE\left(\hat{\beta}^*\right)$ becomes so large when λ_j approaches to 0 and minimum MSE is obtained when λ_j approaches to 1. Segerstedt (1992) had pointed out the problem for the comparison of $\hat{\beta}_R$ with $\hat{\beta}^*$ that the weight matrix $\hat{W}$ is evaluated at different values and suggest the ridge estimator for the GLM cases as

$$\tilde{\beta}_R = \left(\left(Z^T\hat{W}Z\right)^{-1} + KI\right)^{-1}\left(Z^T\hat{W}Z\right)^{-1}, K \geqslant 0. \tag{3.54}$$

He further suggests that for the $\tilde{\beta}_R$, consider the fixed value of K. The approximation of $\tilde{\beta}_R$ to $\hat{\beta}_R$ are also studied by several researchers for the logistic regression model (Schaefer *et al.*, 1984; Schaefer *et al.*, 1986; Lee and Silvapulle, 1988). Segerstedt

(1992) had given the covariance formula with the ridge estimates as

$$Cov\left(\tilde{\beta}_R\right) = \left(Z^T\hat{W}Z + KI\right)^{-1}\left(Z^T\hat{W}Z\right)\left(Z^T\hat{W}Z + KI\right)^{-1}, K \geqslant 0. \tag{3.55}$$

He also noticed that it is not possible to calculate $\tilde{\beta}_R$ because it depends on the unknown value of covariance of the GLM estimate. So the best way to compute the GRRM as under

$$\hat{\beta}_R = \left(Z^T\hat{W}Z + KI\right)^{-1} Z^T\hat{W}Z\hat{\beta}^*. \tag{3.56}$$

3.13 The Selection of Ridge Parameter for the GRRM

There are various approaches to choose ridge parameter (k) are available in literature with different error distributions. Different ridge parameters are found better in different error distributions in the sense that MSE of the ridge estimate smaller than any other GLM estimation method. So there is no exact method for the ridge parameters. The estimation and comparison of the ridge parameters for the Poisson regression, logistic regression and negative binomial regression are available in the literature as we have already given in Chapter 1. No study yet available for the best ridge parameters for the GRM. So here we use the optimum value of k given by Hoerl *et al.* (1975) as $K = \frac{p'\hat{\phi}}{\hat{\beta}^{*T}\hat{\beta}^*}$ for influence diagnostics in the GRRM. As the value of ridge parameter k is increases then the GRRM estimates increases while with increase in k, the variances decrease.

3.14 The GRRM Residuals

With presence of multicollinearity, the structure of the GRM residuals are deviate due to the lage variances of the estimates. So alternatively the computation of these residuals is done with biased (ridge) estimator. For the GRRM, all forms of the residuals are given by.

3.14.1 The GRM Pearson Residuals with Ridge Estimator

The Pearson residuals with ridge estimator for the GRRM are defined by

$$r_{pRi} = \frac{y_i - \hat{\mu}_{Ri}}{\sqrt{Var\left(\hat{\mu}_{Ri}\right)}}, \tag{3.57}$$

where $\hat{\mu}_{Ri} = \frac{1}{z_i^T \hat{\beta}_R}$ and $V\left(\hat{\mu}_{Ri}\right) = \left(\frac{1}{z_i^T \hat{\beta}_R}\right)^2$ are the mean and variance functions respectively for the GRRM. The standardized Pearson residuals for the GRM with collinear independent variables are defined by

$$sr_{pRi} = \frac{r_{pRi}}{\sqrt{\hat{\phi}_R m_{Rii}}}, \tag{3.58}$$

where $\hat{\phi}_R = \frac{1}{n-p'} \sum \frac{(y_i-\hat{\mu}_{Ri})^2}{V(\hat{\mu}_{Ri})}$ is the dispersion parameter with ridge estimate, $m_{Rii} = 1 - h_{Rii}$ and h_{Rii} are the ith diagonal elements of the hat matrix with ridge estimator H_{Rii}, $H_R = \hat{W}^{\frac{1}{2}} Z \left(Z^T \hat{W} Z + KI\right)^{-1} Z^T \hat{W}^{\frac{1}{2}}$ defined in upcoming section.

3.14.2 The GRM Deviance Residuals with Ridge Estimator

The deviance residuals for the GRM with ridge estimator are defined by

$$r_{d_{Ri}} = sign\left(y_i - \hat{\mu}_{Ri}\right) \sqrt{|d_{Ri}|}, \tag{3.59}$$

where $d_{Ri} = -\frac{2}{\hat{\phi}_{Ri}} \left\{ \ln\left(\frac{y_i}{\hat{\mu}_{Ri}}\right) - \frac{y_i - \hat{\mu}_{Ri}}{\hat{\mu}_{Ri}} \right\}$ is the deviance function of the GRM with ridge estimator.

The standardized deviance residuals of the GRM with ridge estimator are defined by

$$sr_{dRi} = \frac{r_{dRi}}{\sqrt{\hat{\phi}_R m_{Rii}}}. \tag{3.60}$$

This form of deviance residual is used in identifying the influential observation, when there is multicollinearity in the independent variables.

3.14.3 The Likelihood Residuals of the GRRM

For the GRRM, likelihood residuals by Eq. (3.12) are reformulated as

$$r_{lRi} = sign\left(y_i - \hat{\mu}_{Ri}\right) \sqrt{h_{Rii}\left(sr_{pRi}\right)^2 + \left(1 - h_{Rii}\right)\left(sr_{dRi}\right)^2}, \tag{3.61}$$

where sr_{pRi} & sr_{dRi} are the standardized Pearson and deviance residuals with ridge estimates as defined in Eq. (3.58) and Eq. (3.60). The influence diagnostics are also computed with the likelihood ridge residuals as defined Eq. (3.61), when the independent variables are multicollinear.

3.14.4 The Anscombe Residuals of the GRRM

The Anscombe residuals of the GRM with ridge estimators by following Eq. (3.14) are defined by

$$r_{aRi} = \frac{3\left(y_i^{\frac{1}{3}} - \hat{\mu}_{Ri}^{\frac{1}{3}}\right)}{\hat{\mu}_{Ri}^{\frac{1}{3}}}. \tag{3.62}$$

The standardized Anscombe residuals of the GRM with ridge estimator are defined by

$$sr_{aRi} = \frac{r_{aRi}}{\sqrt{\hat{\phi}_R m_{Rii}}}. \tag{3.63}$$

This type of residuals are also used for the detection of influential observation for testing their significance.

3.14.5 The GRRM Working Residuals

The GRRM working residuals are the extension of Eq. (3.17) under the assumption of independent variables and are defined as

$$r_{WRi} = (y_i - \hat{\mu}_{Ri})\left(-\frac{1}{\hat{\mu}_{Ri}^2}\right). \tag{3.64}$$

Similarly, the standardized working residuals with ridge estimator are defined by

$$sr_{wRi} = \frac{r_{wRi}}{\sqrt{\hat{\phi}_R m_{Rii}}}. \tag{3.65}$$

The standardized working residuals with ridge estimator are also used for the identification of influential observation.

3.15 Hat Matrix and Leverages for the GRRM

Hat matrix for the GRRM is defined by

$$H_R = \hat{W}^{\frac{1}{2}} Z \left(Z^T \hat{W} Z + KI \right)^{-1} Z^T \hat{W}^{\frac{1}{2}}. \tag{3.66}$$

While the leverages with ridge estimator are the diagonal elements of the hat matrix H_R, given in Eq. (3.66), $h_{Rii} = diag\left(H_R\right)$. The leverages with ridge estimator can also be helpful in detection of influential observation but some modifications are needed because leverages with ridge estimator are smaller than the leverages of the GRM i.e. $h_{Rii} \leqslant h_{ii}$. So we suggest that, the observations with leverages greater than $\frac{p'}{n}$ are declared as influential observation.

3.16 Influence Diagnostics in the GRRM

There are the situations where the problem of multicollinearity and influential observations occurred simultaneously. These two problems separately or jointly may hide the role of independent variables which are really playing critical role in the GRM model building process. Therefore there is need to develop or modify the influence diagnostics for positively skewed data with collinear independent variables. In this section, we modify some of the influence diagnostics with ridge estimate of the GRM. Kamruzzaman and Imon (2002) have shown that influential observation is the new source of multicollinearity. These ridge parameters may be affected by outliers, high leverages and influential observations in the data set. Diagnostics of the influential observation with multicollinearity improves the GRM estimates. These diagnostics are based on the GRRM residuals, dispersion parameter and the leverages. These diagnostics may produce different influential observations and their impact on the GRRM estimates as the GRM estimates. The GRRM residuals are now reformulated in the subsequent section.

3.16.1 Cook's Distance in the GRRM

By following Eq. (3.37), CD for the GRRM is defined by

$$CD_{Ri} = \frac{\left(\hat{\beta}_R - \hat{\beta}_{R(i)}\right)^T Z^T \hat{W} Z \left(\hat{\beta}_R - \hat{\beta}_{R(i)}\right)}{p' \hat{\phi}_R}. \tag{3.67}$$

Eq. (3.67) can be used to measure the distance $\hat{\beta}_{R(i)}$ to $\hat{\beta}_R$. Larger the distance between these two points indicated the ith observation is influential observation. Further Eq. (3.67) can be simplified as

$$CD_{Ri} = \frac{sr_{PRi}^2}{p'} \frac{h_{Rii}}{m_{Rii}}. \tag{3.68}$$

Influence detection criteria of is similar as for the method.

3.16.2 Modified Cook's Distance in the GRRM

The MCD diagnostic with ridge estimator in the GRM to diagnose the influential observation is defined as

$$MCD_{Ri} == \left[\frac{\left(\hat{\beta}_R - \hat{\beta}_{R(i)}\right)^T Z^T \hat{W} Z \left(\hat{\beta}_R - \hat{\beta}_{R(i)}\right)}{\frac{p'}{n-p'}\hat{\phi}_{(i)}}\right]^{\frac{1}{2}} = |t_{Ri}| \left[\frac{(n-p')}{p'}\frac{h_{Rii}}{m_{Rii}}\right]^{\frac{1}{2}}, \tag{3.69}$$

where $t_{Ri} = sr_{PRi}\sqrt{\frac{n-p'}{n-p-sr^2_{PRi}}}$ are the standardized Pearson residuals with ridge estimator. The value of t_{Ri} for the other GRRM residuals are given in Table 3.2. These values of t_{Ri} are further used in the influence diagnostics using all the GRM residuals with ridge estimator.

Table 3.2: The value of t_{Ri} for different GRRM Residuals

GRRM Residuals	Standardized	Adjusted
Pearson	$t_{Ri} = sr_{PRi}\sqrt{\frac{n-p'}{n-p-sr^2_{PRi}}}$	$t_{Ri} = Ar_{PRi}\sqrt{\frac{n-p'}{n-p-Ar^2_{PRi}}}$
Deviance	$t_{Ri} = sr_{dRi}\sqrt{\frac{n-p'}{n-p-sr^2_{dRi}}}$	$t_{Ri} = Ar_{dRi}\sqrt{\frac{n-p'}{n-p-Ar^2_{dRi}}}$
Likelihood	$t_{Ri} = r_{lRi}\sqrt{\frac{n-p'}{n-p-r^2_{lRi}}}$	$t_{Ri} = Ar_{lRi}\sqrt{\frac{n-p'}{n-p-Ar^2_{lRi}}}$
Anscombe	$t_{Ri} = sr_{aRi}\sqrt{\frac{n-p'}{n-p-sr^2_{aRi}}}$	$t_{Ri} = Ar_{aRi}\sqrt{\frac{n-p'}{n-p-Ar^2_{aRi}}}$
Working	$t_{Ri} = sr_{wRi}\sqrt{\frac{n-p'}{n-p-sr^2_{wRi}}}$	$t_{Ri} = Ar_{wRi}\sqrt{\frac{n-p'}{n-p-Ar^2_{wRi}}}$

3.16.3 Andrew's Pregibon in the GRRM

From the Eq. (3.41), we can retransformed the AP with ridge estimator as

$$AP_{Ri} = m_{Rii} - \frac{sr^2_{PRi}}{\sum_{i=1}^{n} sr^2_{PRi}}, \tag{3.70}$$

where $m_{Rii} = 1 - h_{Rii}$ and the influence diagnostic criteria for the AP_{Ri} is same as for the AP.

3.16.4 Covariance Ratio in the GRRM

Like the GRM, the CVR_i can also be used as an influence diagnostic statistic in the GRRM for detecting the influence on variance-covariance matrix of the estimates. From Eq. (3.41), CVR_i using standardized Pearson residuals with ridge estimate is re-transformed as

$$CVR_{Ri} = \frac{\left[\frac{\left(n-p-sr^2_{PRi}\right)}{n-p'}\right]^{p'}}{m_{Rii}}. \tag{3.71}$$

Similar computation of CVR_{Ri} is done with the other types of the GRRM residuals. The influence detection criteria is same as CVR_i.

3.16.5 Welsch Distance in the GRRM

Following Eq. (3.4), the WD for the GRRM is defined as

$$WD_{Ri} = (n-1)\, t^2_{Ri} \frac{h_{Rii}}{m^2_{Rii}}, \tag{3.72}$$

where t_{Ri} is given in Table 3.2. This diagnostic is used to detect the influential observation which influence on the predicted values of the dependent variable. Similarly, WD_{Ri} is computed with other forms of the GRRM residuals. The cut-off- point for the WD_{Ri} is similar as the WD.

3.16.6 DFFITS in the GRRM

The DFFITS for the GRRM is obtained by following Eq. (3.44) as given by

$$DFFITS_{Ri} = \frac{\hat{\mu}_{Ri} - \hat{\mu}_{Ri(i)}}{\sqrt{\hat{\phi}_{(i)} h_{Rii}}} = \frac{W_{ii}^{\frac{1}{2}} z_i^T \left(\hat{\beta}_R - \hat{\beta}_{R(i)}\right)}{\sqrt{\hat{\phi}_{(i)} h_{Rii}}}. \tag{3.73}$$

According to SMW theorem, Eq. (3.73) further can be simplified as

$$DFFITS_{Ri} = |t_{Ri}| \sqrt{\frac{h_{Rii}}{m_{Rii}}}. \tag{3.74}$$

The cut-off value of DFFITS with ridge estimator is similar as defined earlier in Chapter 2.

3.16.7 The H_d Method in the GRRM

From the Eq. (3.46), Hadi's influence diagnostic with ridge estimator is defined by

$$Hd_{Ri} = \frac{p'}{m_{Rii}} \frac{d_{Ri}^2}{1 - d_{Ri}^2} + \frac{h_{Rii}}{m_{Rii}}, \tag{3.75}$$

where $d_{Ri} = \frac{sr_{PRi}^2}{\sum sr_{PRi}^2}$ based on the standardized Pearson residuals of the GRRM. The observation is an influential observation, if $H_{dRi} > Mean\left(H_{dRi}\right)+c\sqrt{Var\left(H_{dRi}\right)}$,where c may be 2 or 3. In the similar way, the H_{dRi} is computed with other forms of the GRRM residuals.

All these diagnostics for the GRM and the GRRM with different residuals are computed for comparison purposes to test their influence diagnostic performance. Also determine the performance of residuals in detecting the influential observations.

3.17 Numerical Results

In this section, we compare the GRM influence diagnostics methods with ridge estimators using a Monte Carlo study and a real data set.

3.17.1 Monte Carlo Study

We conduct a Monte Carlo study to compare the performance of the GRRM influence diagnostic methods with the different forms of the GRM residuals. The data generation process for the GRRM with $p = 4$ explanatory variables as $y \sim G\left(\mu_i, \phi\right)$, where $\mu_i = E\left(y_i\right) = \left(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4}\right)^{-1}$. The dispersion parameter ϕ assumes the following arbitrary values as $\phi = 0.04,\ 0.11,\ 0.33,\ 0.67,\ 2,\ 4\ ,\ 10,$ and we take the following arbitrary values for the true parameters as $\beta_0 = 0.05, \beta_1 = 0.0025, \beta_2 = 0.005, \beta_3 = \beta_4 = 0.0001$. The number of replications each of size n=50 are set to be 1000. Following Kibria *et al.* (2012) and Aslam (2014), the explanatory variables with no influential observations were generated as $X_{ij} = \sqrt{(1-\rho^2)}Z_{ij} + \rho Z_{i5}$

and $Z_{ij} \sim G\left(10, 0.5\right)$, $i = 1, 2, ..., n$, $j = 1, 2, 3, 4.$,where ρ is the collinearity between two independent variables and we take the following arbitrary values 0.75, 0.85, 0.95 and 0.99. Then we generate the influential observation in X's by $X_{15j} = X_{15j} + a_0, j = 1, 2, 3, 4,$ where $a_0 = \overline{X}_j + 100$. Now, the performance of these diagnostics for the identification of generated influential observation with various collinearity levels and with different values of dispersions are performed on the basis of the gamma generated samples. These simulation results are perform using the R software.

3.17.1.1 *Influence Diagnostics in the GRM and the GRRM with Pearson Residuals*

In this section, we analyze the influence diagnostics with four forms of the Pearson residuals. These forms of Pearson residuals include the standardized Pearson, standardized Pearson residuals with ridge estimators, the adjusted Pearson residuals and the adjusted Pearson residuals with ridge estimator. The simulation results of the GRM influence diagnostics with ridge estimates using standardized Pearson residuals are presented in Table 3.10. From Table 3.10, we observe that with multicollinearity, the performance of MCD is a little better than the CD. Performance of all the GRM diagnostics are better with ridge estimates but most important are the AP and CVR method because these methods detect the influential observations as approximately 100%. While the poorer method is the H_d method. As the hr are smaller than the h values, thats why detection with ridge estimates smaller than without ridge estimates. Another result we observed that the performance of the CD, MCD and DFFITS with the ridge estimates are identical. The performance of the

GRM diagnostic methods increases with the increase in ϕ except the H_d method as shown in Table. Similar results are also seen with the ridge estimates. The influence detection percentages of CD method with and without ridge estimates increases as the multicollinearity increases. Similar results are observed from the other measures i.e. MCD, MCD_R, DFFITS and $DFFITS_R$ except H_d and H_{dR} methods. For the moderate multicollinear data with $\phi < 0.33$, the AP method detects larger influential observations than that of the AP_R. While higher degree of multicollinearity, the detection of influential observations of these methods is 100%. Increase in both multicollinearity and ϕ has a little positive effect on the AP and AP_R influential detections. Similar results are also seen for the CVR and CVR_R methods. While for $\phi < 0.11$,performance of the both WD and WD_R seems to be identical in detecting the influential observation. A little bit effect of the multicollinearity on WD and WD_R is observed and it is found that performance of the both measures increases with increase in multicollinearity. These results reveal that the H_d and H_{dR} methods are not better for the influence diagnostics in the GRM multicollinear data.

On considering the influence diagnostics using the adjusted Pearson residuals, we observed that detection rate is higher than that of using the standardized Pearson residuals. These results are given Table 3.15. Another better results which we observed in the detection of H_{dR} increase with the increase in ϕ and is decrease with increase in multicollinearity. The detection (%) of the H_{dR} method is identical to other methods when $\phi \geqslant 10$. So the simulation results of the adjusted Pearson residuals with and without ridge estimator are better than the standardized Pearson residuals.

3.17.1.2 *Influence Diagnostics in the GRM and the GRRM with Deviance Residuals*

In this section, like Pearson residuals, we analyze the influence diagnostics with four forms of the deviance residuals. The simulation results of the GRM influence diagnostics with ridge estimates using standardized deviance residuals are given in Table 3.11. From Table 3.11, we observe that, the performance of all the GRM and the GRRM influence diagnostic measures are approximately identical except the H_ds method for $\phi > 0.04$. While for $\phi = 0.04$, the detection of CD and CD_R methods are 8% and 12%. respectively smaller than the detection of MCD, AP, CVR, DFFITS and WD methods along with the ridge estimator. A moderate positive effect of multicollinearity have observed on the CD and the CD_R influence detection methods. For severe multicollinearity $\rho = 0.99$ and $\phi > 0.04$, the detection rate of all the diagnostic measures are identical except the H_ds method. The simulation results are also indicated that the detection performance with standardized deviance residuals are better than the standardized Pearson residuals.

Now considering the influence diagnostics with the adjusted deviance residuals, we observed that influence detection is higher than that of using standardized Pearson residuals and the standardized deviance residuals. These results are given Table 3.16. Multicollinearity has positive effect on the influence diagnostics only for dispersion value i.e. $\phi = 0.04$. The remaining values of the dispersion parameter has similar effect on the influence diagnostics as of computed with the standardized deviance residuals. So the simulation results of the adjusted deviance residuals with and without ridge estimator are better than the standardized deviance residuals.

3.17.1.3 *Influence Diagnostics in the GRM and the GRRM with Likelihood Residuals*

The simulation results of the GRM influence diagnostics with ridge estimates using likelihood residuals are presented in Table 3.12. From Table 3.12, we observe that with multicollinearity, the performance of CD is not good as compare to the CD_R. Performance of the CD method becomes poor with the increase in ϕ . While with the ridge estimates, the influence detection increases with the increase in ϕ . The influence detection percentages of CD decreases as the multicollinearity increases while the CD_R detection percentages increases with increase of multicollinearity level. Similar results are observed from the other measures i.e. MCD, MCD_R, DFFITS and $DFFITS_R$. For the multicollinear data, the AP method detects larger influential observations than that of the AP_R. Increases in both multicollinearity and ϕ has a little positive effect on the AP and AP_R influential detections. Similar results are also seen for the CVR and CVR_R methods. For $\phi < 0.11$, the influential observation detection of the WD method is larger than that of the WD_R. While for $\phi \geqslant 0.11$, performance of the both WD and WD_R seems to be identical in detecting the influential observation. A little bit effect of the multicollinearity on WD and WD_R is observed and it is found that performance of the both measures increases with increase in multicollinearity. The detection of influential observation by H_d and H_{dR} methods are very poor than the other GRM and GRRM diagnostic methods.

A positive effect of the multicollinearity and ϕ are observed on the H_d method in detecting the influential observation. While negative effect of multicollinearity and ϕ are seen on the H_{dR} method. These results reveal that the H_d and H_{dR} methods with likelihood residuals are not better for the influence diagnostics.

Now the influence diagnostics with the adjusted likelihood residuals of the CD and CD_R methods are totally opposite to the likelihood residuals as indicated in Table 3.17. These methods diagnose well the influential observation with increase in dispersion and identical for the severe multicollinearity. The remaining diagnostic methods has similar performance as with the likelihood residuals.

3.17.1.4 *Influence Diagnostics in the GRM and the GRRM with Anscombe Residuals*

The influence diagnostics of the GRM procedures with the standardized Aonscombe residuals are similar to the standardized Pearson residuals but better than the likelihood residuals. While with the ridge estimates of the Anoscombe residuals, GRRM diagnostics are not better than the GRRM with the standardized Pearson residuals. We also find that the influence diagnostic performance of the MCD_R and $DFFITS_R$ seems to be identical. These results are given in Table 3.13. The results with the Adjusted Anscombe residuals are identical to the results computed with the adjusted likelihood residuals, the adjusted deviance and the adjusted Pearson residuals. These results are given in Table 3.18.

3.17.1.5 *Influence Diagnostics in the GRM and the GRRM with Working Residuals*

The influential observation detection of the GRM and GRRM procedures using standardized working residuals are so interesting. From Table 3.14, we find that influence detection using the CD is smaller than the CD_R method for $\phi = 0.04$. For this dispersion and low multicollinearity (0.75), other the GRM and GRRM procedures perform better than CD method. For all multicollinearity levels and $\phi \geqslant 0.67$, the CD_R and MCD_R methods influence detection decrease. For all dispersion values and multicollinearity levels, other the GRM and GRRM influence diagnostic measures detect 100%. This indicated that the multicollinearity and dispersion does not affect the influential observation detection by AP, AP_R, CVR, CVR_R, WD and WD_R methods.

The GRM and the GRRM influence diagnostics using the adjusted working residuals are given in Table 3.18. From Table 3.18, we find that the detection by AP, AP_R, CVR, CVR_R, WD and WD_R methods are similar to that of using the standardized working residuals with additional results of CD_R and MCD_R which are better with the adjusted working residuals. Another interesting result which we observed that the detection by the H_{dR} method increases with increases dispersion values. We also find that with the increase in multicollinearity results that the H_{dR} detect in a better way. For severe multicollinearity and $\phi \geqslant 2$, the results of H_{dR} method are identical to other the GRM and the GRRM influence diagnostic methods.

3.17.2 Application: Reaction Rate data

Reaction rate data set is well defined in chapter2 and is reproduced in Appendix Table A1. This data set is well fitted to the gamma distribution as tested by the Goodness of Fit Tests and the results are given in Table 3.3. Reaction rate data is multicollinear as we tested by condition index (CI) as

$$CI = \sqrt{\frac{\max(\lambda_j)}{\min(\lambda_j)}}\ ,\ j = 1, 2, ..., p,$$

where $\lambda_1, \lambda_2, ..., \lambda_p$ are the eigenvalues of $X^T W X$. The value of CI of the reaction rate data 1795 shows that the data is multicollinear. From Table 3.3, we find that this data is well fitted to the gamma distribution. Thats the reason to use this data for the gamma regression influence diagnostics. The influential observation are detected in the GRM and GRRM with five types of residuals by seven methods. Each residual is then categorized to standardize and the adjusted forms of residuals and then these seven diagnostic methods are computed for each form of residuals.

Table 3.3: Distribution goodness of fit tests for reaction rate data

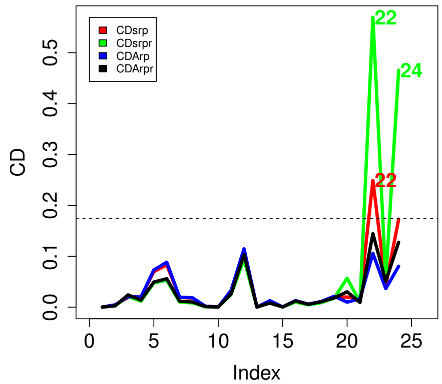
		Probability Distributions				
Goodness of Fit Tests		Normal	Exponential	Gamma	IG	Weibull
Anderson-darling	Statistic	1.2462	0.7129	0.2519	0.7926	0.2973
	p-value	0.0027	0.2624	**0.7538**	0.0965	0.6288
Cramer-Von Mises	Statistic	0.2127	0.1265	0.0432	0.1503	0.0521
	p-value	0.0033	0.2115	*0.6259*	0.0859	0.4772
Pearson	Statistic	10.6670	6.6667	2.0000	7.3333	6.0000
	p-value	0.0584	0.3528	**0.8491**	0.1970	0.3062

Table 3.4: Influential Observations with the GRM and the GRRM Pearson Residuals

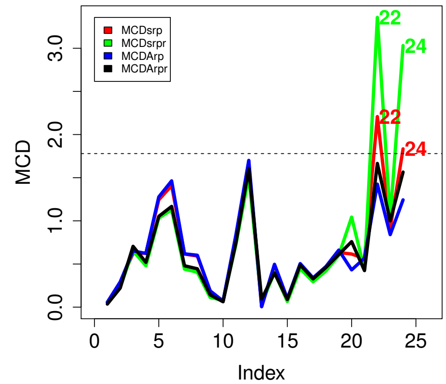
	Standardized Pearson Residuals		Adjusted Pearson Residuals	
Methods	sr_{Pi}	sr_{PRi}	Ar_{Pi}	Ar_{PRi}
CD	22	22,24	Nil	Nil
MCD	22,24	22,24	Nil	Nil
AP	1,13,20,22,24	1,9,13,15,20,22,24	1,13,20,22,24	1,13,20,22,24
CVR	12,13,20,22	13,20,22,23,24	12,13,20,22,23,24	6,12,13,20,22,23,24
DFFITS	22	22,24	Nil	Nil
WD	22,24	20,22,24	Nil	Nil
H_d	20,22,24	22,24	20,22,24	22,24

From Table 3.4 and Fig. 3.2 (a-g), we find that four forms of Pearson residuals the observation 1$^{\text{st}}$, 6$^{\text{th}}$, 9$^{\text{th}}$, 12$^{\text{th}}$, 13$^{\text{th}}$, 15$^{\text{th}}$, and 22-24$^{\text{th}}$ are identified as influential observations by various procedures. These detection with standardized Pearson residuals, CD and DFFITS identified the 22$^{\text{nd}}$ observation as influential observation. Similarly the performance of MCD and WD are identical and detected that 22$^{\text{nd}}$ and 24$^{\text{th}}$ observations are the influential observations. The H_d method using standardized Pearson residuals additionally identified the 20$^{\text{th}}$ observation as influential observation. The AP and CVR additionally detected that the observation 1$^{\text{st}}$, 12$^{\text{th}}$ and 13$^{\text{th}}$ are influential observations. These detections with the ridge Pearson residuals are similar with the addition of the observation 9$^{\text{th}}$, 15$^{\text{th}}$ and 23$^{\text{rd}}$ which are identified by the AP and CVR methods. With the ridge Pearson residuals, the diagnostics methods CD, MCD, DFFITS and H_d identified the similar observation (20$^{\text{th}}$, 22$^{\text{nd}}$ and 24$^{\text{th}}$) are the influential observations. The CD, MCD, DFFITS and WD methods with the adjusted Pearson residuals and the adjusted Pearson residuals with ridge estimates are unable to detect any influential observation. On the other hand, the AP, CVR and H_d methods successfully diagnose the influential observation as already detected with the standardized Pearson residuals. Additionally 6^{th} observation detected by the CVR method using the adjusted Pearson residuals with the ridge estimates.

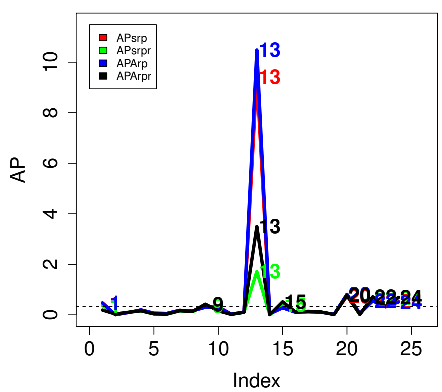
From Table 3.5 and Fig. 3.2 (a-g), we find with four forms of deviance residuals that the observation 1^{st}, 5^{th}, 6^{th}, 9^{th}, 10^{th}, 13^{th}, 20^{th} and 22-24^{th} are identified as the influential observations by various procedures. These detection with the standardized deviance residuals, all methods detects different influential observations but most common in all diagnostic methods is the 22^{nd} observation. The influence diagnostics with ridge deviance residuals indicated that the detection performance of MCD and DFFITS are identical and detected that 6^{th}, 22^{nd} and 24^{th} observations are influential observations. With the ridge deviance residuals, CD, AP, WD and H_d methods are unable to detect 6^{th} observation as influential observation. The detection of influential observation with adjusted deviance residuals shows that MCD and DFFITS methods detected the 6th observation as influential while the CD and WD methods with adjusted deviance residuals are unable to detect any of the influential observation. Similarly with the adjusted ridge deviance residuals, CD, MCD and DFFITS are fail to detect any influential case. While H_d, AP, CVR and WD detected the similar influential observations as detected by the standardized Pearson and standardized deviance residuals as already detected.



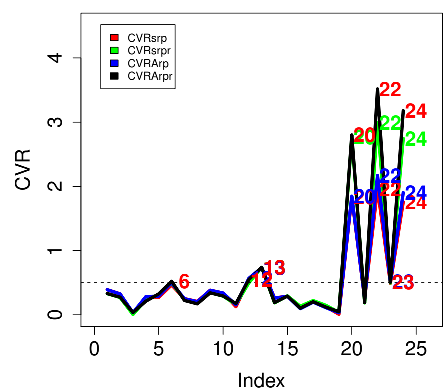
(a) GRMs CD Index plots



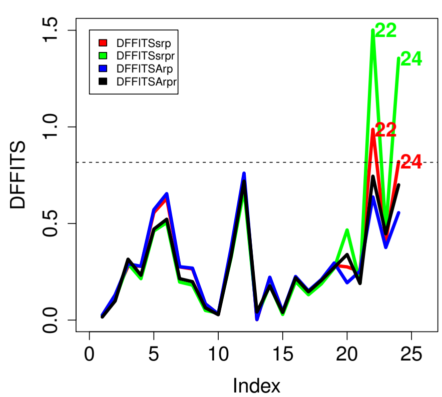
(b) GRMs MCD Index plots



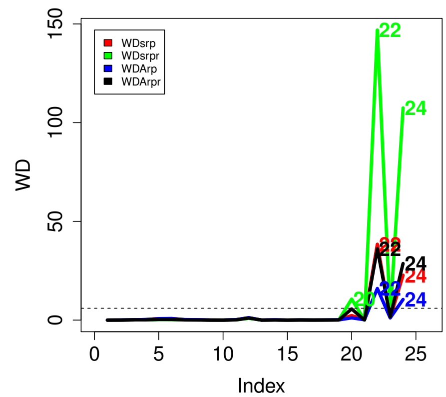
(c) GRMs AP Index plots



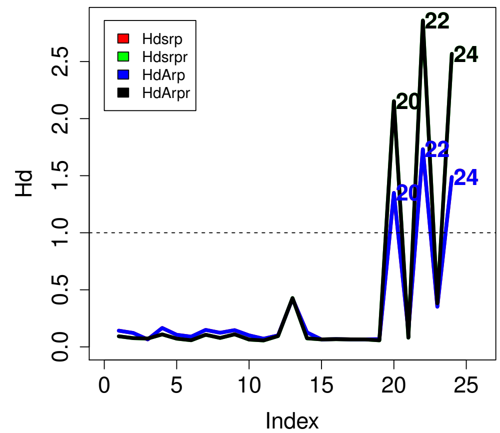
(d) GRMs CVR Index plots



(e) GRMs DFFITS Index plots



(f) GRMs WD Index plots



(a) GRMs H_d Index plot

Figure 3.2: The GRMs Influence Diagnostic Methods Index Plots with Pearson Residuals

Table 3.5: Influential Observations with the GRM and the GRRM Deviance Residuals

	Standardized Deviance Residuals		Adjusted Deviance Residuals	
Methods	sr_{di}	sr_{dRi}	Ar_{di}	Ar_{dRi}
CD	6,22,24	22,24	Nil	Nil
MCD	5,6,22,24	6,22,24	6	Nil
AP	1,13,20,22,24	1,9,13,20,22,23,24	1,10,13,20,22,24	1,10,13,20,22,24
CVR	5,6,13,20,22,23,24	5,6,13,20,22,23,24	6,13,20,22,23,24	6,13,20,22,23,24
$DFFITS$	6,22	6,22,24	6	Nil
WD	22,24	20,22,24	Nil	20,22,24
H_d	20,22,24	22,24	20,22,24	22,24

The GRM diagnostics with the likelihood residuals detected the similar observations as identified by the Pearson and deviance residuals but in more consistent way. These diagnostics with the likelihood ridge residuals additionally detected 12$^{\text{th}}$ observation as influential observation diagnosed by the MCD and CVR methods. From Table 3.6 and Fig. 3.2 (a-g), we also find that CD method with the adjusted likelihood residuals is unable to detect any influential observation.

Table 3.6: Influential Observations with the GRM and the GRRM Likelihood Residuals

	Likelihood Residuals		Adjusted Likelihood Residuals	
Methods	r_{li}	r_{lRi}	Ar_{li}	Ar_{lRi}
CD	6,22,24	5,6,22,24	Nill	22,24
MCD	5,6,22,24	5,6,12,22,24	6	22,24
AP	1,13,20,22,24	13,20,24	1,10,13,20,22,24	9,13,15,20,22,24
CVR	5,6,13,20,22-24	5,6,11-13,19,20,22,24	6,13,20,22,24	6,13,20,22,24
DFFITS	6,22	5,6,22,24	6	22
WD	22,24	20,22,24	22,24	20,22,24
H_d	20,22,24	20,22,24	20,22,24	20,22,24

Table 3.7: Influential Observations with the GRM and the GRRM Anscombe Residuals

	Standardized Anscombe Residuals		Adjusted Anscombe Residuals	
Methods	sr_{ai}	sr_{aRi}	Ar_{ai}	Ar_{aRi}
CD	6,22,24	22,24	Nil	Nil
MCD	5,6,22,24	6,22,24	6	Nil
AP	1,13,20,22,24	1,9,13,15,20,22,23,24	1,10,13,20,22,24	1,10,13,20,22,24
CVR	5,6,13,20,22,23,24	5,6,13,20,22,23,24	6,13,20,22,23,24	6,13,20,22,23,24
DFFITS	6,22	6,22,24	6	Nil
WD	22,24	20,22,24	22	20,22,24
H_d	20,22,24	22,24	20,22,24	22,24

While the other methods detect the influential observations but superior are the AP and CVR methods. The influential observation detection with the adjusted ridge likelihood residuals are better than the likelihood and adjusted likelihood residuals. The influential observation detected by the GRM diagnostics with four forms of Anscombe residuals are given in Table 3.7 and Fig. 3.2 (a-g). We find that similar detection of influential observation by the Anscombe residuals as with the deviance residuals except the with the WD method using adjusted Anscombe residuals. More

Table 3.8: Influential Observations with the GRM and the GRRM Working Residuals

	Standardized Working Residuals		Adjusted Working Residuals	
Method	sr_{Wi}	sr_{WRi}	Ar_{Wi}	Ar_{WRi}
CD	Nil	Nil	Nil	Nil
MCD	6	Nil	6	Nil
AP	1,10,13,20,22,24	1,10,13,20,22,24	1,10,22,24	1,10,22,24
CVR	6,13,20,22,23,24	6,13,20,22,23,24	6,13,20,22,23,24	6,13,20,22,23,24
DFFITS	6	Nil	6	Nil
WD	22	20,22,24	Nil	Nil
H_d	20,22,24	22,24	20,22,24	22,24

influential observations are diagnosed by the AP and CVR methods. The influential observation detection using four forms of working residuals are given in Table 3.8. From Table 3.8 and Fig 3.2 (a-g), we find detection of influential observations are approximately similar to that of using the Pearson residuals. With the addition using adjusted working residuals, the MCD and DFFITS methods detected the observation 6$^{\text{th}}$ as influential observation which was not diagnosed using the adjusted Pearson residuals.

Now we check the actual effect of these detected influential observations with different residuals on the GRRM estimates and inferences. These effects are given in Table 3.9 and now we compare the diagnostic methods with GRM different residuals to determine which one detects the influential observation accurately. And also compare the performance of all GRM residuals in detecting the influential observation. From Table 3.9, we find that top most influential observation is the 22$^{\text{nd}}$ as detected by the all methods with all forms of the GRM and the GRRM residuals except the working residuals. This observation is better detected with four forms of the likelihood residuals. This observation affects the GRM and the GRRM estimates of β_1. The

2nd most influential observation is the 12th observation as diagnosed by the MCD

Table 3.9: Absolute Change (%) in the GRM and the GRRM Estimates after Deleting Influential Observations

Influential	GRM Estimates				GRRM Estimates			
Observations	β_0	β_1	β_2	β_3	β_{R0}	β_{R1}	β_{R2}	β_{R3}
1	0.12	0.47	0.55	0.07	0.38	0.6	0.96	0.53
5	3.98	8.39	6.93	**10.4**	3.58	8.39	6.97	**9.95**
6	4.81	7.94	3.34	**10.76**	4.09	8.31	2.74	**9.28**
9	0.54	2.77	1.28	1.50	0.08	0.93	0.22	0.23
10	0.24	1.25	0.01	0.30	0.68	2.45	0.19	1.27
11	3.00	0.54	**5.90**	4.42	2.79	0.44	**6.06**	3.67
12	**5.66**	9.64	0.51	**14.61**	**5.68**	9.13	0.33	**16.56**
13	0.00	0.05	0.02	0.05	0.02	1.73	0.63	2.42
15	0.41	0.42	0.41	0.75	0.12	0.10	0.09	0.51
19	2.35	0.76	**4.92**	2.32	2.28	0.85	**5.16**	1.79
20	0.09	**11.49**	3.24	1.37	0.03	**15.57**	4.4	3.03
21	1.27	**8.38**	2.21	1.83	1.35	**8.06**	2.36	2.07
22	0.29	**31.98**	**16.35**	4.96	0.26	**21.41**	**11.82**	4.23
23	0.68	3.12	4.96	**7.34**	0.83	3.18	5.84	**9.85**
24	0.29	**13.45**	7.85	4.44	0.28	**12.67**	8.36	4.42

and CVR methods using likelihood residuals with ridge estimates. This observation affects on the GRM and the GRRM estimates of β_3. The 3rd influential observation is the 20th observation as identified by all the diagnostic methods except the CD, MCD and DFFITS methods and better detected using likelihood residuals with ridge estimates. Similarly the observation 24th is also 4th positional influential observation. While observations 5th, 6th, 21st and 23rd are ranked at 5th positional influential observations. These are identified by various procedures and all form of the GRM and the GRRM residuals except with the Pearson residuals. we also find that 1st, 9th, 10th, 13th and 5th observations has minorsubstantiall affect on the GRM and the GRRM estimates. These observations are detected by the AP and CVR methods with all forms of the GRM and the GRRM residuals. These results are indicating that these methods are detected wrong influential observations especially the AP method.

Table 3.10: Performance of the GRM and the GRRM Influence Measures with Standardized Pearson Residuals

ρ	ϕ	CD	MCD	AP	CVR	DFFITS	WD	H_d	CD_R	MCD_R	AP_R	CVR_R	$DFFITS_R$	WD_R	H_{dR}
0.75	0.04	90.6	100	97.3	100	100	100	21.1	80.2	80.2	92.9	95.7	80.2	90	24.7
	0.11	97.9	97.9	99.8	99.8	97.9	99.6	10.9	95.6	95.6	99.5	99.8	95.6	99.5	20.2
	0.33	99.9	99.9	100	100	99.9	100	2.6	99.6	99.6	100	100	99.6	100	7.7
	0.67	99.4	99.4	100	100	99.4	100	2	98.9	98.9	100	100	98.9	100	5.4
	2.00	100	100	100	100	100	100	1.9	99.5	99.5	100	100	99.5	100	4
	4.00	99.8	99.8	100	100	99.8	100	2.4	99.8	99.8	100	100	99.8	100	3.3
	10.0	99.9	99.9	100	100	99.9	100	1.9	100	100	100	100	100	100	2.8
0.85	0.04	94.4	100	98.6	100	100	100	22	85.8	85.9	94.8	96.1	85.9	93.4	25
	0.11	98.3	98.3	100	100	98.3	99.5	9.6	95	95	100	99.9	95	99.1	17.6
	0.33	99.8	99.8	100	100	99.8	100	3.5	99.4	99.4	100	100	99.4	100	8
	0.67	100	100	100	100	100	100	2.4	99.5	99.5	100	100	99.5	100	5.2
	2.00	99.7	99.7	100	100	99.7	100	1.6	99.8	99.8	100	100	99.8	100	5.1
	4.00	99.9	99.9	100	100	99.9	100	1.2	99.9	99.9	100	100	99.9	100	3.3
	10.0	99.9	99.9	100	100	99.9	100	1.8	99.9	99.9	100	100	99.9	100	3
0.95	0.04	95	100	99.4	100	100	100	21	89.2	89.2	97.5	98.3	89.2	95.6	24.6
	0.11	98.7	98.7	99.9	99.9	98.7	99.9	7	98	98	99.9	99.9	98	99.8	12.4
	0.33	99.9	99.9	100	100	99.9	100	2.2	99.2	99.2	100	100	99.2	100	6.3
	0.67	99.9	99.9	100	100	99.9	100	2.7	99.4	99.4	100	100	99.4	100	4.7
	2.00	99.8	99.8	100	100	99.8	100	2.3	99.7	99.7	100	100	99.7	100	2.8
	4.00	99.8	99.8	100	100	99.8	100	3.5	100	100	100	100	100	100	3
	10.0	100	100	100	100	100	100	2.1	99.9	99.9	100	100	99.9	100	2.5
0.99	0.04	96.6	96.6	99.7	99.2	96.6	98.4	19.9	94.8	94.8	99.2	99.5	94.8	98	21.9
	0.11	99.5	99.5	100	100	99.5	100	8	99	99	100	100	99	100	9.2
	0.33	99.7	99.7	100	100	99.7	100	2.9	99.2	99.2	100	100	99.2	100	5.4
	0.67	**100**	100	100	100	100	100	3.2	99.7	99.7	100	100	99.7	100	3.5
	2.00	**99.8**	99.8	100	100	99.8	100	3	99.8	99.8	100	100	99.8	100	3.2
	4.00	99.9	99.9	100	100	99.9	100	2.6	99.8	99.8	100	100	99.8	100	2.9
	10.0	99.8	99.8	100	100	99.8	100	2.4	99.9	99.9	100	100	99.9	100	2.8

Table 3.11: Performance of the GRM and the GRRM Influence Measures with Standardized Deviance Residuals

ρ	ϕ	CD	MCD	AP	CVR	DFFITS	WD	H_d	CD_R	MCD_R	AP_R	CVR_R	$DFFITS_R$	WD_R	H_{dR}
0.75	0.04	91.7	100	97.8	100	100	100	24	87.2	100	92.9	100	100	100	38.2
	0.11	98.6	100	100	100	100	100	9.2	97.3	97.3	99.8	100	97.3	99.3	23
	0.33	99.8	99.8	100	100	99.8	100	1.9	99.3	99.3	100	100	99.3	100	9.7
	0.67	99.9	99.9	100	100	99.9	100	1.6	99.5	99.5	100	100	99.5	100	4.6
	2.00	99.8	99.8	100	100	99.8	100	1.5	99.7	99.7	100	100	99.7	100	3
	4.00	100	100	100	100	100	100	1.2	99.4	99.4	100	100	99.4	100	2.9
	10.0	100	100	100	100	100	100	1.9	100	100	100	100	100	100	1.9
0.85	0.04	95.2	100	98.6	100	100	100	21.7	89.8	100	96.3	100	100	100	32.2
	0.11	99.8	99.8	100	100	99.8	100	9.1	97.9	97.9	99.9	99.9	97.9	99.6	20
	0.33	99.8	99.8	100	100	99.8	100	3.2	99.4	99.4	100	100	99.4	100	8.3
	0.67	99.9	99.9	100	100	99.9	100	1.6	99.6	99.6	100	100	99.6	100	4.5
	2.00	99.8	99.8	100	100	99.8	100	1.5	99.6	99.6	100	100	99.6	100	2.7
	4.00	99.9	99.9	100	100	99.9	100	0.8	99.8	99.8	100	100	99.8	100	1.5
	10.0	99.9	99.9	100	100	99.9	100	2.2	99.7	99.7	100	100	99.7	100	2.6
0.95	0.04	95.9	100	99.3	100	100	100	20.9	93.5	93.6	98.4	97.6	93.6	98	31.2
	0.11	99.4	99.4	100	99.9	99.4	100	6.5	99.5	99.5	100	100	99.5	99.9	15.3
	0.33	99.8	99.8	100	100	99.8	100	3	99.7	99.7	100	100	99.7	100	5.1
	0.67	99.7	99.7	100	100	99.7	100	1.5	100	100	100	100	100	100	5.2
	2.00	99.8	99.8	100	100	99.8	100	1.5	99.8	99.8	100	100	99.8	100	4.1
	4.00	99.8	99.8	100	100	99.8	100	2.2	99.7	99.7	100	100	99.7	100	2.6
	10.0	99.8	99.8	100	100	99.8	100	1.9	99.8	99.8	100	100	99.8	100	1.8
0.99	0.04	97.5	97.5	99.9	99.7	97.5	99.2	17.9	96	96	99.3	99.3	96	98.4	22.8
	0.11	99.5	99.5	100	100	99.5	100	7.7	98.5	98.5	100	100	98.5	100	10.7
	0.33	99.4	99.4	100	100	99.4	100	2.3	100	100	100	100	100	100	4.6
	0.67	99.8	99.8	100	100	99.8	100	3.1	99.8	99.8	100	100	99.8	100	3.2
	2.00	99.9	99.9	100	100	99.9	100	2	99.7	99.7	100	100	99.7	100	2.3
	4.00	99.9	99.9	100	100	99.9	100	2.5	99.9	99.9	100	100	99.9	100	2.6
	10.0	99.8	99.8	100	100	99.8	100	2.7	99.8	99.8	100	100	99.8	100	3.2

Table 3.12: Performance of the GRM and the GRRM Influence Measures with Likelihood Residuals

ρ	ϕ	CD	MCD	AP	CVR	DFFITS	WD	H_d	CD_R	MCD_R	AP_R	CVR_R	$DFFITS_R$	WD_R	H_{dR}
0.75	0.04	71.3	71.3	97.8	97.6	71.3	96.4	35.5	86	86.7	92.7	94	86.7	94.2	38.1
	0.11	57.6	57.6	100	100	57.6	99.1	44.7	97.3	97.3	99.7	100	97.3	99.3	22.8
	0.33	35	34.9	100	100	34.9	99.8	48.7	99.3	99.3	100	100	99.3	100	9.7
	0.67	18.5	18.5	100	100	18.5	99.9	49.2	99.5	99.5	100	100	99.5	100	4.6
	2.00	2.1	2.1	100	100	2.1	99.7	50.4	99.7	99.7	100	100	99.7	100	3
	4.00	0	0	100	100	0	99.9	51.3	99.4	99.4	100	100	99.4	100	2.9
	10.0	0	0	100	100	0	99.3	48.9	100	100	100	100	100	100	1.9
0.85	0.04	71.1	71	98.6	98.3	71	98.2	38.1	89.1	89.1	96.3	96.2	89.1	96.4	31.6
	0.11	52.8	52.8	100	100	52.8	99.9	41.1	97.9	97.9	99.9	99.9	97.9	99.5	20.2
	0.33	30.1	30	100	100	30	99.8	45.4	99.4	99.4	100	100	99.4	100	8.2
	0.67	14.3	14.2	100	100	14.2	99.9	47.7	99.6	99.6	100	100	99.6	100	4.5
	2.00	1.1	1.1	100	100	1.1	99.8	48.9	99.6	99.6	100	100	99.6	100	2.7
	4.00	0	0	100	100	0	99.9	53.1	99.8	99.8	100	100	99.8	100	1.5
	10.0	0	0	100	100	0	99.3	48.7	99.7	99.7	100	100	99.7	100	2.6
0.95	0.04	66.5	66.5	99.3	99.1	66.5	97.8	40.7	93.1	93.1	98.4	97.9	93.1	98	30.1
	0.11	47.8	47.8	100	100	47.8	99.7	43.1	99.5	99.5	100	99.9	99.5	99.9	15.3
	0.33	22.5	22.5	100	100	22.5	99.9	47.6	99.7	99.7	100	100	99.7	100	5.1
	0.67	9.1	9.1	100	100	9.1	99.6	48.4	100	100	100	100	100	100	5.2
	2.00	0.4	0.4	100	100	0.4	99.7	48.2	99.8	99.8	100	100	99.8	100	4.1
	4.00	0	0	100	100	0	99.2	50.5	99.7	99.7	100	100	99.7	100	2.6
	10.0	0	0	100	100	0	98.8	49.9	99.8	99.8	100	100	99.8	100	1.8
0.99	0.04	59.8	59.8	99.9	99.7	59.8	98.8	39.6	95.9	95.9	99.3	99	95.9	98.4	22.9
	0.11	42.2	42.2	100	100	42.2	99.9	46.2	98.5	98.5	100	99.9	98.5	100	10.7
	0.33	17.3	17.3	100	100	17.3	99.2	49.3	100	100	100	100	100	100	4.6
	0.67	5	5	100	100	5	99.6	50.7	99.8	99.8	100	100	99.8	100	3.2
	2.00	0.1	0.1	100	100	0.1	99.6	51.8	99.7	99.7	100	100	99.7	100	2.3
	4.00	0	0	100	100	0	99.5	47.7	99.9	99.9	100	100	99.9	100	2.6
	10.0	0	0	100	100	0	98.9	49.4	99.8	99.8	100	100	99.8	100	3.2

Table 3.13: Performance of the GRM and the GRRM Influence Measures with Standardized Anscombe Residuals

ρ	ϕ	CD	MCD	AP	CVR	DFFITS	WD	H_d	CD_R	MCD_R	AP_R	CVR_R	$DFFITS_R$	WD_R	H_{dR}
	0.04	90.6	100	97.3	100	100	100	21.1	84.9	85.2	93.6	92	85.2	93.8	33.4
	0.11	97.9	97.9	99.8	99.8	97.9	99.6	10.9	96.2	96.2	99.7	99.1	96.2	99.5	22.5
	0.33	99.9	99.9	100	100	99.9	100	2.6	99.6	99.6	100	100	99.6	100	8.1
0.75	0.67	99.4	99.4	100	100	99.4	100	2	98.8	98.8	100	100	98.8	100	5.5
	2.00	100	100	100	100	100	100	1.9	99.5	99.5	100	100	99.5	100	3.9
	4.00	99.8	99.8	100	100	99.8	100	2.4	99.8	99.8	100	100	99.8	100	3.3
	10.0	99.9	99.9	100	100	99.9	100	1.9	100	100	100	100	100	100	2.9
	0.04	94.4	100	98.6	100	100	100	22	89.5	89.6	95	93.6	89.6	96	34.4
	0.11	98.3	98.3	100	100	98.3	99.5	9.6	95.2	95.2	100	100	95.2	99.1	19.9
	0.33	99.8	99.8	100	100	99.8	100	3.5	99.4	99.4	100	100	99.4	100	8.4
0.85	0.67	100	100	100	100	100	100	2.4	99.5	99.5	100	100	99.5	100	5.5
	2.00	99.7	99.7	100	100	99.7	100	1.6	99.8	99.8	100	100	99.8	100	5.3
	4.00	99.9	99.9	100	100	99.9	100	1.2	99.9	99.9	100	100	99.9	100	3.4
	10.0	99.9	99.9	100	100	99.9	100	1.8	99.9	99.9	100	100	99.9	100	3
	0.04	95	100	99.4	100	100	100	21	91.4	91.5	97.8	96.4	91.5	97.1	31.5
	0.11	98.7	98.7	99.9	99.9	98.7	99.9	7	98	98	99.9	99.9	98	99.9	13.9
	0.33	99.9	99.9	100	100	99.9	100	2.2	99.2	99.2	100	100	99.2	100	6.9
0.95	0.67	99.9	99.9	100	100	99.9	100	2.7	99.4	99.4	100	100	99.4	100	5
	2.00	99.8	99.8	100	100	99.8	100	2.3	99.7	99.7	100	100	99.7	100	2.8
	4.00	99.8	99.8	100	100	99.8	100	3.5	99.9	99.9	100	100	99.9	100	3.2
	10.0	100	100	100	100	100	100	2.1	99.9	99.9	100	100	99.9	100	2.6
	0.04	96.6	96.6	99.7	99.2	96.6	98.4	19.9	95.5	95.5	99.4	98.9	95.5	98.2	24.4
	0.11	99.5	99.5	100	100	99.5	100	8	99	99	100	100	99	100	9.7
	0.33	99.7	99.7	100	100	99.7	100	2.9	99.2	99.2	100	100	99.2	100	5.7
0.99	0.67	100	100	100	100	100	100	3.2	99.7	99.7	100	100	99.7	100	3.6
	2.00	99.8	99.8	100	100	99.8	100	3	99.8	99.8	100	100	99.8	100	3.2
	4.00	99.9	99.9	100	100	99.9	100	2.6	99.8	99.8	100	100	99.8	100	2.9
	10.0	99.8	99.8	100	100	99.8	100	2.4	99.9	99.9	100	100	99.9	100	2.8

Table 3.14: Performance of the GRM and the GRRM Influence Measures with Standardized Working Residuals

ρ	ϕ	CD	MCD	AP	CVR	DFFITS	WD	H_d	CD_R	MCD_R	AP_R	CVR_R	$DFFITS_R$	WD_R	H_{dR}
	0.04	91.3	100	95.7	100	100	100	23.3	100	100	100	100	100	100	6.4
	0.11	98.9	98.9	99.9	99.8	98.9	99.7	9.2	99.9	100	100	100	100	100	5.2
	0.33	99.8	99.8	100	100	99.8	100	3.3	99.4	100	100	100	100	100	5.6
0.75	0.67	100	100	100	100	100	100	3	98.5	98.5	100	100	98.5	100	4.5
	2.00	100	100	100	100	100	100	2.1	92.9	92.9	100	100	92.9	100	3.4
	4.00	100	100	100	100	100	100	2.2	80.1	80.1	100	100	80.1	100	2.7
	10.0	99.9	99.9	100	100	99.9	100	1.6	20.4	20.2	100	100	20.2	99.7	1.8
	0.04	92.4	100	97.8	100	100	100	19.7	100	100	100	100	100	100	6
	0.11	99	99	100	99.9	99	99.9	7.7	100	100	100	100	100	100	5.5
	0.33	99.8	99.8	100	100	99.8	100	2.1	99.7	100	100	100	100	100	5.4
0.85	0.67	99.9	99.9	100	100	99.9	100	2	98.9	98.9	100	100	98.9	100	4.3
	2.00	99.9	99.9	100	100	99.9	100	1.7	92.6	92.6	100	100	92.6	99.9	4.5
	4.00	100	100	100	100	100	100	1.1	77.7	77.7	100	100	77.7	99.9	2.7
	10.0	99.9	99.9	100	100	99.9	100	1.2	14	14	100	100	14	99.5	2.6
	0.04	95.4	100	98.8	100	100	100	16.6	100	100	100	100	100	100	5.6
	0.11	99.4	99.4	100	100	99.4	100	6.7	100	100	100	100	100	100	4
	0.33	99.9	99.9	100	100	99.9	100	2.7	99.8	100	100	100	100	100	4.1
0.95	0.67	100	100	100	100	100	100	1.7	98.3	98.3	100	100	98.3	100	4.5
	2.00	99.7	99.7	100	100	99.7	100	3	90.2	90.2	100	100	90.2	99.9	3.3
	4.00	99.9	99.9	100	100	99.9	100	0.9	70.8	70.7	100	100	70.7	99.9	1.7
	10.0	99.8	99.8	100	100	99.8	99.9	1.1	3.5	3.4	100	100	3.4	99.7	1.4
	0.04	97.4	97.4	99.5	99.5	97.4	99.4	16.9	100	100	100	100	100	100	3.9
	0.11	99.2	99.2	100	100	99.2	99.8	6.8	99.9	100	100	100	100	100	4.6
	0.33	99.9	99.9	100	100	99.9	100	2.5	99.6	99.6	100	100	99.6	100	4.4
0.99	0.67	99.6	99.6	100	100	99.6	100	4	98.4	98.4	100	100	98.4	100	3.4
	2.00	99.6	99.6	100	100	99.6	100	2.3	85.2	85.2	100	100	85.2	100	3.1
	4.00	99.3	99.3	100	100	99.3	100	3.3	54	54	100	100	54	99.8	3.7
	10.0	99.7	99.7	100	100	99.7	100	3	0	0	100	100	0	99.3	3.1

Table 3.15: Performance of the GRM and the GRRM Influence Measures with Adjusted Pearson Residuals

ρ	ϕ	CD	MCD	AP	CVR	DFFITS	WD	H_d	CD_R	MCD_R	AP_R	CVR_R	$DFFITS_R$	WD_R	H_{dR}
0.75	0.04	88.2	88.5	97.7	98.5	88.5	94.7	22.9	68.6	68.6	92.2	94.6	68.6	85.7	33.5
	0.11	98.2	98.2	99.8	99.7	98.2	99.4	9.5	89.2	89.2	99.5	99.6	89.2	97.4	47.2
	0.33	100	100	100	100	100	100	2	98.5	98.5	100	100	98.5	99.9	60
	0.67	100	100	100	100	100	100	0.1	99.9	99.9	100	100	99.9	100	65.5
	2.00	100	100	100	100	100	100	0	100	100	100	100	100	100	88.4
	4.00	100	100	100	100	100	100	0	100	100	100	100	100	100	98.3
	10.0	100	100	100	100	100	100	0	100	100	100	100	100	100	99.9
0.85	0.04	89.8	89.8	98	98.4	89.8	95.2	22.4	70.6	70.6	93.9	96.4	70.6	86.6	32
	0.11	99.2	99.2	100	100	99.2	99.5	8.6	91.5	91.5	99.9	100	91.5	98.7	47.7
	0.33	100	100	100	100	100	100	1.3	98.9	98.9	100	100	98.9	99.9	60.6
	0.67	100	100	100	100	100	100	0.1	100	100	100	100	100	100	70.7
	2.00	100	100	100	100	100	100	0	100	100	100	100	100	100	88.6
	4.00	100	100	100	100	100	100	0.1	100	100	100	100	100	100	99.8
	10.0	100	100	100	100	100	100	1.1	100	100	100	100	100	100	100
0.95	0.04	94.2	94.2	99.2	99.3	94.2	97.5	20.5	78.4	78.4	97.2	98.3	78.4	93.9	39.5
	0.11	99.4	99.4	100	100	99.4	99.9	6.4	94.1	94.1	100	100	94.1	99.4	53.7
	0.33	100	100	100	100	100	100	1	99.4	99.4	100	100	99.4	100	70.4
	0.67	100	100	100	100	100	100	0.1	99.8	99.8	100	100	99.8	100	81.8
	2.00	100	100	100	100	100	100	0.2	100	100	100	100	100	100	98.6
	4.00	100	100	100	100	100	100	0.7	99.9	99.9	100	100	99.9	100	99.8
	10.0	100	100	100	100	100	100	20.3	99.6	99.6	100	100	99.6	100	99.4
0.99	0.04	95.4	95.4	99.9	99.5	95.4	98.1	18.1	83.3	83.3	99.3	99.8	83.3	96.5	48.8
	0.11	99.7	99.7	100	100	99.7	100	6.3	95.1	95.1	100	100	95.1	99.6	66.8
	0.33	100	100	100	100	100	100	1.4	98.8	98.8	100	100	98.8	100	87.4
	0.67	100	100	100	100	100	100	0.5	99.6	99.6	100	100	99.6	100	96.8
	2.00	100	100	100	100	100	100	2.3	99.9	99.9	100	100	99.9	100	99.8
	4.00	100	100	100	100	100	100	23.3	98.3	98.3	100	100	98.3	100	98.3
	10.0	100	100	100	100	100	100	95.7	93.5	93.5	100	100	93.5	100	92.7

Table 3.16: Performance of the GRM and the GRRM Influence Measures with Adjusted Deviance Residuals

ρ	ϕ	CD	MCD	AP	CVR	DFFITS	WD	H_d	CD_R	MCD_R	AP_R	CVR_R	$DFFITS_R$	WD_R	H_{dR}
	0.04	89	89.4	98	98	89.4	95	24	83	82.7	93	94.9	82.7	90	24
	0.11	99	98.6	100	100	98.6	100	8.8	98	97.5	100	99.9	97.5	99	11
	0.33	100	99.9	100	100	99.9	100	1.6	100	99.8	100	100	99.8	100	3.4
0.75	0.67	100	100	100	100	100	100	0.1	100	100	100	100	100	100	0.7
	2.00	100	100	100	100	100	100	0	100	100	100	100	100	100	0
	4.00	100	100	100	100	100	100	0	100	100	100	100	100	100	0
	10.0	100	100	100	100	100	100	0.5	100	100	100	100	100	100	0
	0.04	89	89	98	98	89	96	21	85	84.7	94	95	84.7	90	23
	0.11	99	99.1	100	100	99.1	100	8	98	97.8	100	100	97.8	99	12
	0.33	100	99.9	100	100	99.9	100	1.4	100	99.9	100	100	99.9	100	2.4
0.85	0.67	100	100	100	100	100	100	0	100	100	100	100	100	100	0.5
	2.00	100	100	100	100	100	100	0	100	100	100	100	100	100	0
	4.00	100	100	100	100	100	100	0	100	100	100	100	100	100	0
	10.0	100	100	100	100	100	100	0.7	100	100	100	100	100	100	0
	0.04	93	93.2	99	99	93.2	97	21	90	90.5	98	97.7	90.5	95	24
	0.11	100	99.7	100	100	99.7	100	6.8	100	99.5	100	100	99.5	100	9.2
	0.33	100	100	100	100	100	100	0.6	100	100	100	100	100	100	1
0.95	0.67	100	100	100	100	100	100	0.1	100	100	100	100	100	100	0.3
	2.00	100	100	100	100	100	100	0.3	100	100	100	100	100	100	0.1
	4.00	100	100	100	100	100	100	0.8	100	100	100	100	100	100	0
	10.0	100	100	100	100	100	100	21	100	100	100	100	100	100	2.3
	0.04	96	95.5	100	100	95.5	98	17	95	94.7	99	99.7	94.7	97	18
	0.11	100	99.8	100	100	99.8	100	5.9	100	99.6	100	100	99.6	100	6.5
	0.33	100	100	100	100	100	100	2.1	100	100	100	100	100	100	2.4
0.99	0.67	100	100	100	100	100	100	1.5	100	100	100	100	100	100	1.1
	2.00	100	100	100	100	100	100	3.4	100	100	100	100	100	100	0.9
	4.00	100	100	100	100	100	100	25	100	100	100	100	100	100	3.6
	10.0	100	100	100	100	100	100	97	100	100	100	100	100	100	65

Table 3.17: Performance of the GRM and the GRRM Influence Measures with Adjusted Likelihood Residuals

ρ	ϕ	CD	MCD	AP	CVR	DFFITS	WD	H_d	CD_R	MCD_R	AP_R	CVR_R	$DFFITS_R$	WD_R	H_{dR}
0.75	0.04	89.6	89.6	97.5	98	89.6	95.4	20.4	82.5	82.5	93.7	94.7	82.5	90	22.5
	0.11	98.8	98.9	99.9	100	98.9	99.9	10.1	96.2	96.2	99.8	99.8	96.2	98.9	17.9
	0.33	100	100	100	100	100	100	1.2	99.3	99.3	100	100	99.3	100	9.4
	0.67	100	100	100	100	100	100	0.3	99.9	99.9	100	100	99.9	100	5.9
	2.00	100	100	100	100	100	100	0	99.8	99.8	100	100	99.8	100	3.9
	4.00	100	100	100	100	100	100	0.1	99.8	99.8	100	100	99.8	100	2.3
	10.0	100	100	100	100	100	100	0.3	99.9	99.9	100	100	99.9	100	1.5
0.85	0.04	90.1	90.1	97.6	98	90.1	95.2	18.5	83.9	84	94.1	95.2	84	91.8	24.8
	0.11	99.5	99.5	100	100	99.5	99.9	9.3	97.2	97.2	99.8	99.8	97.2	99.3	17.6
	0.33	100	100	100	100	100	100	1.1	99.4	99.4	100	100	99.4	99.9	7
	0.67	100	100	100	100	100	100	0.2	99.6	99.6	100	100	99.6	100	5.7
	2.00	100	100	100	100	100	100	0	99.7	99.7	100	100	99.7	100	3
	4.00	100	100	100	100	100	100	0	99.6	99.6	100	100	99.6	100	2.1
	10.0	100	100	100	100	100	100	1.4	100	100	100	100	100	100	2.4
0.95	0.04	93.1	93.2	99	99.1	93.2	96.9	18.5	88.8	88.9	96.8	97.4	88.9	93.9	23.2
	0.11	99.3	99.3	100	100	99.3	100	5.4	98.7	98.7	100	100	98.7	99.7	13.5
	0.33	100	100	100	100	100	100	1	99.4	99.4	100	100	99.4	100	6
	0.67	100	100	100	100	100	100	0	99.8	99.8	100	100	99.8	100	3.4
	2.00	100	100	100	100	100	100	0.1	99.9	99.9	100	100	99.9	100	3
	4.00	100	100	100	100	100	100	0.7	99.7	99.7	100	100	99.7	100	2.2
	10.0	100	100	100	100	100	100	21	99.8	99.8	100	100	99.8	100	1.6
0.99	0.04	95.8	95.8	99.7	99.8	95.8	98.1	16.3	93.7	93.7	99.5	99.6	93.7	98.4	18.2
	0.11	99.7	99.7	100	100	99.7	99.9	5.2	99	99	100	100	99	99.9	8.7
	0.33	100	100	100	100	100	100	1.3	99.8	99.8	100	100	99.8	100	3.7
	0.67	100	100	100	100	100	100	0.7	99.7	99.7	100	100	99.7	100	4.5
	2.00	100	100	100	100	100	100	2.2	99.7	99.7	100	100	99.7	100	3.6
	4.00	100	100	100	100	100	100	25	99.7	99.7	100	100	99.7	100	2.5
	10.0	100	100	100	100	100	100	94.3	99.8	99.8	100	100	99.8	100	2.2

Table 3.18: Performance of the GRM and the GRRM Influence Measures with Adjusted Anscombe Residuals

ρ	ϕ	CD	MCD	AP	CVR	DFFITS	WD	H_d	CD_R	MCD_R	AP_R	CVR_R	$DFFITS_R$	WD_R	H_{dR}
	0.04	89	89	97	96.7	89	94	26	83	83.3	92	93	83.3	90.5	27
	0.11	98.5	98.5	100	100	98.5	99	11	97	97.2	100	99.7	97.2	98.6	16
	0.33	100	100	100	100	100	100	1.4	100	100	100	100	100	100	3.6
0.75	0.67	100	100	100	100	100	100	0.3	100	100	100	100	100	100	0.8
	2.00	100	100	100	100	100	100	0	100	100	100	100	100	100	0
	4.00	100	100	100	100	100	100	0	100	100	100	100	100	100	0
	10.0	100	100	100	100	100	100	0.1	100	100	100	100	100	100	0
	0.04	91.6	91.6	98	98.1	91.6	96	21	87	86.7	94	95	86.7	92.1	24
	0.11	99.1	99.1	100	99.9	99.1	100	9	98	98.4	100	99.9	98.4	99.4	14
	0.33	100	100	100	100	100	100	1.3	100	100	100	100	100	100	3.2
0.85	0.67	100	100	100	100	100	100	0.3	100	100	100	100	100	100	0.6
	2.00	100	100	100	100	100	100	0	100	100	100	100	100	100	0
	4.00	100	100	100	100	100	100	0	100	100	100	100	100	100	0
	10.0	100	100	100	100	100	100	1.5	100	100	100	100	100	100	0
	0.04	93.4	93.5	99	99.5	93.5	97	18	92	91.6	98	98.1	91.6	96	20
	0.11	99.3	99.3	100	100	99.3	100	5.7	99	99	100	100	99	99.8	9
	0.33	100	100	100	100	100	100	1	100	100	100	100	100	100	1.5
0.95	0.67	100	100	100	100	100	100	0.4	100	100	100	100	100	100	0.5
	2.00	100	100	100	100	100	100	0.2	100	100	100	100	100	100	0.1
	4.00	100	100	100	100	100	100	0.2	100	100	100	100	100	100	0
	10.0	100	100	100	100	100	100	22	100	100	100	100	100	100	2.5
	0.04	94.4	94.4	100	99.9	94.4	98	18	94	93.9	99	99.5	93.9	97.7	21
	0.11	99.8	99.8	100	100	99.8	100	7.2	100	99.8	100	100	99.8	100	8
	0.33	100	100	100	100	100	100	1.3	100	100	100	100	100	100	1.3
0.99	0.67	100	100	100	100	100	100	0.6	100	100	100	100	100	100	0.4
	2.00	100	100	100	100	100	100	4.1	100	100	100	100	100	100	1.3
	4.00	100	100	100	100	100	100	27	100	100	100	100	100	100	4.1
	10.0	100	100	100	100	100	100	96.0	100	100	100	100	100	100	62.0

Table 3.19: Performance of the GRM and the GRRM Influence Measures with Adjusted Working Residuals

ρ	ϕ	CD	MCD	AP	CVR	DFFITS	WD	H_d	CD_R	MCD_R	AP_R	CVR_R	$DFFITS_R$	WD_R	H_{dR}
0.75	0.04	89	89.2	98.1	98.4	89.2	94.8	23.2	97.9	100	100	100	100	100	29.8
	0.11	98.5	98.5	100	100	98.5	99.7	8.8	97.7	100	100	100	100	100	41.3
	0.33	100	100	100	100	100	100	1.4	98.9	98.9	100	100	98.9	100	49.9
	0.67	100	100	100	100	100	100	0.2	99.4	99.4	100	100	99.4	100	51.3
	2.00	100	100	100	100	100	100	0.1	99.9	99.9	100	100	99.9	100	71.5
	4.00	100	100	100	100	100	100	0	100	100	100	100	100	100	92.4
	10.0	100	100	100	100	100	100	0.4	99.9	99.9	100	100	99.9	100	99.9
0.85	0.04	88.9	89	97.2	97.5	89	95	21	98.3	100	100	100	100	100	32.8
	0.11	99.0	99.0	100	99.9	99.0	99.7	7.5	98.3	100	100	100	100	100	41
	0.33	100	100	100	100	100	100	0.5	98.7	98.7	100	100	98.7	100	47.5
	0.67	100	100	100	100	100	100	0.3	100	100	100	100	100	100	53.8
	2.00	100	100	100	100	100	100	0	100	100	100	100	100	100	75.1
	4.00	100	100	100	100	100	100	0.1	99.9	99.9	100	100	99.9	100	96.9
	10.0	100	100	100	100	100	100	1.6	99.9	99.9	100	100	99.9	100	99.9
0.95	0.04	93.1	93.2	99.1	99	93.2	97	16.6	97	100	100	100	100	100	31
	0.11	99.6	99.6	100	100	99.6	100	5.9	97.6	100	100	100	100	100	48.4
	0.33	100	100	100	100	100	100	0.7	99.3	99.3	100	100	99.3	100	57.1
	0.67	100	100	100	100	100	100	0.2	99.8	99.8	100	100	99.8	100	66.4
	2.00	100	100	100	100	100	100	0	100	100	100	100	100	100	93.9
	4.00	100	100	100	100	100	100	0.5	99.9	99.9	100	100	99.9	100	99.8
	10.0	100	100	100	100	100	100	21.4	99.6	99.6	100	100	99.6	100	99.5
0.99	0.04	96.3	96.3	99.5	99.5	96.3	97.6	17.5	97.6	100	100	100	100	100	38.4
	0.11	99.9	99.9	100	100	99.9	99.9	6.2	97.0	97.0	100	100	97.0	100	53.3
	0.33	100	100	100	100	100	100	1.7	98.9	98.9	100	100	98.9	100	73.9
	0.67	100	100	100	100	100	100	1.1	99.3	99.3	100	100	99.3	99.9	89.6
	2.00	100	100	100	100	100	100	2.6	99.5	99.5	100	100	99.5	100	99.5
	4.00	100	100	100	100	100	100	24.5	98.8	98.8	100	100	98.8	100	98.6
	10.0	100	100	100	100	100	100	96.4	94.5	94.5	100	100	94.5	100	94.3

Chapter 4

Influence Diagnostics in the Inverse Gaussian Ridge Regression Model

4.1 Introduction

The inverse Gaussian regression model (IGRM) is suitable to explore the relationship between independent and dependent variables, under the assumption that the dependent variable is positively skewed and well fitted to the inverse Gaussian (IG) distribution. In this Chapter, we present overview and estimation method of the IGRM. Like the GRM, the IGRM is affected by one or more influential observations. So the influential observation detection methods are needed to explore in the IGRM. Some of the IGRM influence diagnostics based on several types of IGRM residuals are given. Amin *et al.* (2015) propose and compare the performance of the IGRM standardized and adjusted residuals with uncorrelated independent variables.

The estimation of the IGRM parameters may inestimable due to multicollinearity among the independent variables. This problem of estimation can be resolved with ridge estimation method. Here again there is a possibility that the multicollinearity and influential observation may occurs simultaneously. Also this chapter covers the formulation of the IGRM residuals with ridge estimator and proposed influence diagnostics with these residuals. Same to the GRM, the performance of IGRM residuals and diagnostic methods are compared to test whether the results are similar as in the GRM. The comparison of some influence diagnostics with and without collinear independent variables of the IGRM under different residual structures using simulation and a real data set are presented.

4.2 The Inverse Gaussian Regression Model

When the response variable is positively skewed but larger skewed than the gamma distribution, then the IGRM is applied. The IG distribution introduced by famous physicists, Schrodinger (1915), who applied this distribution in one dimensional Brownian motion. The particular name (inverse Gaussian distribution) was given by Tweedie (1945) and further statistical properties of this distribution were given in the pioneer papers of Tweedie (1957a,b). Chhikara and Folks (1989) recommended that the IG distribution is used for positively skewed data. The IG distribution is also called Wald distribution (Heinzl and Mittlbock, 2002). The IG is looking to be similar as the gamma except larger skewness and high-pitched as gamma (de Jong and Heller, 2008). The IG model has remarkable applications in the field of engineering, medical and health sciences, chemical, and social sciences (Ferrai *et al.*, 2002; Balka *et al.*, 2011; Wu, 2012; Hanagal and Debade, 2013). The IG distribution has remarkable application in regression analysis. In the literature, various researchers have studied the IGRM with different dimensions. Whitmore (1986) originated an inverse Gaussian ratio estimation model for simple linear models. Iwase (1989) about constant coefficient of variation observed a simple linear regression model. Woldie and Folks (1995) studied about the inverse Gaussian regression model as a method of calibration. Upadhyay *et al.* (1996) studies a Bayesian analysis for the inverse Gaussian non-linear regression model. Woldie *et al.* (2001) presented some IGRM power functions to test the hypothesis of slope parameter under different situations. They find that in certain conditions when the assumption of normal distribution is not fulfilled then by using the IGRM a better estimation is achieved. The IGRM have particularly importance when variance of the explanatory variables is a function of response variable. Ducharme (2001) proposed some goodness-of-fit tests for the IGRM. Hanagal and Dabade (2013) had given frailty model for bivariate survival data about the IGRM. Stogiannis and Caroni (2013) discussed some issues for fitting of the IG first hitting time regression model for lifetime data. Lastly Meshkani *et al.* (2014) discussed the problem of analysis of covariance (ANCOVA) under the IG distribution for the dependent variable. It is given that for real life situations symmetry and ANCOVA violating certain assumptions such as independent mean and variance of normal response variable, homoscedasticity, which are not applicable. In this condition, the IG distribution leading to closed form reliable maximum likelihood estimation of parameters and giving accuracy of results for data analysis.

4.3 Derivation of the IGRM

Let y_i be the dependent variable, which comes from the IG distribution i.e. $IG(\mu, \phi)$ and the probability density function of the IG distribution is given as

$$f(y_i/\mu_i, \phi) = \frac{1}{\sqrt{2\pi\phi y_i^3}} \exp\left[-\frac{1}{2\mu_i^2\phi y_i}(y_i - \mu_i)^2\right] \;;\; y_i, \mu_i, \phi > 0. \tag{4.1}$$

The IG distribution has different shapes with diverse parameters. Also $\phi = \sigma^2 = \frac{E(Y)^3}{V(Y)}$ is assumed to be fixed shape parameter and unknown and so estimated for the IGRM using sample observations. There are different form of IG distribution but most common one is the Eq. (4.1) which is the reparameterization of Johnson *et al.* (1994) page 261 Eq. (15.4a) is obtained by setting the dispersion parameter $\phi = \frac{1}{\lambda}$. If $\mu = 1$, then the IG distribution is referred as the standardized IG or Wald distribution. The dispersion parameter determines the shape of the given density. For moderate value of ϕ, the density is highly positively skewed while increasing ϕ, the density approaches normal distribution. These shapes of the IG density are shown in Fig. 4.1. In exponential family of distribution, Eq. (4.1) can be written as

$$f(y; \mu, \phi) = \exp\left[-\frac{y}{2\mu^2\phi} + \frac{1}{\mu\phi} - \frac{1}{2y\phi} - \frac{1}{2}\ln\left(2\pi y^3\right) - \ln\phi\right]. \tag{4.2}$$

So the mean and variance for the IG distribution are given as $E(y) = b'(\theta) = \frac{\partial b}{\partial\mu}\frac{\partial\mu}{\partial\theta} = \left(\frac{-1}{\mu^2}\right)(-\mu^3) = \mu = \frac{1}{\sqrt{\eta}}$, and $V(y) = \phi b''(\theta) = \phi V(\mu) = \phi\mu^3$. So the mean and variance of the observed random vector y are $E(y) = \mu$ and $V(y) = \mu^3\phi$, respectively.

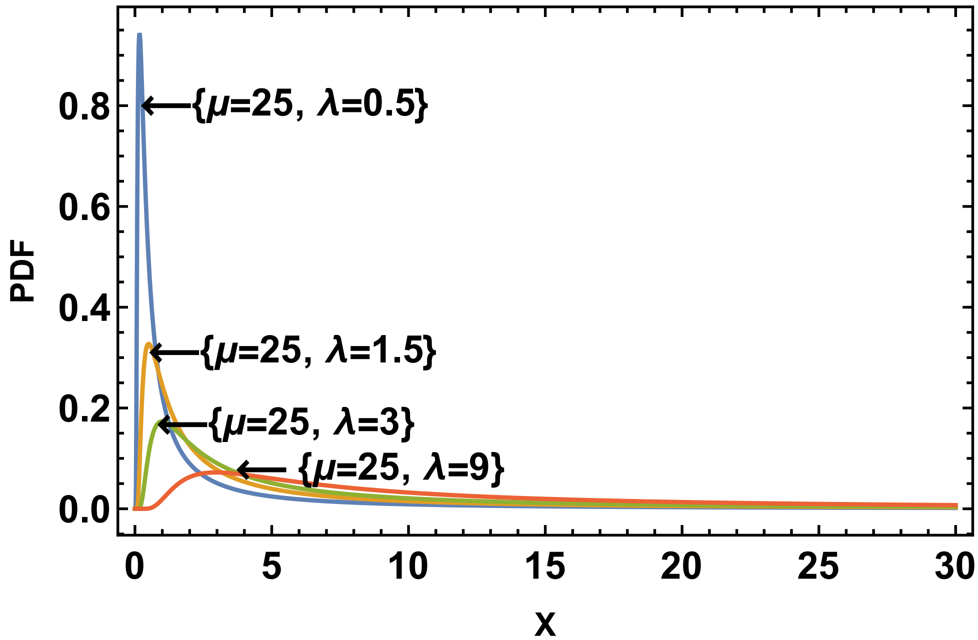


Figure 4.1: The Density Curves of the IG Distribution

This indicated that the variance of the IG random variable is proportional to cube power of its expectation (Tsou, 2011).

4.4 Link Function for the IGRM

Generally, for the IGRM the most suitable link function is the quadratic reciprocal. Wu (2012) has shown with the help of simulation study that log-link is more suitable than quadratic reciprocal link function. While here we consider quadratic reciprocal link function because mostly authors (McCullagh and Nelder, 1989; Lindsey, 2007; Hardin and Hilbe, 2012) recommend this link function for fitting the IGRM.

4.5 Estimation of the IGRM

Let $y = (y_1, ..., y_n)^T$ be the response vector of the dependent variable which is assumed to follow IG distribution. Let $X = (x_{i1}, x_{i2}, ..., x_{ip})^T$ is the $n \times p$ data matrix of center and standardized independent variables with full rank such that $n > p$. Let $Z = (1, X)$ be the design matrix and $E(y_i) = \mu_i$ be the mean function of the response variable, $g(\mu_i) = \frac{1}{\mu_i^2} = \eta_i$ be the link function, where $\eta_i = z_i^T \beta^*$ be the linear predictor and β^* are the $n \times p'$ vector of unknown parameters which includes intercept. The log likelihood of Eq.(4.2) is given as

$$l(\mu_i, \phi) = \sum_{i=1}^{n} \left\{ \frac{\left[\frac{(y_i)}{2\mu_i^2} - \frac{1}{\mu_i}\right]}{\phi} + \frac{1}{(-2y_i\phi)} - \frac{1}{2} \ln\left(2\pi y_i^3 \phi\right) \right\}. \tag{4.3}$$

As $\mu = \frac{1}{\sqrt{\eta}} = \frac{1}{\sqrt{Z\beta^*}}$, so Eq. (4.3) can also be expressed as

$$l(\beta^*, \phi) = \sum_{i=1}^{n} \left\{ \frac{\left[\frac{y z_i^T \beta^*}{2} - \sqrt{z_i^T \beta^*}\right]}{\phi} + \frac{1}{(-2y_i\phi)} - \frac{1}{2} \ln\left(2\pi y_i^3 \phi\right) \right\}. \tag{4.4}$$

For the estimation β^* by MLE method has the solution of the system of equations by setting first derivative of Eq. (4.4) equals to zero, we have

$$U\left(\beta_j^*\right) = \frac{\partial l_i}{\partial \beta_j^*} = \frac{1}{2\phi}\left(y - \frac{1}{\sqrt{z_i^T \beta^*}}\right) z_i^T = 0. \tag{4.5}$$

Since solution of the system of equations in Eq.(4.5) is non-linear, so the Newton-Raphson iterative procedure is used to estimate the unknown parameter. For iterative procedure of the IGRM, initial values and full algorithm for the estimation of unknown parameter can be found in Hardin and Hilbe (2012). Let $\beta^{*(m)}$ be the approximated ML value of β^* at the mth iteration with convergence, the iterative method (Green, 1984) gives the relation as

$$\beta^{*(m+1)} = \beta^{*(m)} + \left\{ I\left(\beta^{*(m)}\right) \right\}^{-1} U\left(\beta^{*(m)}\right), \tag{4.6}$$

where $I\left(\beta^{*(m)}\right)$ is the information matrix and $U\left(\beta^{*(m)}\right)$ is the score vector with dimension $p' \times 1$ and both information and score vectors are evaluated at $\beta^{*(m)}$. At convergence, the unknown parameter can be estimated as

$$\hat{\beta}^* = \left(Z^T \hat{W} Z\right)^{-1} Z^T \hat{W} y^*, \tag{4.7}$$

where $y_i^* = \hat{\eta}_i + \frac{(y_i - \hat{\mu}_i)}{\hat{\mu}_i^3}$ be the adjusted response variable and $\hat{W} = diag\left(\hat{\mu}_1^3, \hat{\mu}_2^3, ..., \hat{\mu}_n^3\right)$. Here $\hat{\mu}_i = \frac{1}{\sqrt{z_i^T \hat{\beta}^*}}$, $i = 1, 2, ..., n$. and $\hat{\eta}_i = z_i^T \hat{\beta}^*$. Both y^* and $\hat{W}$ are found by the iterative methods and for the detail derivations and procedure, readers are referred to Hardin and Hilbe (2012). At the final iteration, the estimators maximize the likelihood of Eq. (4.4). The weights are used due to the presence of heteroscedasticity in the GLMs as the variances of the response variable are not constant.

Now and onward we use the estimators which are obtained at the final iteration. The estimated value of the response variable at final iteration of the IGRM is

$$\hat{\mu}_i = \frac{1}{\sqrt{z_i^T \hat{\beta}^*}}. \tag{4.8}$$

and $\hat{\phi}$ is the estimated dispersion parameter at the final iteration and is computed as

$$\hat{\phi} = \frac{\sum_{i=1}^{n} (y_i - \hat{\mu}_i)^2}{n - p'} = \frac{\sum_{i=1}^{n} \left(y_i - \frac{1}{\sqrt{z_i^T \beta^*}} \right)^2}{n - p'}. \tag{4.9}$$

This dispersion parameter later on use for the computation of the IGRM residuals and influence diagnostics.

4.6 Hat matrix and Leverages in the IGRM

Like the GRM, it is possible to define the hat matrix for IGRM as

$$H = \hat{W}^{\frac{1}{2}} Z \left(Z^T \hat{W} Z \right)^{-1} Z^T \hat{W}^{\frac{1}{2}}, \tag{4.10}$$

where $\hat{W} = diag\left(\hat{\mu}_1^3, ..., \hat{\mu}_n^3\right)$ is the weight matrix of the IGRM. The leverages are the i-th diagonal elements of the hat matrix H, given in Eq. (4.10), and defined as $h_{ii} = diag\left(H\right)$. These leverages are latterly used for the computation of the IGRM residuals and their influence diagnostics.

4.7 The IGRM Residuals

In this section, we will give the functional forms of the IGRM residuals like Pearson, deviance, likelihood etc. Also in this section, we will give the standardized and adjusted forms of the IGRM residuals with uncorrelated independent variables.

4.7.1 The IGRM Pearson Residuals

The Pearson residuals in the IGRM are defined as

$$r_{pi} = \frac{y_i - \hat{\mu}_i}{\sqrt{Var(\hat{\mu}_i)}} = \frac{y_i - \hat{\mu}_i}{\sqrt{\hat{\mu}_i^3}}. \tag{4.11}$$

The standardized and adjusted Pearson residuals of the IGRM are similar as we define in Chapter 3 and in Table 4.1.

4.7.2 The IGRM Deviance Residuals

The deviance residuals for the IGRM is defined as

$$r_{di} = sign(y_i - \hat{\mu}_i)\sqrt{|d_i|}, \tag{4.12}$$

where $d_i = \frac{(y_i - \hat{\mu}_i)^2}{\hat{\mu}_i^2 y_i}$ is the deviance function for the IGRM. This deviance function can also be used in estimating the dispersion parameter. The standardized and adjusted deviance residuals of the IGRM are similar as define in Chapter 3 and in Table 4.1.

4.7.3 The IGRM Likelihood Residuals

The likelihood residuals (Fox, 2002) are the weighted residuals of the standardized deviance and standardized Pearson residuals and are defined as

$$r_{li} = sign(y_i - \hat{\mu}_i)\sqrt{h_{ii}(sr_{pi})^2 + (1 - h_{ii})(sr_{di})^2}, \tag{4.13}$$

where sr_{pi} and sr_{di} are the standardized Pearson and deviance residuals which are given in Table 4.1.

4.7.4 The IGRM Anscombe Residuals

From Eq. (2.35), the Anscombe residuals for the IGRM are defined by

$$r_{ai} = \frac{\ln(y_i) - \ln(\hat{\mu}_i)}{\sqrt{(\hat{\mu}_i)}}. \tag{4.14}$$

The standardized and adjusted forms of the Anscombe residuals for the IGRM are given in Table 4.1.

4.7.5 The IGRM Working Residuals

This type of the GLM residuals by Hardin and Hilbe (2012) are defined as

$$r_{Wi} = (y_i - \hat{\mu}_i)\left(\frac{\partial \eta}{\partial \mu}\right), \tag{4.15}$$

where for the IGRM, $\frac{\partial \eta}{\partial \mu} = -\frac{2}{\mu^3}$ So the estimated working residuals for the IGRM, Eq. (4.15) becomes as

$$r_{Wi} = (y_i - \hat{\mu}_i)\left(-\frac{2}{\hat{\mu}_i^3}\right). \tag{4.16}$$

The standardized and adjusted working residuals of the IGRM are similar as define in Chapter 3 and in Table 4.1.

In Table 4.1 for the adjusted forms of the IGRM residuals, $\hat{E}(\cdot)$ and $\hat{V}(\cdot)$ are obtained by replacing μ_i with estimated value $\hat{\mu}_i$.

Table 4.1: Summary of the IGRM Residuals

Residuals	Formula	Standardized	Adjusted
Pearson	$r_{pi} = \frac{y_i - \hat{\mu}_i}{\sqrt{\hat{\mu}_i^3}}$	$sr_{P_i} = \frac{r_{Pi}}{\sqrt{\hat{\phi} m_{ii}}}$	$Ar_{Pi} = \frac{r_{Pi} - \hat{E}(r_{Pi})}{\sqrt{\hat{V}(r_{Pi})}}$
Deviance	$r_{di} = sign\left(y_i - \hat{\mu}_i\right)\sqrt{\left\lvert\frac{(y_i - \hat{\mu}_i)^2}{\hat{\mu}_i^2 y_i}\right\rvert}$	$sr_{di} = \frac{r_{di}}{\sqrt{\hat{\phi} m_{ii}}}$	$Ar_{di} = \frac{r_{di} - \hat{E}(r_{di})}{\sqrt{\hat{V}(r_{di})}}$
Likelihood	$r_{li} = sign\left(y_i - \hat{\mu}_i\right)\sqrt{h_{ii}(sr_{pi})^2 + (m_{ii})\left(sr_{di}\right)^2}$		$Ar_{li} = \frac{r_{li} - \hat{E}(r_{li})}{\sqrt{\hat{V}(r_{li})}}$
Anscombe	$r_{ai} = \frac{\ln(y_i) - \ln(\hat{\mu}_i)}{\sqrt{(\hat{\mu}_i)}}$	$sr_{ai} = \frac{r_{ai}}{\sqrt{\hat{\phi} m_{ii}}}$	$Ar_{ai} = \frac{r_{ai} - \hat{E}(r_{ai})}{\sqrt{\hat{V}(r_{ai})}}$
Working	$r_{Wi} = (y_i - \hat{\mu}_i)\left(-\frac{2}{\hat{\mu}_i^3}\right)$	$sr_{Wi} = \frac{r_{Wi}}{\sqrt{\hat{\phi}(m_{ii})}}$	$Ar_{Wi} = \frac{r_{Wi} - \hat{E}(r_{Wi})}{\sqrt{\hat{V}(r_{Wi})}}$

4.8 Influence Diagnostics in the IGRM

A little attention had given for influence diagnostics in the IGRM. For the IGRM, influence diagnostics is also important as the other GLMs like Poisson regression Beta regression etc. The only study given by Xie and Wei (2008), is available for the influence analysis of the Poisson IG mixture modeling and no research has been conducted in the literature concerning influence diagnostics in the IGRM. Lin *et al.* (2004) proposed diagnostics using score test for different dispersion parameters in the IGRM. Also, the literature indicated that no study has been conducted for the influence assessment based on the different IGRM residuals. The comparison to determine which the IGRM residual is more appropriated or they all are equally important for the influence diagnostics yet to be explored.

4.9 The Inverse Gaussian Ridge Regression Model

There are the situations, where dependent variable assumes to follow IG distribution and independent variables are multicollinear or non-orhtogonal. So when IRLS estimation method is applied to non-orthogonal data, then the IGRM estimates are doubtful for the model inferences. To remedy this problem, we use the ridge estimation by adopting Schaefer *et al.* (1984) method. This estimation with ridge estimator of the IGRM, we call the IGRRM. The IGRRM estimates are biased but smaller MSE than the MSE of unbiased estimates of the IGRM. The details estimation with ridge estimator of the IGRM is similar to the GRM as we discuss in Chapter 3 but different functional form of mean, variance and weights. So the IGRM estimates with ridge estimator by Eq. (3.48) are given by

$$\hat{\beta}_R = \left(Z^T \hat{W} Z + kI\right)^{-1} Z^T \hat{W} Z \hat{\beta}^*, \tag{4.17}$$

where k is the ridge parameter which is already defined and detail is given in Chapter 3. The estimated mean function with ridge estimator of the IGRRM is given by

$$\hat{\mu}_{Ri} = \frac{1}{\sqrt{z_i^T \hat{\beta}_R}}. \tag{4.18}$$

Now this estimated mean function with ridge estimate is used for the computations of the IGRRM residuals and influence diagnostics.

4.10 The IGRRM Residuals

In this section, we will give the functional forms of the IGRRM residuals like Pearson, deviance, likelihood etc. Also in this section, we will give the standardized and adjusted IGRM residuals with correlated independent variables. These forms of the

IGRRM residuals we study for the influence diagnostics to determine their importance under the assumption that independent variables are multicollinear.

4.10.1 The IGRRM Pearson Residuals

The Pearson residuals in the IGRM with ridge estimator are defined as

$$r_{PRi} = \frac{y_i - \hat{\mu}_{Ri}}{\sqrt{Var\left(\hat{\mu}_{Ri}\right)}} = \frac{y_i - \hat{\mu}_{Ri}}{\sqrt{\hat{\mu}_{Ri}^3}}. \tag{4.19}$$

The standardized and adjusted Pearson residuals of the IGRRM are given in Table 4.2.

4.10.2 The IGRRM Deviance Residuals

The deviance residuals of the IGRM with ridge estimator are defined as

$$r_{dRi} = sign\left(y_i - \hat{\mu}_{Ri}\right)\sqrt{|d_{Ri}|}, \tag{4.20}$$

where $d_{Ri} = \frac{(y_i - \hat{\mu}_{Ri})^2}{\hat{\mu}_{Ri}^2 y_i}$ is the deviance function of the IGRM with ridge estimator. This deviance function can also be used in estimating the dispersion parameter with ridge estimator. The standardized and adjusted deviance residuals of the IGRRM are given in Table 4.2.

4.10.3 The IGRRM Likelihood Residuals

The IGRRM likelihood residuals are the weighted residuals of the standardized deviance and standardized Pearson residuals with ridge estimator and are defined as

$$r_{lRi} = sign\left(y_i - \hat{\mu}_{Ri}\right)\sqrt{h_{Rii}\left(sr_{PRi}\right)^2 + \left(1 - h_{Rii}\right)\left(sr_{dRi}\right)^2}, \tag{4.21}$$

where sr_{PRi} and sr_{dRi} are the standardized Pearson and standardized deviance residuals with ridge estimator which are given in table 4.2. The adjusted likelihood residuals with the ridge estimator are also given in Table 4.2.

4.10.4 The IGRRM Anscombe Residuals

From Eq. (4.14), the Anscombe residuals of the IGRM with ridge estimator are defined by

$$r_{aRi} = \frac{\ln\left(y_i\right) - \ln\left(\hat{\mu}_{Ri}\right)}{\sqrt{\left(\hat{\mu}_{Ri}\right)}}. \tag{4.22}$$

The standardized and adjusted Anscombe residuals of the IGRM with ridge estimates are given in Table 4.2.

4.10.5 The IGRRM Working Residuals

By following Eq. (4.16), the estimated working residuals for the IGRM with ridge estimator are defined by

$$r_{WRi} = \left(y_i - \hat{\mu}_{Ri}\right)\left(-\frac{2}{\hat{\mu}_{Ri}^3}\right) \tag{4.23}$$

Table 4.2: Summary of the IGRRM Residuals

Residuals	Formula	Standardized	Adjusted
Pearson	$r_{PRi} = \frac{y_i - \hat{\mu}_{Ri}}{\sqrt{\hat{\mu}_{Ri}^3}}$	$sr_{PRi} = \frac{r_{PRi}}{\sqrt{\hat{\phi} m_{Rii}}}$	$Ar_{PRi} = \frac{r_{PRi} - \hat{E}(r_{PRi})}{\sqrt{\hat{V}(r_{PRi})}}$
Deviance	$r_{dRi} = sign\left(y_i - \hat{\mu}_{Ri}\right) \sqrt{\left\lvert \frac{(y_i - \hat{\mu}_{Ri})^2}{\hat{\mu}_{Ri}^2 y_i} \right\rvert}$	$sr_{dRi} = \frac{r_{dRi}}{\sqrt{\hat{\phi} m_{Rii}}}$	$Ar_{dRi} = \frac{r_{dRi} - \hat{E}(r_{dRi})}{\sqrt{\hat{V}(r_{dRi})}}$
Likelihood	$r_{lRi} = sign\left(y_i - \hat{\mu}_{Ri}\right) \sqrt{h_{Rii}(sr_{PRi})^2 + (m_{Rii})\left(sr_{dRi}\right)^2}$		$Ar_{lRi} = \frac{r_{lRi} - \hat{E}(r_{lRi})}{\sqrt{\hat{V}(r_{lRi})}}$
Anscombe	$r_{aRi} = \frac{\ln(y_i) - \ln(\hat{\mu}_{Ri})}{\sqrt{(\hat{\mu}_{Ri})}}$	$sr_{aRi} = \frac{r_{aRi}}{\sqrt{\hat{\phi} m_{Rii}}}$	$Ar_{aRi} = \frac{r_{aRi} - \hat{E}(r_{aRi})}{\sqrt{\hat{V}(r_{aRi})}}$
Working	$r_{WRi} = (y_i - \hat{\mu}_{Ri})\left(-\frac{2}{\hat{\mu}_{Ri}^3}\right)$	$sr_{WRi} = \frac{r_{WRi}}{\sqrt{\hat{\phi} m_{Rii}}}$	$Ar_{WRi} = \frac{r_{WRi} - \hat{E}(r_{WRi})}{\sqrt{\hat{V}(r_{WRi})}}$

The standardized and adjusted working residuals of the IGRM with ridge estimator are given in Table 4.2. In Table 4.2, for the adjusted forms of residuals with ridge estimate, $\hat{E}\,(.)$ and $\hat{V}\,(.)$ are obtained by replacing μ_{Ri} with estimated value $\hat{\mu}_{Ri}$.

4.11 Influence Diagnostics in the IGRRM

When there are more than one independent variables in the IGRM, then the problem of multicollinearity may arise in the independent variables. The IGRM estimates may be affected by two problems i.e. multicollinearity and influential observations. Under multicollinearity, the variances of the IGRM estimates are so large and results are misleading for the influence assessment and inferences on the IGRM fitting. To overcome this problem, the IGRRM is used for the estimation and influence assessment. In this section, we modify and evaluate some influence diagnostics for the IGRRM by following the notations as we discuss in Chapter 3. These formulations of influence measures are same but different residual functional form.

The functional form of the IGRM and IGRRM is well defined in Section 4.5 and Section 4.9 respectively. In the upcoming section, we are giving the simulation layout for studying the performance of the IGRM and IGRRM influence diagnostics with different residuals.

4.12 Numerical Results

In this section, we evaluate and test our proposed influence diagnostics in the IGRM and IGRRM with the help of simulations and a real data set.

4.12.1 Simulation Study

In this section, we present simulation results to test the performance of influence diagnostics defined in previous sections for the IGRM and IGRRM with various conditions. These conditions include with different dispersion parametric assumed values, with different sample sizes, with different multicollinear levels and with different the IGRM and the IGRRM residuals.

We consider the following Monte Carlo scheme to compare the performance of IGRRM influence diagnostic methods with all the IGRM residuals. The data generation process for the IGRRM is generated for $p = 4$ explanatory variables as follows

$$y \sim IG\left(\mu_i, \phi\right),$$

where $\mu_i = E(y_i) = (\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4})^{-1/2}$, where β_j are the slope coefficients for respective jth column and assumes the following arbitrary values for the true parameters as $\beta_0 = 0.05, \beta_1 = 0.0025, \beta_2 = 0.005, \beta_3 = \beta_4 = 0.0001$. The number of observations are set at $n = 50$ and the number of replications are 1000. The dispersion parameter ϕ assumes the following arbitrary values as $\phi = 0.06$, 0.11, 0.25, 0.5, 2, 20 &200, and the design matrix X is generated with no influential observations as

$$X_{ij} = \left(1 - \rho^2\right)^{\frac{1}{2}} Z_{ij} + \rho Z_{i5},$$

where $Z_{ij} \sim IG(3, 0.5)$, $i = 1, 2, ..., n$ and $j = 1, 2, 3, 4$. where ρ is the correlation between two explanatory variables and we take the following values 0.75, 0.85, 0.95 and 0.99. Then we make influential observation in X's that is the 30ˆth observation by following chapter3 method and is replaced in the complete data set. Now the performance of these diagnostics for the identification of generated influential observation with various collinearity levels and with different values of dispersions are performed on the basis of IG generated samples. These simulation studies are performed on R software.

Influence Diagnostics in the IGRM and the IGRRM with Pearson Residuals

In this section, we analyze the influence diagnostics with four forms of the Pearson residuals. The simulation results of the IGRM influence diagnostics with ridge estimates using standardized Pearson residuals are presented in Table 4.11. From Table 4.11, we observe that the performance of all the IGRM and the IGRRM diagnostics are identical and approximately detection is 100% when $\phi \leqslant 2$. For $\phi > 2$, the influential observation detection of these methods reduces. Another result we find is that for this dispersion, the detections of the IGRM diagnostics are more than the IGRRM diagnostics. These smaller detections of the IGRRM diagnostics are due to the leverages with ridge estimates. Like the GRM influence diagnostics, the H_d method in the IGRM and IGRRM performs poorly in detecting the generated influential observation. However, the influence detection performance of CD, MCD and $DFFITS$ is identical. Moreover, the detection performance of the AP, CVR and WD methods are better than the CD, MCD and $DFFITS$ methods. Similar results are also found with the ridge estimates. Now substantial effect of the multicollinearity with different levels on the IGRM and IGRRM diagnostics using standardized Pearson residuals are seen except the H_d method.

On studying the influence diagnostics with adjusted Pearson residuals, we find the similar results as with standardized Pearson residual. Additionally for $\phi > 2$, we find the detections of influential observation with the adjusted Pearson residuals are slightly better than standardized Pearson residuals. These results are given in Table 4.16. From Table 4.16, we also find that a moderate effect of multicollinearity on the influence diagnostic methods (CD, MCD, $DFFITS$ and WD) with and without ridge estimates.

Influence Diagnostics in the IGRM and the IGRRM with Deviance Residuals

Like the Pearson residuals, now we study the influence diagnostics with four forms of the Deviance residuals in IGRM and the IGRRM with the help of simulation. The simulation results of the IGRM influence diagnostics with ridge estimates using standardized Deviance residuals are given in Table 4.12. From Table 4.12, we observe that, the performance of all the IGRM and the IGRRM influence diagnostic methods is approximately identical except the H_d method when $\phi \leqslant 2$. While for $\phi > 2$, the detection of all the IGRM diagnostics decreases and is better than the IGRRM diagnostics. For $\phi > 2$ and $\rho = 0.99$, we also find that the detection rate with deviance residuals are slightly less than the diagnostics with Pearson residuals. When $\phi > 2$ and $\rho = 0.95$ & $\rho = 0.99$, a moderate positive effect of multicollinearity are observed on all the IGRRM influence detection techniques except the H_{dR}. The simulation results also direct that the detection performance of the IGRRM with standardized deviance residuals are better than the standardized Pearson residuals. These results are true only, if $\phi > 2$ and the multicollineaity is lower to severe.

Now considering the influence diagnostics with the adjusted deviance residuals, we observe that the influence detection is identical to that of using standardized Pearson residuals and less good than that of using standardized deviance residuals. These results are shown in Table 4.17. Multicollinearity has a little positive effect on the IGRRM influence diagnostics for $\phi > 2$. So the simulation results of the adjusted deviance residuals with and without ridge estimator are not better than that of IGRRM influence diagnostics with standardized deviance residuals.

Influence Diagnostics in the IGRM and the IGRRM with Likelihood Residuals

The simulation results of the IGRM influence diagnostics with ridge estimates using the likelihood residuals are presented in Table 4.13. From Table 4.13, we find that for $\phi \leqslant 2$, the performance of CD is not better as compared to the CD_R. In other words for $\phi \leqslant 2$,the influence detections of all the IGRRM diagnostics are better than the IGRM diagnostics except the CD, MCD and H_d methods. While with the ridge estimates, the influence detection are decreasing with the increasing in ϕ i.e. $\phi > 2$. There is no effect on the IGRM diagnostics but a little effect on the IGRRM diagnostics is observed, when $\phi > 2$. For the multicollinear data, the AP method detects larger influential observations than that of the AP_R. Increase in multicollinearity has a little positive effect on the IGRRM influence detection techniques while ϕ has a little negative effect on the IGRRM influential observation detection techniques. These results are true only, if $\phi > 2$. Detection of the influential observation by H_d and H_{dR} methods are very poor than the other IGRM and IGRRM diagnostic methods. These results indicates that the H_d and H_{dR} methods with likelihood residuals are not better for the influence diagnostics. While the detection by AP, CVR and WD methods are superior to the other diagnostic measures.

On studying the influence diagnostics in the IGRM and IGRRM with the adjusted likelihood residuals, we find the influential observation detections are slightly less than the likelihood as well as from the standardized Pearson and deviance residuals. From Table 4.18, we also find that WD method is better than the other IGRM influence diagnostics for the detection of influential observation. The multicollinearity does not affect the IGRM diagnostics while mild positive effect on the IGRRM influence diagnostics is observed. All these results with the adjusted likelihood residuals hold only for $\phi > 2$.

Influence Diagnostics in the IGRM and the IGRRM with Anscombe Residuals

The influence diagnostics of the IGRM methods with the standardized Anscombe residuals are similar to the standardized Pearson residuals but not better than the likelihood residuals. While the Anscombe residuals with the ridge estimates, IGRRM diagnostics perform to some extent better than the IGRRM with the standardized Pearson residuals for $\phi > 2$. We also find that the influence diagnostic performance of all the IGRM and the IGRRM diagnostics seem to be identical except the H_d method, when $\phi \leqslant 2$. These results are given in Table 4.14. The IGRRM diagnostics results with the Adjusted Anscombe residuals are better than the results computed with the adjusted likelihood residuals, the adjusted deviance and the adjusted Pearson residuals. These results are indicated in Table 4.19.

Influence Diagnostics in the IGRM and the IGRRM with Working Residuals

The detections of influential observation in the IGRM and IGRRM using standardized working residuals are so interesting. From Table 4.15 for $\phi \leqslant 2$ and all multicollinearity levels, we find that the influence detections by all the IGRM and IGRRM diagnostics are similar to all the types of residuals except the likelihood residuals. While for $\phi > 2$, the influential observation detection by the IGRM and IGRRM reduces except the H_d method. For this range of dispersion, multicollinearity affects positively on the IGRRM influence diagnostics methods. This indicates that the influential observation detection by the IGRRM diagnostics increases with the increase in multicollinearity. On comparison, the diagnostic procedures, we find that AP, CVR and WD methods are better than the CD, MCD and $DFFITS$ methods. Again the H_d method detects the influential observations poorly as compared to other methods.

The influence detections by the IGRM and the IGRRM methods using the adjusted working residuals are given in Table 4.20. From Table 4.20 for $\phi \leqslant 2$, the detection by all the IGRM and IGRRM methods are identical (100%). While for $\phi > 2$,the influence detections by the IGRM diagnostics (CD, MCD, AP, CVR, WD and $DFFITS$) methods are similar to that of using standardized working residuals. For $\phi > 2$, the detection by the IGRRM diagnostics (CD_R, MCD_R, AP_R, CVR_R, WD_R and $DFFITS_R$) methods with the adjusted working residuals are less than that of with standardized working residuals. Another interesting result which we observe that the detection by the H_d and H_{dR} method increases to 50% . We also find that multicollinearity has similar effect on the IGRRM diagnostics with adjusted working residuals to that of with standardized working residuals.

4.12.2 Application: Stack Loss Data

It is interesting to note that on testing the probability distribution of the response variable (stack loss) using Crammer-Von-Mises test, we find that the stack loss is well fitting to the IG distribution and results are shown in Table 4.3. Another interesting result we observe that the CI of the $\left(X'\hat{W}X\right)$ is 5157.01. This indicates the existence of multicollinearity in the independent variables.

Table 4.3: Distribution goodness of fit tests for stack loss data

		Probability Distributions				
Goodness of Fit Test		Normal	Exponential	Gamma	IG	Weibull
Anderson-darling	Statistic	1.4157	2.5777	0.6751	0.4486	0.8721
	p-value	0.001	0.0023	0.0781	0.3934	0.0241
Cramer-Von Mises	Statistic	0.2363	0.4898	0.0992	0.0613	0.1394
	p-value	0.0016	0.0012	0.1151	**0.5147**	0.0295
Pearson	Statistic	14.000	15.333	4.6667	6.000	6.0000
	p-value	0.0073	0.009	0.3232	0.1991	0.1991

Table 4.4: Influential observations of stack loss data set in the Literature

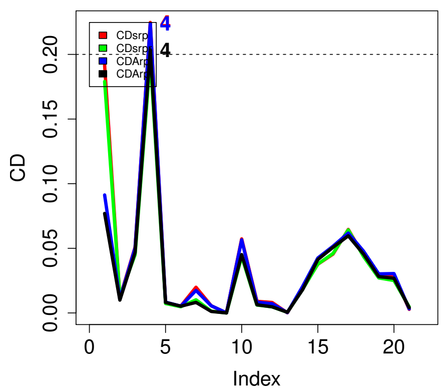
Sr#	References	Influential Observations
1	Balasooriya *et al.* (1987)	17,21
2	Hossain and Naik (1991)	17,21
3	Li *et al.* (2001)	1,2,3,4,21
4	Meloun and Militky (2001)	1,2,3,4,17,21
5	Hoaglin *et al.* (2006)	1,2,3,4,21
6	Nurunnabi *et al.* (2014)	1,2,3,4,17,21

From Table 4.4, we find in the literature related to the LM that the observations 1^{st}, 2^{nd}, 3^{rd}, 4^{th}, 17^{th} and 21^{st} observations are detected as influential observations. Now we study this data with the IGRM and the IGRRM for testing the performance of difference residuals and diagnostics methods for the detection of influential observation.

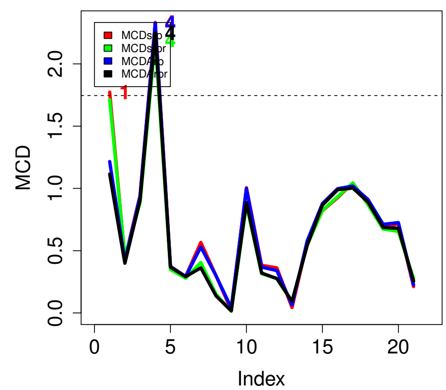
Now we study the influence diagnostics using different forms of the IGRM residuals with ridge estimates. From Tables 4.5, we find that with four forms of Pearson residuals, the observations 1^{st}, 2^{nd}, 4^{th}, 7^{th}, 8^{th}, 9^{th}, 13^{th} and 21^{st} are diagnosed as the influential observations by different IGRM diagnostic methods. With the standardized Pearson residuals, influence diagnostics CD, $DFFITS$ and H_d fail to detect any influential observations. While with this form of the Pearson residuals, the diagnostic methods (MCD, AP, CVR and WD) detect 1^{st}, 2^{nd}, 4^{th}, 7^{th}, 8^{th}, 9^{th}, 13^{th} and 21^{st} observations as the influential observations. The observation 1^{st} commonly detected by MCD, CVR and WD methods but AP fails to detect this observation as influential observation. These detections with the ridge Pearson residuals are similar to that of using standardized Pearson residuals with the addition that 13^{th} observation is not diagnosed as influential observation by any of the given methods. And again with the ridge Pearson residuals, CD and H_d methods fail to detect any influential observation. Additionally with this form of residuals, $DFFITS$ detects the 4^{th} observation as influential observation. On studying the influence diagnostics with the adjusted Pearson residuals, we find better influence diagnostics because all methods detect influential observations. Similar results are found using the ridge adjusted Pearson residuals. The influence detection of AP with standardized Pearson residuals are identical to the influence detection with adjusted Pearson residuals. The performance of CVR and WD methods with four forms of Pearson residuals is identical. The MCD and $DFFITS$ methods diagnosed the influential observations with three form of Pearson (standardized ridge, adjusted and adjusted ridge) residuals is identically better than with standardized Pearson residuals. The CD method detected the influential observation in a better way with two forms of adjusted Pearson residuals. These results are also shown in Fig. 4.2 (a-g) to indicate the detected influential observations.

Table 4.5: Influential Observations with the IGRM and the IGRRM Pearson Residuals

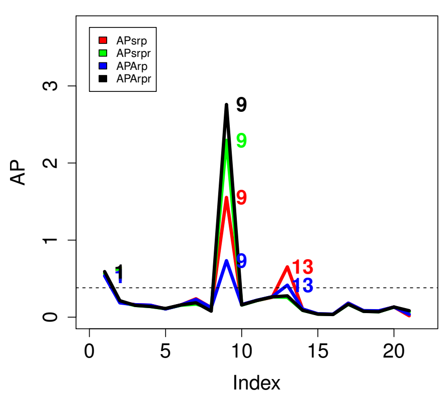
	Standardized Pearson Residuals		Adjusted Pearson Residuals	
Methods	sr_{Pi}	sr_{PRi}	Ar_{Pi}	Ar_{PRi}
CD	Nil	Nil	4	4
MCD	1	4	4	4
AP	9,13	1,9	1,9,13	1,9
CVR	1,2,4,7,8,21	1,2,4,7,8,21	1,2,4,7,8,21	1,2,4,7,8,21
$DFFITS$	Nil	4	4	4
WD	1	1	1	1
H_d	Nil	Nil	1	1



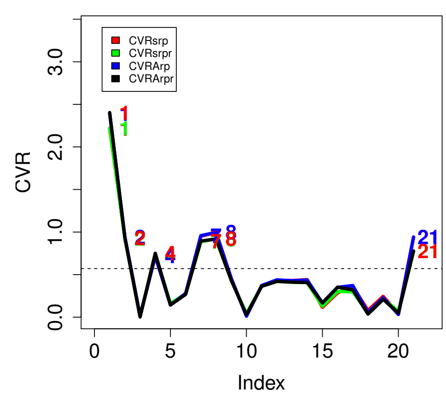
(a) IGRM CD index plot



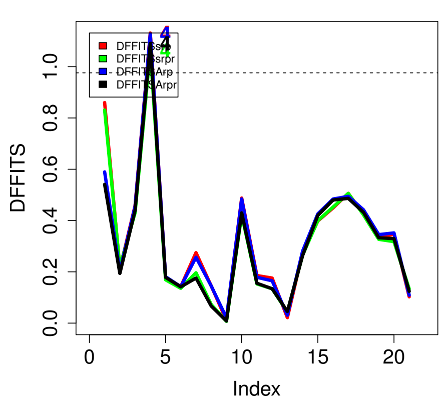
(b) IGRM MCD index plot



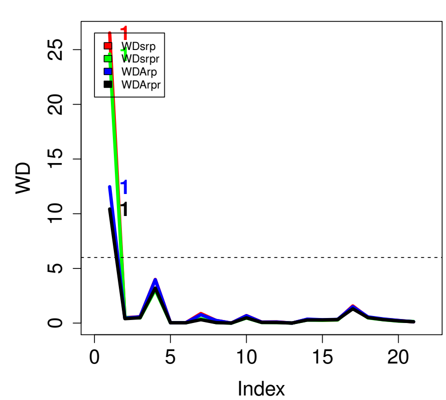
(c) IGRM AP index plot



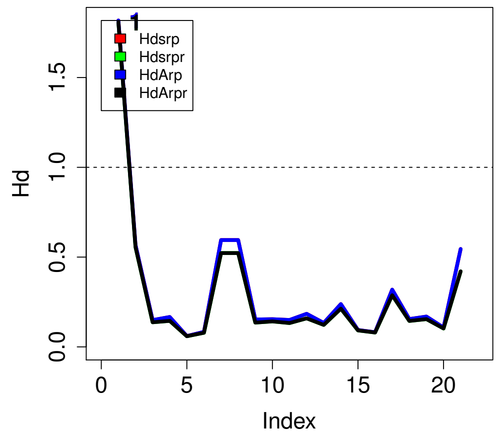
(d) IGRM CVR index plot



(e) IGRM DFFITS index plot



(f) IGRM WD index plot



(a) IGRM H_d index plot

Figure 4.2: The IGRM's Influence Diagnostic Methods Index Plots

Table 4.6: Influential Observations with the IGRM and the IGRRM Deviance Residuals

	Standardized Deviance Residuals		Adjusted Deviance Residuals	
Methods	sr_{di}	sr_{dRi}	Ar_{di}	Ar_{dRi}
CD	1	1	Nil	Nil
MCD	1	1	Nil	Nil
AP	9,13	1,9	1,9,13,21	1,8,9
CVR	1,2,4,7,8,16,21	1,2,4,7,8,16,21	1,2,4,7,8,16,21	1,2,4,7,8,16,21
$DFFITS$	1	1	4	4
WD	1	1	1	1
H_d	Nil	Nil	1	1

On studying the influence diagnostics with four forms of deviance residuals, Table 4.6 results and Fig.4.2 (a-g) have shown the similar detection of the influential observations as identified with four forms of Pearson residuals with the addition of 16$^{\text{th}}$ observation. These observations diagnose differently by each method as compared to the detection with Pearson residuals. With the standardized deviance residuals, influence measure H_d fails to detect any influential observation. The CD

and MCD fail to detect any influential observations with adjusted deviance and the adjusted ridge deviance residuals. While these two methods along with $DFFITS$ perform identically and diagnosed 1^{st} observation as the influential observation with standardized deviance as well as with the standardized ridge deviance residuals. We also observe that $DFFITS$ and WD methods with all forms of deviance residuals identically detect 1^{st} observation as the influential observation. The best method is the CVR, which diagnoses identical and maximum influential observations with all forms of deviance residuals. This method with the all forms of deviance residuals detects the similar influential observation as computed with the all forms of Pearson residuals but with addition of 16^{th} observation diagnosed as the influential observation. When we studying the influence diagnostics with the likelihood residuals in the stack loss data, we find the similar observation as we diagnosed with the Pearson residuals. From Table 4.7 and Fig.4.2 (a-g), we find the influence diagnostic performance of WD method is identical to results as we find with Pearson and deviance residuals. This detects only method

Table 4.7: Influential Observations with the IGRM and the IGRRM Likelihood Residuals

	Likelihood Residuals		Adjusted Likelihood Residuals	
Methods	r_{li}	r_{lRi}	Ar_{li}	Ar_{lRi}
CD	1	Nil	Nil	Nil
MCD	1	1	4	4
AP	9,13	1,9	1,9,13,21	1,8,9
CVR	1,2,7,8,21	1,2,7,8,21	1,2,7,8,21	1,2,7,8,21
$DFFITS$	1	Nil	4	Nil
WD	1	1	1	1
H_d	Nil	Nil	1	1

Table 4.8: Influential Observations with the IGRM and the IGRRM Anscombe Residuals

	Standardized Anscombe Residuals		Adjusted Anscombe Residuals	
Methods	sr_{ai}	sr_{aRi}	Ar_{ai}	Ar_{aRi}
CD	1	1	Nil	Nil
MCD	1	1	4	Nil
AP	9,13	1,9	1,9,13,21	1,8,9,13
CVR	1,2,7,8,21	1,2,7,8,21	1,2,7,8,21	1,2,7,8,21
$DFFITS$	1	1	Nil	Nil
WD	1	1	1	1
H_d	Nil	Nil	1	1

one observation i.e. 1st as influential observation. Again the CVR is the best method for the detection of influential observation as computed with the likelihood residuals. The CD method detects the 1st observation only with the likelihood residuals but fails to detect any influential observation as computed with three forms (r_{lR}, Ar_l and Ar_{lR}) of the likelihood residuals. The H_d method influence detection performance with the likelihood residuals is similar as computed with the Pearson and deviance residuals. When we study the influence diagnostics using Anscombe residuals, we find the similar influential observation as detected with Pearson and likelihood residuals. Also the detection performance of AP, CVR, WD and H_d methods is identical to that as computed with the Pearson and the likelihood residuals. From Table 4.8 and Fig.4.2 (a-g), we find the influential observation detection performance of CD, MCD and $DFFITS$ methods is identical and diagnoses 1st observation as influential observation.

On studying the influence diagnostics with the working residuals, we find that the influential observation detection is not better. Because AP method with ridge working residuals and with the adjusted working residuals indicates that all observations are

Table 4.9: Influential Observations with the IGRM and the IGRRM Working Residuals

	Standardized Working Residuals		Adjusted Working Residuals	
Methods	sr_{Wi}	sr_{WRi}	Ar_{Wi}	Ar_{WRi}
CD	Nil	Nil	Nil	Nil
MCD	Nil	Nil	Nil	Nil
AP	Nil	2-6,9-12,14-20	1,2,7,8,9,13,21	20-Jan
CVR	1,2,7,8,17,21	1,2,7,8,17,21	1,2,7,8,21	1,2,7,8,21
DFFITS	Nil	Nil	Nil	Nil
WD	Nil	Nil	Nil	Nil
H_d	Nil	Nil	1	1

influential, this is not true. From Table 4.9 and Fig.4.2 (a-g), we also find another drawback with this type residuals that the CD, MCD, $DFFITS$ and WD fail to detect any influential observation with any form of working residuals. The only diagnostic method that is the CVR method which diagnose the influential observation in better way. The CVR method diagnoses additionally the 17$^{\text{th}}$ observation as detected with the standardized Pearson and the standardized ridge Pearson residuals. This observation is not diagnosed by any form of the IGRM and IGRRM residuals and by any influence diagnostic methods. Also note that the 17$^{\text{th}}$ observation is also detected in the literature as given in Table 4.4.

Now we check the actual effect of these diagnosed influential observations with different residuals on the IGRRM estimates and inferences. These effects are given in Table 4.10 and now we compare the diagnostic methods with IGRM different residuals to determine which one detected the influential observation accurately. And also compare the performance of all IGRM residuals for the detection of influential observation. From Table 4.10, we find that top most influential is the 4$^{\text{th}}$ observation as identified by some IGRM methods with some residuals form. The CD method detects this top influential observation only with the adjusted Pearson and adjusted ridge Pearson residuals. While this method is unable to detect the other forms of IGRM residuals. This indicates that the CD method may be good in detecting the influential observations with the adjusted forms of the Pearson residuals. The MCD and $DFFITS$ equally diagnose the 4^{th} observation with the three forms of Pearson (sr_{pR}, Ar_p and Ar_{pR}) residuals. While these two methods do not detect this influential observation consistently with all other forms of the IGRM and the IGRRM residuals. This observation is detected efficiently by CVR method with the Pearson and deviance forms of residuals. While this method fails to detect 4^{th} observation with all forms of likelihood, Anscombe and working residuals. We also find that this most influential observation is not detected by the AP, WD and H_d methods with any kind of the IGRM residuals. This top most influential observation affects the IGRM and the IGRRM estimates of β_1 and β_{1R}. The 2^{nd} most influential obsevation is the 17^{th} observation as diagnosed by the CVR method only with the standardized working residuals and also with ridge standardized working residuals. While the other IGRM and IGRRM diagnostics techniques with all other residuals fail to detect the 2^{nd} most influential observation. This observation affects the IGRM and the IGRRM estimates of β_3. The 3^{rd} influential observation is the 7^{th} observation as detected only by CVR method with all forms of the IGRM and IGRRM residuals. Similarly the 8^{th} observations is the 4^{th} ranked influential observation. This observation is identified by CVR method with all IGRM and IGRRM residuals and also this observation is diagnosed by AP method with all adjusted forms of residuals with ridge estimates except the Ar_{pR}. This observation affects the IGRM estimate of β_1. While 4^{th}

most influential observations for the IGRRM estimate is the 21st observation. This observation is detected again by the CVR method with all forms of the IGRM and the IGRM residuals. This observation affects the IGRRM estimate of β_1. We also find that 2nd, 9th and 13th observations have minor substantial effect on the IGRM and the IGRRM estimates. These observations are detected by the AP method with all forms of the IGRM and the IGRRM residuals. These results are showing that this method detected wrong influential observations. The best method is the CVR which perform excellently to detect the influential observation with all form of the IGRM residuals.

Table 4.10: Absolute percentage change in the IGRM d the IGRRM estimates due to Influential Observations

Influential	IGRM Estimates				IGRRM Estimates			
Observations	β_0	β_1	β_2	β_3	β_{0R}	β_{R1}	β_{R2}	β_{R3}
1	0.45	**12.45**	0.95	3.26	0.36	**9.3**	1.49	3.07
2	0.29	2.06	1.25	**4.77**	0.36	2.92	0.75	**5.34**
4	**8.37**	**105.22**	**40.45**	18.4	**8.55**	**77.69**	**34.72**	17.92
7	0.77	**20.45**	6.84	9.45	0.6	**12.62**	5.07	7.7
8	0.42	**11.11**	3.71	5.12	0.27	**4.5**	1.71	2.7
9	0.09	1.08	0.4	0.18	0.15	0.02	0.22	0.51
13	0.21	1.03	0.77	0.65	0.56	2.18	2.06	1.95
16	5.04	4.75	**6.61**	4.69	4.94	**6.36**	5.74	4.01
17	4.06	14.32	4.68	**27.43**	4.15	9.53	2.97	**28.47**
21	0.31	**9.92**	5.81	2.53	0.59	**10.76**	8.08	2.65

Table 4.11: Performance of the IGRM and the IGRRM Influence Measures with Standardized Pearson Residuals

ρ	ϕ	CD	MCD	AP	CVR	$DFFITS$	WD	H_d	CD_R	MCD_R	AP_R	CVR_R	$DFFITS_R$	WD_R	H_{dR}
0.75	0.06	100	100	100	100	100	100	52.4	100	100	100	100	100	100	0.1
	0.11	99.8	99.8	100	100	99.8	100	48.8	100	100	100	100	100	100	1.1
	0.25	99.8	99.8	100	100	99.8	100	43.4	99.8	99.8	100	99.7	99.8	100	5.0
	0.50	98.5	98.5	100	100	98.5	100	40.1	99.0	99.0	99.4	98.9	99.0	99.3	12.5
	2.00	99.8	99.8	100	100	99.8	100	44.3	99.9	99.9	100	100	99.9	100	4.1
	20.0	87.0	87.5	98.3	97.1	87.5	93.3	12.2	54.0	54.4	62.5	67.2	54.4	63.4	9.1
	200	80.9	80.9	93.5	92.3	80.9	88.6	10.2	38.7	38.7	49.8	54.1	38.7	46.6	11.3
0.85	0.06	99.3	99.3	100	100	99.3	100	50.6	100	100	100	100	100	100	0.0
	0.11	99.6	99.6	100	100	99.6	100	47.8	100	100	100	100	100	100	0.6
	0.25	99.3	99.3	100	100	99.3	99.9	41.0	99.6	99.6	99.9	99.3	99.6	99.9	3.2
	0.50	98.4	98.4	100	100	98.4	99.9	39.1	99.1	99.1	99.4	98.5	99.1	99.3	12.6
	2.00	99.6	99.6	100	100	99.6	100	42.9	100	100	100	100	100	100	3.1
	20.0	87.9	88.0	97.7	96.3	88.0	94.9	12.4	52.1	52.3	59.3	64.1	52.3	61.1	9.4
	200	77.2	77.4	90.3	90.4	77.4	87.4	10.0	40.9	41.1	50.8	54.9	41.1	48.7	9.3
0.95	0.06	100	100	100	100	100	100	48.8	100	100	100	100	100	100	0.1
	0.11	99.7	99.7	100	100	99.7	100	45.4	100	100	100	99.9	100	100	1.4
	0.25	99.1	99.1	100	100	99.1	100	39.8	100	100	100	99.6	100	100	3.3
	0.50	98.7	98.7	100	100	98.7	100	36.4	99.7	99.7	99.9	99.3	99.7	99.9	10.9
	2.00	99.7	99.7	100	100	99.7	100	45.8	100	100	100	99.9	100	100	2.7
	20.0	86.8	86.9	98.2	97.1	86.9	94.8	14.7	59.9	60.0	65.5	70.1	60	67.3	9.1
	200	78.6	78.7	90.7	90.9	78.7	86.3	12.3	46.1	46.3	53.7	58.2	46.3	52.5	10.3
0.99	0.06	99.7	99.7	100	100	99.7	100	47.3	100	100	100	100	100	100	0.2
	0.11	99.8	99.8	100	100	99.8	100	46.7	100	100	100	99.9	100	100	1.1
	0.25	99.3	99.3	100	100	99.3	99.9	41.9	99.9	99.9	100	99.9	99.9	100	4.3
	0.50	98.3	98.3	100	100	98.3	100	37.8	99.7	99.7	100	99.4	99.7	100	8.0
	2.00	99.7	99.7	100	100	99.7	100	44.2	100	100	100	99.7	100	100	2.3
	20.0	86.5	86.6	97.8	96.7	86.6	94.8	18.8	70.8	70.8	75	78.3	70.8	76.0	10.8
	200	77.6	77.7	91.5	90.7	77.7	87.8	13.4	54.8	54.9	61.7	65.8	54.9	61.6	13.6

Table 4.12: Performance of the IGRM and the IGRRM Influence Measures with Standardized Deviance Residuals

ρ	ϕ	CD	MCD	AP	CVR	$DFFITS$	WD	H_d	CD_R	MCD_R	AP_R	CVR_R	$DFFITS_R$	WD_R	H_{dR}
0.75	0.06	99.7	99.7	100	100	99.7	100	47.5	100	100	100	100	100	100	0
	0.11	99.7	99.7	100	100	99.7	100	49.1	100	100	100	100	100	100	0.5
	0.25	99.4	99.5	100	100	99.5	100	42.5	99.9	99.9	100	99.6	99.9	100	2.4
	0.50	98.7	98.7	100	100	98.7	99.8	38.9	99.0	99.0	99.2	97.9	99.0	99.2	6.4
	2.00	99.5	99.5	100	100	99.5	100	45.7	100	100	100	99.8	100	100	1.4
	20.0	86.1	86.3	98.2	96.8	86.3	94.8	12.7	56.2	56.4	57.6	57	56.4	65.0	13
	200	76.3	76.3	89.2	89.1	76.3	85.9	9.4	49.8	50	45.7	46.1	50	56.4	11.6
0.85	0.06	99.4	99.4	100	100	99.4	100	46.8	100	100	100	100	100	100	0
	0.11	99.7	99.7	100	100	99.7	100	45.1	100	100	100	100	100	100	0.7
	0.25	99.4	99.4	100	100	99.4	100	42.5	100	100	100	100	100	100	1.9
	0.50	98.7	98.7	100	100	98.7	100	37.4	98.9	98.9	99.3	98.7	98.9	99.2	5
	2.00	99.5	99.5	100	100	99.5	100	42.9	99.9	99.9	100	99.8	99.9	100	1.2
	20.0	86.0	86.3	98.6	96.9	86.3	94.2	12.5	59.6	59.7	62	59.7	59.7	67.9	13.2
	200	77.7	78.0	91.5	90.6	78.0	87.1	10.2	54.9	55.0	49.8	49.7	55.0	61.4	12.0
0.95	0.06	99.5	99.5	100	100	99.5	100	49.4	100	100	100	100	100	100	0.0
	0.11	99.4	99.4	100	100	99.4	100	45.8	100	100	100	100	100	100	0.4
	0.25	99.4	99.4	100	100	99.4	100	38.4	99.8	99.8	100	99.7	99.8	100	1.3
	0.50	98.3	98.3	100	100	98.3	99.8	35.9	100	100	100	99.3	100	100	3.6
	2.00	99.3	99.3	100	100	99.3	100	43.4	99.9	99.9	100	100	99.9	100	0.7
	20.0	85.8	85.8	97.9	96.8	85.8	93.6	14.7	66.2	66.3	65.6	63.4	66.3	72.1	18.0
	200	78.4	78.4	91.9	91.1	78.4	87.7	13.3	60.3	60.5	53.6	53.2	60.5	65.2	14.5
0.99	0.06	99.9	99.9	100	100	99.9	100	46.8	100	100	100	100	100	100	0.0
	0.11	99.3	99.3	100	100	99.3	100	46.1	100	100	100	100	100	100	0.4
	0.25	99	99.1	100	100	99.1	100	40.2	100	100	100	99.8	100	100	1.1
	0.50	97.9	97.9	100	100	97.9	99.9	35.5	99.8	99.8	99.9	99.5	99.8	99.9	2.1
	2.00	99.5	99.5	100	100	99.5	100	44.2	100	100	100	99.8	100	100	0.9
	20.0	87.1	87.1	97.9	96.9	87.1	93.4	14.3	73.7	73.8	74.9	72.3	73.8	79.4	16.9
	200	80.0	80.3	92.5	91.6	80.3	88.9	15.0	69.4	69.5	63.8	63.3	69.5	75.3	16.1

Table 4.13: Performance of the IGRM and the IGRRM Influence Measures with Likelihood Residuals

ρ	ϕ	CD	MCD	AP	CVR	$DFFITS$	WD	H_d	CD_R	MCD_R	AP_R	CVR_R	$DFFITS_R$	WD_R	H_{dR}
0.75	0.06	1.1	1.2	100	100	1.2	99.5	47.7	100	100	100	100	100	100	0.0
	0.11	4.4	4.6	100	100	4.6	99.1	47.1	100	100	100	100	100	100	0.5
	0.25	18.0	18.7	100	100	18.7	99.1	42.6	99.9	99.9	99.9	99.7	99.9	99.9	2.8
	0.50	36.0	36.8	100	100	36.8	98.7	40.0	99.6	99.6	99.6	99	99.6	99.6	7.1
	2.00	12.4	12.9	100	100	12.9	99.1	44.2	99.9	99.9	100	99.7	99.9	100	1.3
	20.0	87.0	87.1	98.2	97.9	87.1	95.9	11.2	56.4	56.6	58.7	56.9	56.6	63.1	17.4
	200	88.2	88.2	90.2	95.1	88.2	92.7	11.7	48.8	48.9	44.3	43.4	48.9	55.5	14.9
0.85	0.06	0.7	0.7	100	100	0.7	99.2	49.0	100	100	100	100	100	100	0.1
	0.11	5.1	5.5	100	100	5.5	99.4	45.0	100	100	100	100	100	100	0.4
	0.25	16.5	16.8	100	100	16.8	99.4	42.1	99.9	99.9	100	99.6	99.9	100	2.0
	0.50	36.0	36.4	100	100	36.4	98.9	38.6	99.3	99.3	99.8	99.0	99.3	99.8	5.6
	2.00	11.9	12.8	100	100	12.8	99.4	45.3	100	100	100	99.8	100	100	1.5
	20.0	86.8	87.	98.7	97.6	87.0	96.2	14.7	56.3	56.3	59.2	59.2	56.3	64.1	16.9
	200	89.6	89.6	92.4	95.9	89.6	93.5	10.3	54.1	54.3	52.1	50.2	54.3	63.1	15.1
0.95	0.06	1.1	1.1	100	100	1.1	99.3	46.6	100	100	100	100	100	100	0.1
	0.11	4.1	4.1	100	100	4.1	99.4	42.9	100	100	100	100	100	100	0.5
	0.25	13.9	14.4	100	100	14.4	98.7	40.4	99.7	99.8	100	99.5	99.8	100	1.6
	0.5	30.6	31.3	100	100	31.3	98.8	35.9	99.9	99.9	99.9	99.2	99.9	99.9	4.1
	2.00	10.8	11.0	100	100	11.0	98.8	42.6	99.9	99.9	100	99.7	99.9	100	0.8
	20.0	84.4	84.5	98.3	97.1	84.5	95.8	14.7	65.5	65.5	66.7	64.7	65.5	70.3	20.6
	200	89.7	89.7	92.3	96.1	89.7	94.2	14.7	57.2	57.4	53.1	54.1	57.4	64.9	18.3
0.99	0.06	1.1	1.1	100	100	1.1	99.1	46.9	100	100	100	100	100	100	0.1
	0.11	4.3	4.5	100	100	4.5	99.5	44.3	100	100	100	100	100	100	0.5
	0.25	14.8	15.1	100	100	15.1	99.1	41.9	100	100	99.9	99.5	100	100	0.9
	0.50	29.1	29.8	100	100	29.8	98.9	35.9	99.9	99.9	100	98.9	99.9	100	3.0
	2.00	10.7	10.8	100	100	10.8	98.4	40.9	99.9	99.9	100	99.9	99.9	100	0.8
	20.0	83.6	83.7	98.2	97.2	83.7	96.3	14.9	74.4	74.5	75.7	75.6	74.5	79.5	23.1
	200	89.3	89.3	92.3	95.8	89.3	93.9	14.1	64.0	64.2	61.3	62.7	64.2	70.3	22.4

Table 4.14: Performance of the IGRM and the IGRRM Influence Measures with Standardized Anscombe Residuals

ρ	ϕ	CD	MCD	AP	CVR	$DFFITS$	WD	H_d	CD_R	MCD_R	AP_R	CVR_R	$DFFITS_R$	WD_R	H_{dR}
0.75	0.06	99.6	99.6	100	100	99.6	100	51.2	100	100	100	100	100	100	0.2
	0.11	99.6	99.6	100	100	99.6	100	48.8	100	100	100	100	100	100	0.3
	0.25	99.6	99.6	100	100	99.6	100	47.1	99.9	99.9	100	99.6	99.9	100	2.2
	0.50	99.4	99.4	100	100	99.4	100	40.2	98.9	98.9	99.2	97.8	98.9	99.2	5.2
	2.00	99.3	99.4	100	100	99.4	100	43.2	99.9	99.9	100	99.9	99.9	100	2.1
	20.0	88.5	88.7	98.8	98.0	88.7	95.3	12.6	57.0	57.3	60.3	55	57.3	64.3	15.7
	200	79.6	79.7	92.0	91.4	79.7	87.6	9.8	49.0	49.2	51.2	44.3	49.2	56.1	11.8
0.85	0.06	99.7	99.7	100	100	99.7	100	47.3	100	100	100	100	100	100	0.0
	0.11	99.5	99.5	100	100	99.5	100	46.2	100	100	100	100	100	100	0.6
	0.25	99.6	99.6	100	100	99.6	100	41.1	99.8	99.8	100	99.9	99.8	100	2.0
	0.50	99.0	99	100	100	99	99.9	34.2	99.2	99.2	99.3	98.7	99.2	99.3	3.7
	2.00	99.6	99.6	100	100	99.6	100	43.0	99.8	99.8	100	100	99.8	100	1.3
	20.0	86.3	86.3	98.6	97.7	86.3	94.8	12.8	58.7	58.9	61.3	55.8	58.9	64.9	16.8
	200	78.9	79.0	91.2	89.9	79.0	86.6	10.6	49.2	49.3	52.6	45.3	49.3	56.2	11.7
0.95	0.06	99.9	99.9	100	100	99.9	100	47.5	100	100	100	100	100	100	0.1
	0.11	99.8	99.8	100	100	99.8	100	45.8	100	100	100	100	100	100	0.6
	0.25	99.0	99.0	100	100	99.0	100	39.4	99.9	99.9	100	99.4	99.9	100	0.6
	0.5	98.6	98.7	100	100	98.7	99.9	34.5	99.5	99.5	99.5	98.6	99.5	99.6	3.2
	2.0	99.4	99.4	100	100	99.4	100	42.7	100	100	100	99.8	100	100	0.7
	20 .0	85.6	85.7	98.0	96.4	85.7	94.8	16.0	66.8	66.8	70.2	64.2	66.8	72.8	18.5
	200	79.4	79.5	90.6	89.8	79.5	88.2	11.4	54.4	54.5	56.3	49.4	54.5	60.2	16.3
0.99	0.06	99.7	99.7	100	100	99.7	100	47.6	100	100	100	100	100	100	0.1
	0.11	99.8	99.8	100	100	99.8	100	43.2	100	100	100	100	100	100	0.3
	0.25	98.7	98.7	100	100	98.7	100	40.9	100	100	100	100	100	100	2.0
	0.50	98.9	98.9	100	100	98.9	99.6	33.4	99.8	99.8	99.9	99.1	99.8	99.9	1.7
	2.00	99.2	99.2	100	100	99.2	100	42.6	100	100	100	100	100	100	1.2
	20.0	87.2	87.2	98.5	97.3	87.2	95.2	15.1	74.1	74.2	77.3	74	74.2	79.1	17.4
	200	79.7	79.9	92.5	92.2	79.9	87.9	15.4	65.0	65.1	68.9	61	65.1	71.2	17.3

Table 4.15: Performance of the IGRM and the IGRRM Influence Measures with Standardized Working Residuals

ρ	ϕ	CD	MCD	AP	CVR	$DFFITS$	WD	H_d	CD_R	MCD_R	AP_R	CVR_R	$DFFITS_R$	WD_R	H_{dR}
0.75	0.06	99.7	99.7	100	100	99.7	100	49.2	100	100	100	100	100	100	0
	0.11	99.4	99.4	100	100	99.4	100	47.4	99.9	99.9	100	100	99.9	100	0.1
	0.25	99.2	99.2	100	100	99.2	99.9	42.0	99.9	99.9	100	99.9	99.9	100	1.5
	0.50	98.4	98.4	100	100	98.4	99.9	39.0	99.1	99.1	99.1	98.6	99.1	99.1	4.5
	2.00	99.3	99.3	100	100	99.3	100	45.1	100	100	100	99.9	100	100	1.5
	20.0	86.3	86.6	98.6	97.5	86.6	93.9	13.00	71.3	71.3	70.2	72	71.3	72.4	11.0.0
	200	80.1	80.3	89.9	89.8	80.3	87.5	9.8	66.8	66.8	62.0.0	68.8	66.8	68.4	5.4
0.85	0.06	99.5	99.5	100	100	99.5	100	46.7	100	100	100	100	100	100	0.0
	0.11	99.9	99.9	100	100	99.9	100	49.2	100	100	100	100	100	100	0.2
	0.25	99.2	99.2	100	100	99.2	99.9	41.1	100	100	100	99.9	100	100	1.6
	0.50	98.5	98.5	100	100	98.5	99.8	36.4	99.4	99.4	99.5	99.3	99.4	99.5	2.7
	2.00	99.1	99.1	100	100	99.1	100	44.5	99.9	99.9	100	99.9	99.9	100	1.1
	20.0	87.4	87.6	97.8	96.6	87.6	94.4	12.7	69.5	69.5	67.4	70.7	69.5	70.5	9.6
	200	79.2	79.4	90.6	89.8	79.4	86.9	11.1	68.2	68.2	65.4	69.9	68.2	68.9	5.7
0.95	0.06	100	100	100	100	100	100	47.6	100	100	100	100	100	100	0.0
	0.11	99.1	99.1	100	100	99.1	100	45.2	100	100	100	100	100	100	0.1
	0.25	99.2	99.2	100	100	99.2	99.9	41.4	99.9	99.9	100	99.8	99.9	100	1.2
	0.50	99.0	99.0	100	100	99	99.9	36.8	99.6	99.6	99.7	99.4	99.6	99.7	1.5
	2.00	99.8	99.8	100	100	99.8	100	43.9	100	100	100	100	100	100	1
	20.0	85.2	85.4	98	96.9	85.4	93.4	14.7	77.9	77.9	74	77.4	77.9	78.5	8.8
	200	79.9	80.0	92.2	91.4	80	88.4	14.1	73.7	73.7	68.3	75.1	73.7	74.7	6.6
0.99	0.06	99.7	99.7	100	100	99.7	100	45.9	100	100	100	100	100	100	0
	0.11	99.3	99.4	100	100	99.4	100	44.1	100	100	100	100	100	100	0.1
	0.25	99.3	99.3	100	100	99.3	100	39.7	99.9	99.9	100	99.9	99.9	100	0.9
	0.50	99.1	99.1	100	99.8	99.1	100	36.9	99.8	99.8	99.9	99.6	99.8	99.9	1.5
	2.00	99.3	99.3	100	100	99.3	100	43.8	100	100	100	100	100	100	0.4
	20.0	89.1	89.1	97.6	96.8	89.1	94.9	15.3	84.9	85.0	81.2	84.8	85	85.4	5.2
	200	80.3	80.4	92.5	90.9	80.4	88.7	13.5	78.1	78.1	73.2	79.7	78.1	79.3	2.1

Table 4.16: Performance of the IGRM and the IGRRM Influence Measures with Adjusted Pearson Residuals

ρ	ϕ	CD	MCD	AP	CVR	$DFFITS$	WD	H_d	CD_R	MCD_R	AP_R	CVR_R	$DFFITS_R$	WD_R	H_{dR}
0.75	0.06	100	100	100	100	100	100	46.5	99.7	99.7	100	100	99.7	100	12.0
	0.11	100	100	100	100	100	100	48.8	99.5	99.5	100	100	99.5	100	13.0
	0.25	100	100	100	100	100	100	46.9	98.7	98.7	99.9	99.9	98.7	99.9	30.2
	0.50	100	100	100	100	100	100	39.6	97.8	97.8	99.4	99.4	97.8	99.4	52.5
	2.00	100	100	100	100	100	100	45.4	98.9	98.9	100	100	98.9	100	23.4
	20.0	90.8	90.8	98.1	98.1	90.8	96.3	12.8	42.6	43.1	59.4	60.2	43.1	55.9	17.8
	200	82.7	83.0	90.6	91.5	83.0	88.0	10.9	46.4	46.5	53.1	54.4	46.5	54.6	20.1
0.85	0.06	100	100	100	100	100	100	46.5	99.2	99.2	100	100	99.2	100	10.2
	0.11	100	100	100	100	100	100	44.2	98.9	98.9	100	100	98.9	100	12.9
	0.25	99.9	99.9	100	100	99.9	100	42.7	99.1	99.1	100	100	99.1	100	29.7
	0.50	99.6	99.6	100	100	99.6	100	38.9	97.9	97.9	99.2	99.2	97.9	99.2	48.7
	2.00	100	100	100	100	100	100	44.5	99.0	99.0	100	100	99.0	100	23.0
	20.0	90.2	90.2	98.1	98.0	90.2	95.1	11.8	43.8	43.8	61.1	60.7	43.8	57.4	20.4
	200	81.2	81.4	90.9	91.0	81.4	87.8	12.0	48.0	48.0	54.1	53.4	48.0	55.4	22.3
0.95	0.06	100	100	100	100	100	100	46.2	99.1	99.1	100	100	99.1	100	8.8
	0.11	100	100	100	100	100	100	45.2	99.5	99.5	100	100	99.5	100	11.8
	0.25	99.9	99.9	100	100	99.9	99.9	42.6	99.1	99.1	100	100	99.1	100	21
	0.50	99.7	99.7	100	100	99.7	99.9	35.1	98.6	98.6	99.7	99.7	98.6	99.7	47.3
	2.00	100	100	100	100	100	100	44.8	99.2	99.2	100	100	99.2	100	15.5
	20.0	90.2	90.3	98.3	97.8	90.3	96.2	15.	50.6	50.9	67.4	67.7	50.9	63.1	23.2
	200	82.5	82.6	92.4	92.5	82.6	90	11.3	50.5	50.7	57.5	57.1	50.7	57.5	23.4
0.99	0.06	100	100	100	100	100	100	49.8	99.1	99.1	100	100	99.1	100	7.8
	0.11	99.9	99.9	100	100	99.9	100	46	99.3	99.3	100	100	99.3	100	10.1
	0.25	99.8	99.8	100	100	99.8	99.9	39.4	99.1	99.1	100	100	99.1	100	19.2
	0.50	99.8	99.8	100	100	99.8	100	37.7	99.0	99.0	99.9	99.9	99	99.9	42.1
	2.00	99.9	99.9	100	100	99.9	100	43.3	99.2	99.2	100	100	99.2	100	13.9
	20.0	90.9	91	98.2	98.4	91	96.3	15.9	60.1	60.3	75.1	74.5	60.3	71.7	26.5
	200	81.6	81.7	90.9	91.3	81.7	87.8	12.3	58.9	58.9	63.6	61.5	58.9	66.3	27.3

Table 4.17: Performance of the IGRM and the IGRRM Influence Measures with Adjusted Deviance Residuals

ρ	ϕ	CD	MCD	AP	CVR	$DFFITS$	WD	H_d	CD_R	MCD_R	AP_R	CVR_R	$DFFITS_R$	WD_R	H_{dR}
0.75	0.06	100	100	100	100	100	100	48.0	99.	98.9	100	100	98.9	100	15.0
	0.11	100	100	100	100	100	100	45.0	98.0	98.2	100	100	98.2	100	19.0
	0.25	100	100	100	100	100	100	42.0	98.0	98.2	100	100	98.2	100	34.0
	0.50	100	99.8	100	100	99.8	100	40.0	96	96.1	99.0	99.3	96.1	99.0	36.0
	2.00	100	100	100	100	100	100	43.0	99.0	98.7	100	100	98.7	100	28.0
	20.0	91.0	90.8	98.0	98.0	90.8	95.0	15.0	45.0	45.2	59.0	61.0	45.2	55.0	16.0
	200	83.0	82.6	90.0	90.0	82.6	88	9.6	34.0	33.9	48.0	50.6	33.9	44.0	12.0
0.85	0.06	100	100	100	100	100	100	46.0	99.0	99.3	100	100	99.3	100	15.0
	0.11	100	100	100	100	100	100	46.0	99.0	98.5	100	100	98.5	100	17.0
	0.25	100	99.7	100	100	99.7	100	45.0	97.0	97.3	100	99.9	97.3	100	31.0
	0.50	100	99.8	100	100	99.8	100	37.0	96.0	96.2	100	99.5	96.2	99.0	36.0
	2.00	100	100	100	100	100	100	44.0	99.0	98.5	100	100	98.5	100	28.0
	20.0	91.0	90.5	99.0	98.0	90.5	96.0	14.0	45.0	45.3	60.0	61.0	45.3	55.0	18.0
	200	82.0	82.1	91.0	92.0	82.1	88.0	10.0	37.0	36.8	49.0	51.0	36.8	47.0	14.0
0.95	0.06	100	100	100	100	100	100	45.0	100	99.5	100	100	99.5	100	14.0
	0.11	100	100	100	100	100	100	48.0	99.0	98.6	100	100	98.6	100	17.0
	0.25	100	100	100	100	100	100	42.0	98.0	98.4	100	100	98.4	100	28.0
	0.50	100	99.8	100	100	99.8	100	35.0	97.0	96.7	100	100	96.7	100	36.0
	2.00	100	100	100	100	100	100	45.0	98.0	98.4	100	100	98.4	100	24.0
	20.0	91.0	91.5	99.0	99.0	91.5	96	14.0	53.0	53.2	67	68.9	53.2	63.0	20.0
	200	83.0	83.0	92.0	92.0	83.0	90.0	14.0	38.0	37.7	49.0	51.8	37.7	47.0	17.0
0.99	0.06	100	100	100	100	100	100	44.0	99.0	99.2	100	100	99.2	100	14.0
	0.11	100	99.9	100	100	99.9	100	45.0	99.0	98.9	100	100	98.9	100	16.0
	0.25	100	99.9	100	100	99.9	100	43.0	98.0	98.4	100	100	98.4	100	32.0
	0.50	99.0	99.5	100	100	99.5	100	35.0	97.0	97.2	100	99.8	97.2	100	36.0
	2.00	100	100	100	100	100	100	43.0	99.0	98.5	100	100	98.5	100	20.0
	20.0	90.0	90.0	99.0	98.0	90.0	96.0	17.0	59.0	59.3	74	74.1	59.3	70.0	24.0
	200	82.0	82.2	92.0	93.0	82.2	89.0	15.0	49.0	49.0	63.0	64.7	49.0	61.0	22.0

Table 4.18: Performance of IGRM and IGRRM Influence Measures with Adjusted Likelihood Residuals

ρ	ϕ	CD	MCD	AP	CVR	$DFFITS$	WD	H_d	CD_R	MCD_R	AP_R	CVR_R	$DFFITS_R$	WD_R	H_{dR}
0.75	0.06	100	100	100	100	100	100	50.1	100	100	100	100	100	100	0.1
	0.11	100	100	100	100	100	100	47.1	100	100	100	99.0	100	100	0.7
	0.25	100	100	100	100	100	100	43.9	99.9	99.9	100	99.5	99.9	100	2.3
	0.50	99.9	99.9	100	100	99.9	100	40.1	98.0	98.0	98.6	97	98	98.5	5.9
	2.00	100	100	100	100	100	100	46.1	100	100	100	99.7	100	100	0.8
	20.0	90.4	90.5	97.7	97.9	90.5	96.3	11.6	42.4	42.5	58.2	61.1	42.5	52	14.6
	200	82.6	82.6	91.6	92.2	82.6	89.6	10.7	34.8	34.9	46.5	49.6	34.9	42.1	10.5
0.85	0.06	100	100	100	100	100	100	46.9	100	100	100	100	100	100	0
	0.11	100	100	100	100	100	100	46.3	100	100	100	99.8	100	100	0.5
	0.25	100	100	100	100	100	100	42.7	99.7	99.7	100	99.6	99.7	100	2
	0.50	99.9	99.9	100	100	99.9	100	39.3	99.3	99.3	99.5	98.7	99.3	99.5	6.2
	2.00	100	100	100	100	100	100	45.6	100	100	100	99.3	100	100	0.4
	20.0	89.3	89.3	97.7	97.2	89.3	94.5	14.2	45.1	45.1	62.1	63.9	45.1	54.6	14.1
	200	82.9	82.9	90.9	91.3	82.9	88.8	9.5	34.7	34.7	48.9	51.5	34.7	44.9	11.8
0.95	0.06	100	100	100	100	100	100	44.9	100	100	100	99.9	100	100	0.1
	0.11	100	100	100	100	100	100	44.7	100	100	100	99.7	100	100	0.4
	0.25	99.9	99.9	100	100	99.9	100	40.1	99.9	99.9	100	99.6	99.9	100	1.5
	0.50	99.6	99.6	100	100	99.6	100	37.5	99.6	99.6	99.9	99.4	99.6	99.9	4.3
	2.00	100	100	100	100	100	100	42.5	99.9	99.9	100	99.4	99.9	100	0.9
	20.0	90.3	90.3	98.1	98.1	90.3	95.7	14.9	49.3	49.3	63.1	64.8	49.3	57.9	14.1
	200	81.7	81.7	91.4	92.1	81.7	88.9	12.7	42.3	42.4	55.2	56.6	42.4	50.5	15.9
0.99	0.06	100	100	100	100	100	100	46.8	100	100	100	100	100	100	0
	0.11	100	100	100	100	100	100	45.4	100	100	100	99.6	100	100	0.1
	0.25	100	100	100	100	100	100	42	99.9	99.9	100	99.4	99.9	100	1
	0.50	99.6	99.6	100	100	99.6	100	35.7	99.3	99.3	99.4	99	99.3	99.4	2.5
	2.00	100	100	100	100	100	100	43	100	100	100	99.5	100	100	0.9
	20.0	90.7	90.8	98.3	98.4	90.8	95.1	17.6	60.7	60.8	74	74.6	60.8	70.4	18.7
	200	83.8	83.9	92.3	92.6	83.9	90.1	13	48.9	49.1	63.6	63.9	49.1	59.7	18.3

Table 4.19: Performance of IGRM and IGRRM Influence Measures with Adjusted Anscombe Residuals

ρ	ϕ	CD	MCD	AP	CVR	$DFFITS$	WD	H_d	CD_R	MCD_R	AP_R	CVR_R	$DFFITS_R$	WD_R	H_{dR}
0.75	0.06	100	100	100	100	100	100	49	99	98.8	100	100	98.8	99.9	15
	0.11	100	100	100	100	100	100	48	99	98.6	100	100	98.6	100	19
	0.25	100	100	100	100	100	100	45	98	98	100	100	98	100	31
	0.50	100	100	100	100	100	100	39	96	95.8	99	99.1	95.8	98.9	33
	2.00	100	100	100	100	100	100	45	98	97.8	100	100	97.8	99.9	26
	20.0	90.5	90.5	99	98.5	90.5	97	13	52	51.9	62	60	51.9	59.7	19
	200	83	83	93	93.1	83	89	10	51	51.4	59	56.1	51.4	59.1	24
0.85	0.06	100	100	100	100	100	100	47	99	98.6	100	100	98.6	100	13
	0.11	100	100	100	100	100	100	46	99	98.8	100	100	98.8	100	17
	0.25	99.9	99.9	100	100	99.9	100	42	98	98.1	100	99.9	98.1	99.9	30
	0.50	99.4	99.4	100	100	99.4	100	37	97	96.8	100	99.6	96.8	99.2	34
	2.00	100	100	100	100	100	100	44	99	98.6	100	100	98.6	99.9	25
	20.0	90.4	90.4	98	98.2	90.4	96	14	56	56.1	67	64.7	56.1	64	24
	200	83.3	83.4	92	92	83.4	89	11	56	56	64	62.7	56	61.8	22
0.95	0.06	100	100	100	100	100	100	50	98	98.5	100	100	98.5	100	14
	0.11	100	100	100	100	100	100	46	99	99.1	100	100	99.1	100	16
	0.25	99.9	99.9	100	100	99.9	100	41	98	98.3	100	100	98.3	100	28
	0.50	99.9	99.9	100	100	99.9	100	34	97	96.7	100	99.7	96.7	99.7	35
	2	99.9	99.9	100	100	99.9	100	44	98	97.9	100	100	97.9	99.8	23
	20	91.6	91.6	99	99.2	91.6	97	15	59	60	70	68.7	60	68.4	26
	200	80.4	80.6	91	92.1	80.6	87	14	59	59.5	65	63	59.5	64.3	28
0.99	0.06	100	100	100	100	100	100	52	99	99.4	100	100	99.4	100	14
	0.11	100	100	100	100	100	100	45	99	98.6	100	100	98.6	100	15
	0.25	99.9	99.9	100	100	99.9	100	40	99	99.1	100	100	99.1	100	28
	0.50	100	100	100	100	100	100	36	97	97.1	100	99.8	97.1	99.6	30
	2.00	100	100	100	100	100	100	44	99	98.9	100	100	98.9	100	20
	20.0	90	90.2	98	98.2	90.2	96	18	71	70.9	79	75.1	70.9	77	32
	200	80	80	90	90.4	80	88	13	64.0	63.8	70.0	67.4	63.8	68.4	31

Table 4.20: Performance of IGRM and IGRRM Influence Measures with Adjusted Working Residuals

ρ	ϕ	CD	MCD	AP	CVR	$DFFITS$	WD	H_d	CD_R	MCD_R	AP_R	CVR_R	$DFFITS_R$	WD_R	H_{dR}
0.75	0.06	100	100	100	100	100	100	51.5	97	97	100	100	97	100	52.9
	0.11	100	100	100	100	100	100	46.4	97	97.1	100	100	97.1	100	46.1
	0.25	100	100	100	100	100	100	42.2	96.3	96.3	100	100	96.3	99.9	43.4
	0.50	99.9	99.9	100	100	99.9	100	41.9	95.5	95.5	99.3	99.2	95.5	99.3	47.5
	2.00	100	100	100	100	100	100	44.1	96.2	96.2	100	100	96.2	100	42.7
	20.0	91.6	91.6	99	98.8	91.6	96.2	12.7	64	64.1	67.6	67.3	64.1	67	23.9
	200	81.8	81.8	91.2	91.1	81.8	87.9	10.7	62.9	62.9	65	64.9	62.9	63.8	16.7
0.85	0.06	100	100	100	100	100	100	48.8	96.3	96.3	100	100	96.3	100	57
	0.11	100	100	100	100	100	100	46.1	96.3	96.3	100	100	96.3	100	51.3
	0.25	99.8	99.8	100	100	99.8	100	43.7	97.1	97.2	100	100	97.2	100	45.5
	0.50	99.8	99.8	100	100	99.8	100	38.3	97.8	97.8	99.7	99.7	97.8	99.7	45.9
	2.00	100	100	100	100	100	100	45.6	97.2	97.2	100	100	97.2	100	44.9
	20.0	91.9	91.9	97.7	97.8	91.9	95.6	12.9	69.2	69.2	71.6	70.6	69.2	71	24.7
	200	82	82	91.9	92.3	82	88.7	10.2	64.7	64.7	67.6	67.4	64.7	65.7	16.1
0.95	0.06	100	100	100	100	100	100	48.6	95.8	95.8	100	100	95.8	99.9	59.7
	0.11	100	100	100	100	100	100	46.4	95.7	95.7	100	100	95.7	100	47.5
	0.25	99.9	99.9	100	100	99.9	100	42.6	96.3	96.4	100	100	96.4	99.9	43.9
	0.50	99.6	99.6	100	100	99.6	100	37.8	97.3	97.3	99.8	99.6	97.3	99.8	43.7
	2.00	100	100	100	100	100	100	41.3	95.9	95.9	100	100	95.9	100	45.4
	20.0	92.6	92.7	98.8	98.5	92.7	96.3	15	72.8	72.9	74.7	73	72.9	74.6	23.4
	200	81.8	81.9	91.8	92.2	81.9	88.5	12.6	67.4	67.4	69.1	68.4	67.4	68.2	16.9
0.99	0.06	100	100	100	100	100	100	44.9	95.7	95.8	100	100	95.8	100	59.2
	0.11	100	100	100	100	100	100	48.9	97.1	97.2	100	100	97.2	99.9	51.4
	0.25	100	100	100	100	100	100	39.5	95.9	95.9	100	100	95.9	100	41.5
	0.50	99.5	99.5	100	100	99.5	100	36.5	96.6	96.7	99.9	99.8	96.7	99.9	41.9
	2.00	99.9	99.9	100	100	99.9	100	39.2	96.4	96.5	100	100	96.5	100	43.7
	20.0	91.8	91.9	98.2	98.6	91.9	97.1	16	78.7	78.7	79.8	78.8	78.7	79.8	20.2
	200	80.7	80.7	91.7	92.4	80.7	89	17.5	71	71.1	72.2	71.3	71.1	71.6	10.7

Chapter 5

Conclusions and Future Research

This chapter summarizes some conclusions which are drawn from the previous Chapters (3 and 4) regarding the influential observation diagnostics in the GLM. It also covers the future dimensions regarding the influential observation detection in the GLM. The observation is said to be influential, if deletion of such observation substantially changes the predicted/forecasted values and other inferences about the GLM. In the GLM fitting, there are the situations, where the problems of influential observations and multicollinearity exist simultaneously. Both the influential observation and multicollinearity disturb the GLM estimates and other statistical measures substantially (Walker and Birch, 1988). However, if the researchers interest is to find the factors which are responsible to change the regression model, then due to these problems, significant role of some factor(s) remains hidden (Dunteman and Ho, 2006; Bonate, 2011; Amin *et al.*, 2015). The main interests of this thesis are the assessment of influential observations in some of the GLM with ridge estimate under various forms of the GLM residuals. There are two forms of residuals named as standardized and adjusted residuals. Cordeiro (2004) has shown that the adjusted Pearson residuals have zero mean and unit variance and also directs that the model diagnostics can be computed with adjusted forms of residuals. We extend the work of Cordeiro (2004) and derive the adjusted Pearson and adjusted deviance residuals (Sections 3.8). Then the computation of the influence diagnostics based on these two forms of residuals is proposed for the GRM and IGRM. For the assessment of influential observation, our emphasis is only on the detection of influential observations but not how to deal with them after detection. In these context, following conclusions are as under:

1. Sometimes, the influential observation may or may not arise due to misspecification of the statistical model. This problem is handled before the influence diagnostics as we include those data sets which are best fitted to the GRM and the IGRM. As we find in the stack loss data Chapter 4, 7^{th} and 8^{th} observations are the influential observations, which are not detected in the literature (see Table 4.4). These two observations are 3^{rd} ranked influential observations which are affecting the IGRM and IGRRM estimates.

2. Cook's distance (Cook, 1977) is the most popular influence diagnostic measure in the literature. On the comparison of influence diagnostics, we find in both GLM cases (GRM & IGRM), the influence detection by CD method is not better than the CVR and WD methods. We also find that the influence detection performance of CD method is same as the MCD and $DFFITS$ methods (see Tables 3.10-3.19 and Tables 4.11-4.20).

3. The simulation results of AP method for both the models along with ridge estimate indicate that the generated influential observation is detected successfully (see Tables 3.10-3.19 and Tables 4.11-4.20). While real data sets show that the AP detects the observations as influential observations which is actually not influencing the output of the GRM. As in the GRM with reaction rate data, this method detected observation 1^{st}, 9^{th}, 10^{th}, 13^{th} and 15^{th} are the influential observations. But on deleting these observations, we find no substantial changes in the GRM and the GRRM estimates (see Table 3.9). Similarly, for the IGRM with stack loss data we find the observations 9^{th} and 13^{th} are detected as influential observations which are actually uninfluential observations. Because on deleting, these observations have no change on the IGRM and IGRRM estimates (see Table 4.10).

4. Chatterjee and Hadi (2012) have shown that the CVR method is a good method for the detection of influential observations. We also find in both models with help of simulation and real data sets the similar conclusions for the detection of influential observations. Another advantage of the CVR method is that it is working well and identically diagnoses the similar influential observations with all forms of the GRM and IGRM residuals. There are rare chances to detect the wrong influential observation by the CVR method.

5. Hadi's influence diagnostic measure Eq. (3.46) page 74 is proposed for the GRM and IGRM for the detection of influential observations. We find with the help of simulation and a real data set that the influential observation detection

performance in both models seem to be very poor. This diagnostic measure needs further modification especially for these two models.

6. The deviance residuals with all GRM diagnostics perform better than that of using the Pearson residuals (See Tables 3.4 and 3.5). While deviance residual performance in the IGRM diagnostics is not better than the IGRM diagnostics using the Pearson residuals (see Tables 4.5 and 4.6).

7. The order of influential observation may change on applying the ridge regression alternative to the LM (Cook, 1977; Walker and Birch, 1988, Jahufer and Jianbao, 2009). Similar results are also seen in the influence diagnostics of the GRRM and the IGRRM (see Table 3.9 and 4.10).

8. Simulation results for the GRM residuals shows that influence measures performance with all forms of the GRM residuals is same except the likelihood residuals, when $\phi > 0.11$. However if $\phi \leqslant 0.11$, these measures gives different results with various forms of the GRM residuals. While, GRM and GRRM with working residuals give highest percentage of detecting potential influential observations as compared to the other GRM residuals for the detection of influential observations (see Table 3.14 and 3.19). Moreover, the influence diagnostics using likelihood residuals are found to be better with ridge estimates (see Table 3.12). We also find that the influence diagnostics of the GRM and the GRRM with standardized residuals are better than the adjusted residuals.

9. Now on comparing the IGRM residuals in detecting the influential observations, we find opposite results as the GRM residuals related to dispersion parameter. For $\phi \leqslant 2$, with all standardized and adjusted residuals the influence detection are identical except the IGRM likelihood residuals with all multicollinearity levels. While for $\phi > 2$, the performance of all IGRM and IGRRM residuals for the detection of influential observation is different. We observe that the influence detection with working residuals seems to be better than the other IGRM residuals (see Table 4.11-4.20). For this dispersion, we also find that the influence diagnostics of the IGRRM standardized (deviance, likelihood anscombe and working) residuals perform well than the adjusted forms of the IGRRM residuals. While with Pearson residuals, adjusted Pearson residuals perform better than standardized Pearson residuals but not better than IGRRM working residuals.

10. The results of real applications are also directed how much the results are important and the effect of influential observations on the GRM and IGRM estimates. Real applications (Reaction rate data and stack loss data) results indicate that in the GRM influence diagnostics, the likelihood residuals perform well. While in the IGRM influence diagnostics, the Pearson residuals perform well.

11. Our results suggest that the likelihood residuals are more appropriate for studying influence diagnostics in the GRM with ridge estimates. While for the IGRM, the most appropriate residuals are the Pearson residuals.

12. On the basis of simulation and real applications, we suggest that CVR is the appropriate diagnostic method, which performs better with all forms of the GLM residuals. This diagnostic is reported to perform better in dealing with the problem of multicollinearity.

5.0.0.1 Future Research

There are dimensions which still need to be explored:

1. This thesis covers the influence diagnostics with different GLM residuals along with one biased (ridge) estimation method. These can be extended to GLM influence diagnostics with modified ridge estimation, Liu estimation, modified Liu estimation, and Stein estimation methods.

2. As this thesis covers influence diagnostics in the gamma regression and inverse Gaussian regression using ridge estimator with different residuals. This dimension can be extended for the other GLM cases.

3. Beyaztas and Alin (2013) extended the LM influence diagnostics to the GLM with Jackknife-After-Bootstrap method considering the case of logistic regression (Beyaztas and Alin, 2014). These dimensions can be investigated for the other GLM like binomial, Poisson, gamma, inverse Gaussian etc. Also one can study these directions with biased estimators and compares the performance of these methods with different GLM residuals.

4. Pena (2005) proposed a new diagnostic measure to detect the influential observation in the LM. This method can be extended for the GLM with different residuals using all biased estimation techniques.

5. Emami and Emami (2016) proposed new influence diagnostic methods for the ridge regression model to detect the influential observations. This method can be extended for the GLM with different residuals using all biased estimation techniques.

6. The power of available influence diagnostic methods has not been explored so far in the GLM as well as in the LM with biased estimators. It can also be taken up and explored as a future project of study.

References

Agresti, A. (2015). *Foundations of Linear and Generalized Linear Models.* John Wiley & Sons, Inc.

Amin, M., Akbar, A., Manzoor, M.A. (2015). Fitting regression model with some assumptions for the identification of cotton yield factors. *Pakistan Journal of Life and Social Sciences*, 13(20), 86–90.

Amin, M., Ullah, M.A., Alsam, M. (2015). Empirical evaluation of the inverse gaussian regression residuals for the assessment of influential observations. *Journal of Chemometrics.*

Amin, M., Ullah, M.A., Alsam, M. (2015). Influence diagnostics in the gamma regression model using likelihood residuals. *Communication in Statistics-Theory and Methods.*

Amin, M., Ullah, M.A., Cordeiro, G.M. (2015). Influence diagnostics in the gamma regression model with adjusted deviance residuals. *Communication in Statistics-Simulation and Computation.*

Andrews, D.F., Pregibon, D. (1978). Finding the outliers that matter. *Journal of the Royal Statistical Society, Series B*, 40, 85–93.

Anholeto, T., Sandoval, M.C., Botter, D.A. (2014). Adjusted Pearson residuals in beta regression models. *Journal of Statistical Computation and Simulation*, 84(5), 999–1014.

Anscombe, F.J. (1953). Contribution of discussion paper by H. Hotelling New light on the correlation coefficient and its transformations. *Journal of the Royal Statistical Society, Series B*, 15, 229-230.

Aslam, A. (2014). Performance of Kibrias Method for the heteroscedastic ridge regression model: Some Monte Carlo evidence. *Communication in Statistics-Simulation and Computation*, 43(4), 1–14.

Atkinson, A.C. (1981). Two graphical displays for outlying and influential observations in regression. *Biomerika*, 68, 13–20.

Atkinson, A.C. (1982). Regression diagnostics, transformations and constructed variables. *Journal of the Royal Statistical Society, Series B*, 44, 1–36.

Atkinson, A.C. (1985). *Plots, Transformations and Regression*. Clarendon, Oxford.

Balasooriya, U., Tse, Y.K., Liew, Y.S. (1987). An empirical comparison of some statistics for identifying outliers and influential observations in linear regression models. *Journal of Applied Statistics*, 14(2), 177–184.

Balka, J., Desmond, A.F., McNicholas, P.D. (2011). Bayesian and likelihood inference for cure rates based on defective inverse Gaussian regression models. *Journal of Applied Statistics*, 38(1), 127–144.

Beirlant, J., Goegebeur, Y. (2003). Regression with response distributions of Pareto-type. *Computational Statistics and Data Analysis*, 42, 595–619.

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

Beyaztas, U., Alin, A. (2014). Jackknife-after-bootstrap as logistic regression diagnostic tool. *Communications in Statistics - Simulation and Computation*, 43(9), 2047–2060.

Beyaztas, U., Alin, A. (2013). Jackknife-after-bootstrap method for detection of influential observations in linear regression models. *Communication in Statistics Simulation and Computation*, 42, 1256–1267.

Billor, N. (1999). An application of the local influence approach to ridge regression. *Journal of Applied Statistics*, 26, 177–183.

Bonate, P.L. (2011). *Pharmacokinetic-Pharmacodynamic Modeling and Simulation, 2nd Edition*. Springer Science+Business Media, LLC.

Breslow, N. (1990). Test of hypotheses in overdispersion regression and other quasi likelihood models. *Journal of the American Statistical Association*, 85, 565–571.

Brown, G.C. (1997). Local influence regression diagnostics for gamma data. *Communication in Statistics-Theory and Methods*, 26(12), 3051–3067.

Brownlee, K.A. (1965). *Statistical Theory and Methodology in Science and Engineering*. John Wiley, New York.

Cameron, A.C., Trivedi, P.K. (1998). *Regression Analysis of Count Data*. Econometric Society Monograph No 30, Cambridge University Press, Cambridge.

Cancho, V.G. Ortega, E.M.M. and Bolfarine, H. (2009). The log-exponentiated-Weibull regression models with cure rate: Local influence and residual analysis. *Journal of Data Science*, 7, 433–458.

Chatterjee, S., Hadi, A.S. (2012). *Regression Analysis by Examples, 5th Edition*. John Wiley, New York.

Chatterjee, S., Hadi, A.S. (1988). *Sensitivity Analysis in Linear Regression*. Wiley, New York.

Chatterjee, S., Hadi, A.S. (1986). Influential observations, high leverage points, and outliers in regression. *Statistical Sciences*, 1, 379–416.

Chatterjee, S., Yilmaz, M. (1992). A review of regression diagnostics for behavioral research. *Applied Psychological Measurement*, 16(3), 209–227.

Chhikara, R.S., Folks, J.L., (1989). *The Inverse Gaussian Distribution; Theory, Methodology and Applications*. Marcel Dekker, New York.

Collett, D. (1991). *Modelling Binary Data*. Chapman & Hall.

Consul, P.C. and Famoye, F. (1992). Generalized Poisson regression model. *Communication in Statistics-Theory and Methods*, 21, 81–109.

Cook, R.D. (1977). Detection of inuential observations in linear regression. *Technometrics*, 19, 15–18.

Cook, R.D. (1986). Assessment of local influence (with discussion). *Journal of the Royal Statistical Society, Series B*, 48, 133–169.

Cook, R.D. (1998). *Regression Graphics: Ideas for studying regression through graphics*. John Wiley and Sons, New York.

Cook, R.D., Weisberg, S. (1982). *Residuals and Influence in Regression*. Chapman & Hall, New York.

Cordeiro, G.M. (2004). On Pearsons residuals in generalized linear models. *Statistics and Probability Letters*, 66, 213–219.

Cordeiro, G.M., Cysneiros, A.H.M.A., Cysneiros, F.J.A. (2006): Bartlett adjustments for over dispersed generalized linear models. *Communications in Statistics - Theory and Methods*, 35(5), 937–952.

Cordeiro, G. M., McCullagh, P. (1991). Bias correction in generalized linear models. *Journal of the Royal Statistical Society, Series B*, 53, 629–643.

Cox, D.R., Snell, E.J. (1968). A general definition of residuals (with discussion). *Journal Royal Statistical Society, Series B*, 30, 248–275.

Czado, C. and Raftery, A.E. (2006). Choosing the link function and accounting for link uncertainty in generalized linear models using Bayes factors.

Das, R.N. (2014). *Response Surfaces, Regression and Positive Data Analyses*. CRC Press Taylor & Francis Group.

Das, R.N., Kim, J. (2012). GLM and joint GML techniques in hydrogeology: An illustration. *International Journal of Hydrology Science and Technology*, 2(2), 185–201.

De Jong, P., Heller, G. (2008). *Generalized Linear Models for Insurance Data.* Cambridge University Press, New York.

Ducharme, G.R. (2001). Goodness-of-fit tests for the inverse Gaussian and related distributions. *Test*, 10(2), 271–290.

Dunteman, G.H., Ho, M.R. (2006). *An Introduction to Generalized Linear Models.* Sage Publication, New Dehli. Emami, H., Emami, M. (2016). New influence diagnostics in ridge regression. *Journal of Applied Statistics*, 43(3), 476–489.

Ertas, H., Erisoglu, M., Kaciranlar, S. (2013). Detecting influential observations in Liu and modified Liu estimators. *Journal of Applied Statistics*, 40(8), 1735–1745.

Espinheira, Ferrari, P.L.S.L.P., Cribari-Neto, F. (2008a). On beta regression residuals. *Journal of Applied Statistics*, 35,407–419.

Espinheira, Ferrari, P.L.S.L.P., Cribari-Neto, F. (2008b). Influence Diagnostics in beta regression. *Computational Statistics and Data Analysis*, 52, 4417–4431.

Fahrmeir, L., Tutz, G. (2001). *Multivariate Statistical Modelling based on Generalized Linear Models, 2nd Edition.* Springer-Verlag New York, Inc.

Famoye, F., Wulu, J.T., Singh, K.P. (2004). On the generalized poisson regression model with an application to accident data. *Journal of Data Science*, 2, 287–295.

Famoye, F. (1993). Restricted generalized Poisson regression model. *Communications in Statistics Theory and Methods*, 22, 1335–1354.

Ferrari, S.L.P., Uribe-Opazo, M.A., Cordeiro, G.M. (2002). Bartlett-type corrections for two-parameter exponential family models. *Communications in Statistics - Theory and Methods*, 31(6), 901–924.

Ferrari, S. L. P., Cribari-Neto, F. (2004). Beta regression for modeling rates and proportions. *Journal of Applied Statistics*, 31, 799–815.

Ferrari, S.L.P. (2011). Diagnostic tools in beta regression with varying dispersion. *Statistica Neerlandica*, 65(3), 337–351.

Firth, D. (1988). Multiplicative errors: Log-normal or Gamma. *Journal of the Royal Statistical Society, Series B*, 50(2), 266–268.

Fox, J. (2002). *An R and S-Plus Companion to Applied Regression.* Sage Publication Inc.

Fox, J. (1991). *Regression Diagnostics.* Sage Publication Inc.

Fuller, W.A. (1987). *Measurement Error Models.* Wiley, New York.

Galea-Rojas, M., Bolfarine, H., de Castro M. (2002). Local influence in comparative calibration models, *Biometrical Journal*, 1, 59–81.

Gbur, E.E., Stroup, W.W., McCarter, K.S., Durham, S., Young L.J., Christman, M., West, M., Kramer, M. (2012). *Analysis Generalized Linear Mixed Models in Agricultural and Natural Resources Sciences*. American Society of Agronomy, USA.

Gill, J. (2001). *Generalized Linear Models: A Unified Approach*. Sage Publication Inc.

Gray, G.B. (1989). On the use of regression diagnostics. *Statistician*, 38, 97–105.

Green, P.J. (1984). Iteratively reweighted least squares for maximum likelihood estimation, and some robust and resistant alternatives (with discussion). *Journal of the Royal Statistical Society, Series B*, 46, 149–192.

Gunst, R.F. (1983). Regression analysis with multicollinear predictor variables: definition, detection and effects. *Communications in Statistics, Theory and Models*, 12(19), 2217–2260.

Hadi, A.S. (1992). A new measure of over all potential influence in linear regression. *Computational Statistics and Data Analysis*, 14, 1–27.

Halekoh, U., Hjsgaard, S. (2007). Gamma distributed data. *Faculty of Agricultural Science*, University of Aarhus.

Hanagal, D.D., Dabade, A.D. (2013). A comparative study of shared frailty models for kidney infection data with generalized exponential baseline distribution. *Journal of Data Science*, 11,109–142.

Hardin, J.W., Hilbe, J.M. (2012). *Generalized Linear Models and Extensions.3rd Edition*. Stata Press Publication.

Hashimoto, E.M., Ortega, E.M.M, Cancho, V.G., Cordeiro, G.M. (2013). On estimation and diagnostics analysis in log-generalized gamma regression model for interval-censored data. *Statistics: A Journal of Theoretical and Applied Statistics*, 47(2), 379–398.

Hashimoto, E.M., Ortega, E.M.M., Cancho, V.G., Cordeiro, G.M. (2010). The log-exponentiated Weibull regression model for interval-censored data. *Computational Statistics and Data Analysis*, 54, 1017–1035.

Heinzl, H., Mittlbock, M. (2002). Adjusted measures for the inverse Gaussian regression model. *Computational Statistics*, 17, 525–544.

Hoerl, A.E., Kennard, R.W. (1970a). Ridge regression: Biased estimation for non-orthogonal problems. *Technometrics*, 12, 55–67.

Hoerl, A.E., Kennard, R.W. (1970b). Ridge regression: Application to non-orthogonal problems. *Technometrics*, 12, 69–82.

Hoaglin, D.C., R.E. Welsch (1978). The hat matrix in regression and ANOVA. *The American Statistician*, 32, 17–22.

Hossain, A., Naik D.N. (1991). A comparative on detection of influential observations in linear regression. *Statistical Papers*, 32, 55–69.

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

Huet, S., Bouvier, A., Gruet, M.A., Jolivet, E. (2004). *Statistical Tools for Nonlinear Regression: A Practival Guide with S-plus and R Examples.* Springer-Verlag, New York.

Iwase, K. (1989). Linear regression through the origin with constant coefficient of variation for the inverse Gaussian distribution. *Communications in Statistics-Theory and Methods*, 18(10), 3587-93.

Jahufer, A. (2014). Identifying unusual observations in ridge regression linear model using Box-Cox transformation technique. *Open Journal of Statistics*, 4, 19–26.

Jahufer, A., Jianbao, C. (2013). Detecting Global Influential Observations in Liu Regression Model. *Open Journal of Statistics*, 3(1), 5–11.

Jahufer, A., Jianbao, C. (2012). Identifying Local Influential Observations in Liu Estimator. *Metrika*, 75(3), 425–438.

Jahufer, A., Jianbao, C. (2011). Measuring Local Influential Observations in Modified Ridge Regression. *Journal of Data Science*, 9(3), 359–372.

Jahufer, A., Jianbao, C. (2009). Assessing global influential observations in modified ridge regression. *Statistics and Probability Letters*, 79, 513–518.

Jearkpaporn, D., Montgomery, D. C., Runger G. C., Borror C. M. (2005). Model-based process monitoring using robust generalized linear models. *International Journal of Production Research*, 43,7, 1337–1354.

Jiao, Y., Chen, Y. (2004). An application of generalized linear models in production model and sequential population analysis. *Fisheries Research*, 70, 367–376.

Jorgennsen, B. (1983). Maximum likelihood estimation and large-sample inference dor generalized linear and non-linear regression models. *Biometrika*, 70, 19–28.

Kaiser, M.S. (1997). Maximum likelihood estimation of link function parameters. *Computational Statistics and Data Analysis*, 24, 79–87.

Kamruzzaman, M., Imon, A.H.M.R. (2002). High leverage point: Another source of multicollinearity. *Pakistan Journal of Statistics*, 435–448.

Kelly, G.E. (1984). The influences function in the error in variables problem. *The Annals of Statistics*, 12, 87–100.

Kibria, B.M.G., Mansson, K., Shukur, G. (2012). Performance of some logistic regression estimators. *Computational Econometrics*, 40, 401–414.

Kim, C., Lee, J., Yang, H., Bae, W. (2015). Case influence diagnostics in lasso regression. *Journal of the Korean Statistical Society*, 44, 271–279.

Labra, F.V., Aoki, R., Rojas, F. (2007). An application of the local influence diagnostics to ridge regression under elliptical model. *Communications in Statistics -Theory and Methods*, 36(4), 767–779.

Landwehr, J. M., Pregibon, D. (1993). Comments on 'improved added variable and partial residual plots for the detection of influential observations in generalized linear models' by R. J. O'Hara Hines and E. M. Carter. *Applied Statistics*, 42, 16–19.

Lange, K.L., Little, R.J.A., Taylor, J.M.G. (1989). Robust statistical modeling using the t distribution. *Journal of the American Statistical Association*, 84, 881–896.

Lawless, J.F. (2003). *Statistical Models and Methods for Life Time Data*. Wiley, New York.

Lee, A. H. (1988). Assessing partial influence in generalized linear models. *Biometrics*, 44, 71–77.

Lee, A.H. (1987). Diagnostic displays for assessing leverage and influence in generalized linear models. *Australian Journal of Statistics*, 29(3), 233–243.

Lee, A.H., Zhao, Y. (1996). Assessing local influence in measurement error models. *Biometrical Journal*, 38, 829–841.

Lee, A.H., Silvapulle, M.J. (1988) Ridge estimation in logistic regression. *Communications in Statistics - Simulation and Computation*, 17, 1231–1257.

Lesaffre, E., Marx, B.D. (1993). Collinearity in generalized linear regression. *Communications in Statistics - Theory and Methods* 22, 1933–1952.

Li, B., Martin, E.B., Morris, A.J. (2001). A graphical technique for detecting influential cases in regression analysis. *Communications in Statistics - Theory and Methods*, 30(3), 463–483.

Lindsey, J.K. (2007). *Applying Generalized Linear Models*. Limburgs Universitair Centrum, Diepenbeek.

Lo, S.N., Ronchetti, E. (2009). Robust and accurate inference for generalized linear models. *Journal of Multivariate Analysis*, 100, 2126–2136.

Mackinnon, M.J., Puterman, M.L. (1989). Collinearity in generalized linear models. *Communication in Statistics- Theory and Methods*, 18(9), 3463–3472.

Markovic, R.D. (1965). Probability of best fit to distributions of annual precipitation and runoff (Fort Collins: Co: Colorado State University). *Hydrology Paper*, 8.

Marquardt, D.W. (1980). You should standardized the predictor variables in your regression models. *Journal of the American Statistical Association*, 75, 74–103.

Meloun, M., Militky, J. (2001). Tutorial: Detection of single influential points in OLS regression model building. *Analytica Chimica Acta*, 439,169–191

Mansson, K., Kibria, B.M.G., Sjolander, P., Shukur, G. (2012). Improved Liu estimators for the Poisson regression model. *International Journal of Statistics and Probability*, 1(1), 2–6.

Mason, R.L., Gunst, R.F. (1985). Outlier-Induced Collinearities. *Technometrics*, 27, 401–407.

McCullagh, P., Nelder, J.A. (1989). *Generalized linear models 2nd Edition*. Chapman & Hall, London.

McCullagh, P., Nelder, J. (1983). *Generalized Linear Models*. Chapman & Hall, London.

McCullagh, P. (1980). Regression Models for Ordinal Data. *Journal of Royal Statistical Society, Series B*, 42(2), 109–142.

Meshkani, M.R., Fallah, A., Kavousi, A. (2014). Analysis of covariance under inverse Gaussian model. *Journal of Applied Statistics*, 41(6), 1189–1202.

Minoda, Y., Yanagimoto, T. (2009). Estimation of a common slope in a gamma regression model with multiple strata: An empirical Bayes method. *Computational Statistics and Data Analysis*, 53(12), 4178–4185.

Myers, R.H., Montgomery, D.C., Vining, G.G., Robinson, T.J. (2010). *Generalized Linear Models With Applications in Engineering and the Sciences, 2nd Edition*. John Wiley & Sons, Inc., Hoboken, New Jersey.

Nelder, J.A. (1968). Weighted regression, quantal response data and inverse polynomials. *Biometrics*, 24, 979–985.

Nakamura, T. (1990). Corrected score function for errors-in-variables models: methodology and application to generalized linear models. *Biometrika*, 77, 127–137.

Nelder, J.A., Weddernburn, R.W.M. (1972). Generalized linear models. *Journal of the Royal Statistical Society, Series A*, 135, 370–384.

Nurunnabi, A.A.M., Hadi, A.S., Imon, A.H.M.R. (2014). Procedures for the identification of multiple influential observations in linear regression. *Journal of Applied Statistics*, 41(6), 1315–1331.

Olsson, U. (2002). *Generalized Linear Model: An Applied Approach*. Student literature AB.

Ortega, E.M.M., Cordeiro, G.M., Pascoa, M.A.R., Couto, E.V. (2012). The log-exponentiated generalized gamma regression model for censored data. *Journal of Statistical Computation and Simulation*, 82, 1169–1189.

Ortega, E.M.M., Cancho, V.G., Paula, G.A. (2009). Generalized log-gamma regression models with cure fraction. *Lifetime Data Analysis*, 15, 79–106.

Ortega, E.M.M., Cancho, V.G., Bolfarine, H. (2006). Influence diagnostics in exponentiated-Weibull regression models with censored data. *Statistics and Operations Research Transactions*, 30, 172–192.

Ortega, E. M. M., Bolfarine, H., Paula, G. A. (2003). Influence diagnostics in generalized log-gamma regression models. *Computational Statistics and Data Analysis*, 42, 165–186.

zkale, M.R. (2013) Influence measures in affine combination type regression. *Journal of Applied Statistics*, 40(10), 2219–2243.

Pena, D. (2005). A new statistics for influence in linear regression. *Technometrics*, 47,1–12.

Pierce, D.A., Schafer, D.W. (1986). Residuals in generalized linear models. *Journal of American Statistical Association*, 81, 977–986.

Preisser, J.S., Qaqish, B.F. (1996). Deletion diagnostics for generalized estimating equations. *Biometrika*, 83, 551–562.

Prudente, A.A., Cordeiro, G.M. (2010). Generalized Weibull linear models. *Communications in Statistics - Theory and Methods*, 39(20), 3739–3755.

Rasekh, A.R., Fieller, N.R.J. (2003). Influence functions in functional measurement error models with replicated data. *Statistics*, 37, 169–178.

Ratkowsky, D.A. (1983). *Nonlinear Regression Modeling*. Marcel Dekker, New York.

Ricci, L., Martinez, R. (2008). Adjusted R2-type measures for Tweedie models. *Computational Statistics and Data Analysis*, 52, 1650–1660.

Roozbeh, M., Arashi, M. (2014). Feasible ridge estimator in seemingly unrelated semiparametric models. *Communications in Statistics-Simulation and Computation*, 43(10), 2593–2613.

Schaefer, R.L. (1986). Alternative estimators in logistic regression when the data are collinear. *Journal Statistical Computation and Simulation*, 25, 75–91.

Schaefer, R.L., Roi, L.D., Wolfe, R.A. (1984). A ridge logistic estimator. *Communications in Statistics Theory and Methods*, 13, 99–113.

Schrodinger, V.E. (1915). Zur theorie der fall- und steigversuchea n teilchen mit brown-scher B ewegung. *PhysikalischeZ eitschri*, 16, 289–295.

Seber, G.A.F., Lee, A.J. (2003). *Linear Regression Analysis, 2nd Edition*. Wiley Series in Probability and Statistics.

Segerstedt, B. (1992). On ordinary ridge regression in generalized linear models. *Communications in Statistics - Theory and Methods*, 21(8), 2227–2246.

Segond, M-L., Onof, C., Wheater, H.S. (2006). Spatialtemporal disaggregation of daily rainfall from a generalized linear model. *Journal of Hydrology*, 331, 674–689.

Shi, X., Zhu, H., Ibrahim J.G. (2009). Local influence for generalized linear models with missing covariates. *Biometrics*, 65, 1–22.

Shi, L., Chen, G. (2009). Influence measures for general linear models with correlated errors. *The American Statistician*, 63, 40–42.

Simas, A.B., Cordeiro, G.M. (2009). Adjusted Pearson residuals in exponential family nonlinear models. *Journal of Statistical Computation and Simulation*, 79, 411–425.

Staniswalis, J.G. (2006). On fitting generalized non-linear models with varying coefficients. *Computers and Data Analysis*, 50, 1818–1839.

Stogiannis D. and Caroni C. (2013). Issues in fitting inverse Gaussian first hitting time regression models for lifetime data. *Communications in Statistics - Simulation and Computation*, 42(9), 1948–1960.

Thomas, W., Cook, R.D. (1989). Assessing influence on regression coefficients in generalized linear models. *Biometrika*, 76, 741–749.

Tsou, T-S. (2011). Determining the meanvariance relationship in generalized linear modelsA parametric robust way. *Journal of Statistical Planning and Inference*, 141, 197–203.

Tweedie, M.C.K. (1945). Inverse statistical variate. *Nature*, 155, 453.

Tweedie, M.C.K. (1957a). Statistical properties of inverse Gaussian distributionsI. *Annals of Mathematical Statistics*, 28, 362–377.

Tweedie, M.C.K. (1957b). Statistical properties of inverse Gaussian distributionsII. *Annals of Mathematical Statistics*, 28, 696–705.

Upadhyay, S.K., Agrawal R., Smith, A. F. M. (1996). Bayesian analysis of inverse Gaussian non-linear regression by simulation. *Sankhy: The Indian Journal of Statistics B*, 58(3), 363–378.

Ullah, M.A. Pasha, G.R., Aslam, M. (2013). Local influence diagnostics in the modified ridge regression. *Communications in Statistics - Theory and Methods*, 42(10), 1851–1869.

Ullah, M.A., Pasha, G.R. (2009). The origin and developments of inuence measures in regression. *Pakistan Journal of Statistics*, 25, 295–307.

Urbano, M.R., Demtrio, C.G.B., Cordeiro, G.M. (2012). On Wald residuals in generalized linear models. *Communications in Statistics - Theory and Methods*, 41(4), 741–758.

Vanegas, L.H., Rondn, L.M., Cordeiro, G.M. (2012). Diagnostic tools in generalized Weibull linear regression models. *Journal of Statistical Computation and Simulation*, 1, 1–24.

Velleman, P.F., Welsch, R.E. (1981). Efficient computing of regression diagnostics. *The American Statistician*, 35, 234–242.

Venezuela, M.K., Booter, D.A., Sandoval, M.C. (2007). Diagnostic techniques in generalized estimating equations. *Journal of Statistical Computation and Simulation*, 77, 879–888.

Villegas, C., Paula, G.A., Cysneiros, F.J.A., Galea, M. (2013). Influence diagnostics in generalized symmetric linear models. *Computational Statistics and Data Analysis*, 59, 161–170.

Walker, E., Birch, J.B. (1988). Influence measures in ridge regression. *Technometrics*, 30(2), 221–227.

Wang, M., Song, L., Wang, X. (2010). Bridge estimation for generalized linear models with a diverging number of parameters. *Statistics and Probability Letters*, 80, 1584–1596.

Weisberg, S. 2005. *Applied Linear Regression, 3rd Edition*. John Wiley & Sons, Inc.

Weissfeld, L.A., Sereika, S.M. (1991). A multicollinearity diagnostic for generalized linear models. *Communications in Statistics - Theory and Methods*, 20(4), 1183–1198.

Welsch, R.E. (1982). *Influence functions and regression diagnostics, in: R.* Launer, A. Siegel (Eds.), Modern Data Analysis, Academic Press, New York.

Welsch, R.E., Kuh, E. (1977). *Linear regression diagnostics*. Technical report 923–77, Sloan School of Management, Massachusetts Institute of Technology.

Wellman, J.M., Gunst, R.F. (1991). Influence diagnostics for linear measurement error models. *Biometrika*, 78, 373–380.

Whitmore, G. A. (1986). First-passage-time models for duration data: Regression structures and competing risks. *Statistician*, 35, 207–219.

Williams, D.A. (1984). Residuals in generalized linear models. *Proceedings of the 12th International Biometrics Conference*, Tokyo, 59–68.

Williams, D.A. (1987). Generalized linear model diagnostics using the deviance and single case deletions. *Journal of the Royal Statistical Society, Series C*, 36, 181–191.

Woldie, M., J.L. Folks, J.P. Chandler, 2001. Power function for inverse Gaussian regression models. *Communications in Statistics Theory and Methods*, 30, 787–797.

Woldie, M.; Folks, J.L. (1995). Calibration for inverse Gaussian regression models. *Communications in Statistics Theory and Methods*, 24, 2609–2620.

Wood, G.R. (2005). Confidence and prediction intervals for generalized linear accident models. *Accident Analysis and Prevention*, 37, 267–273.

Wu, K.Y.K. (2012). *On some extensions of generalized linear models with varying dispersion*. PhD theses, The University of Hong Kong.

Wu, X., Luo, Z. (1993). Second-order approach to local influence. *Journal of the Royal Statistical Society, Series B*, 55, 929–936.

Xiang, L., Tse, S., Lee, A.H. (2002). Influence diagnostics for generalized linear mixed models: applications to clustered data. *Computational Statistics and Data Analysis*, 40, 759–774.

Xie, F.C., Wei, B.C. (2007). Diagnostics Analysis in censored generalized Poisson regression models. *Journal of Statistics Computation and Simulation*, 77, 695–708.

Xie, F.C., Wei, B.C., Lin, J.G. (2008). Assessing influence for pharmaceutical data in zero-inflated generalized Poisson mixed models. *Statistics in Medicine*, 27, 3656–3673.

Xie, F.C., Wei, B.C. (2009). Diagnostics generalized Poisson regression models with error in variables. *Journal of Statistics Computation and Simulation*, 79, 909–922.

Xu, J., Lavalley, M. (2011). Pearson residual and efficiency of parameter estimates in generalized linear model. *Journal of Statistical Planning and Inference*, 141, 1014–1020.

Xu, L., Lee, S.Y., Poon, W.Y. (2006). Deletion measures for generalized mixed effects models. *Computational Statistics and Data Analysis*, 51, 1131–1146.

Zare, K., Rasekh, A. (2012). Residuals and leverages in the linear mixed measurement error models. *Journal of Statistical Computation and Simulation*, 1, 1–17.

Zhao, Y.J., Lee, A.H. (1995). Assessment of influence in nonlinear measurement error models. *Journal of Applied Statistics*, 22, 215–225.

Zhao, Y., Lee, A.H., Hui Y.V. (1994). Influence diagnostics for generalized linear measurement error models. *Biometrics*, 50, 1117–1128.

Zhong, X.P., Wei, B.C., Fung, W.K. (2000). Influence analysis for linear measurement error models. *Annals of the Institute of Statistical Mathematics*, 52, 367–379. 185.

Zhu, H.T., Lee, S.Y. (2001). Local influence for incomplete-data models. *Journal of Royal Statistical Society, Series B*, 63, 111–126.

Appendix

Appendix A

Data Sets

Table A.1: Reaction Rate Data (Huet et al., 2004)

Observation No.	Partial Pressure of Hydrogen	partial pressure of n-pentane	partial pressure of iso-pentane	Reaction Rate
1	205.8	90.9	37.1	3.541
2	404.8	92.9	36.3	2.397
3	209.7	174.9	49.4	6.694
4	401.6	187.2	44.9	4.722
5	224.9	92.7	116.3	0.593
6	402.6	102.2	128.9	0.268
7	212.7	186.9	134.4	2.797
8	406.2	192.6	134.9	2.451
9	133.3	140.8	87.6	3.196
10	470.9	144.2	86.9	2.021
11	300	68.3	81.7	0.896
12	301.6	214.6	101.7	5.084
13	297.3	142.2	10.5	5.686
14	314	146.7	157.1	1.193
15	305.7	142	86	2.648
16	300.1	143.7	90.2	3.303
17	305.4	141.1	87.4	3.054
18	305.2	141.5	87	3.302
19	300.1	83	66.4	1.271
20	106.6	209.6	33	11.648
21	417.2	83.9	32.9	2.002
22	251	294.4	41.5	9.604
23	250.3	148	14.7	7.754
24	145.1	291	50.2	11.59

Table A.2: Stack Loss Data (Brownlee, 1965)

Observation No.	Air. Flow	Water. Temperature	Acid Concentration	Stack Loss
1	80	27	89	42
2	80	27	88	37
3	75	25	90	37
4	62	24	87	28
5	62	22	87	18
6	62	23	87	18
7	62	24	93	19
8	62	24	93	20
9	58	23	87	15
10	58	18	80	14
11	58	18	89	14
12	58	17	88	13
13	58	18	82	11
14	58	19	93	12
15	50	18	89	8
16	50	18	86	7
17	50	19	72	8
18	50	19	79	8
19	50	20	80	9
20	56	20	82	15
21	70	20	91	15