A Robust Rescaled Moment Test for Normality in Regression

[1]Md.Sohel Rana, [1]Habshah Midi and [2]A.H.M. Rahmatullah Imon
[1]Laboratory of Applied and Computational Statistics, Institute for Mathematical Research,
University Putra Malaysia, 43400 Serdang, Selangor, Malaysia
[2]Department of Mathematical Sciences, Ball State University, Muncie, IN 47306, USA

Abstract: Problem statement: Most of the statistical procedures heavily depend on normality assumption of observations. In regression, we assumed that the random disturbances were normally distributed. Since the disturbances were unobserved, normality tests were done on regression residuals. But it is now evident that normality tests on residuals suffer from superimposed normality and often possess very poor power. **Approach:** This study showed that normality tests suffer huge set back in the presence of outliers. We proposed a new robust omnibus test based on rescaled moments and coefficients of skewness and kurtosis of residuals that we call robust rescaled moment test. **Results:** Numerical examples and Monte Carlo simulations showed that this proposed test performs better than the existing tests for normality in the presence of outliers. **Conclusion/Recommendation:** We recommend using our proposed omnibus test instead of the existing tests for checking the normality of the regression residuals.

Key words: Regression residuals, outlier, rescaled moments, skewness, kurtosis, jarque-bera test, robust rescaled moment test

INTRODUCTION

In regression analysis, it is a common practice over the years to use the Ordinary Least Squares (OLS) method mainly because of tradition and ease of computation. The assumption of normality, that is, the data are a random sample from a normal distribution is the most important assumption for many statistical procedures. Under a normal assumption of regression errors, the OLS method has many desirable properties in both estimation of parameters and in testing of hypotheses. But in practice we often deal with data sets which are not normal in nature. Non-normality may occur because of their inherent random structure or because of the presence of outliers. Nevertheless, evidence is available that such departures can have unfortunate effects in a variety of situations. In regression problems, the effect of departure from normality in estimation was studied by Huber[12]. In testing hypotheses, the effect of departure from normality has been investigated by many statisticians[15]. When the errors are not normally distributed, the estimated regression coefficients and estimated error variances are no longer normal and chi-square and consequently the t and F tests are not generally valid in finite samples. Koenker[16] pointed out that the powers of t and F tests are extremely sensitive to the hypothesized error distribution and may deteriorate very rapidly as the error distribution becomes long-tailed. Furthermore, Bera and Jarque[1] have found that homoscedasticity and serial independence tests suggested for normal errors may result in incorrect conclusions under non-normality. It may be also essential to have proper knowledge of the error distribution in prediction and in confidence limits of predictions. Most of the standard results of this particular study are based on the normality assumption and the whole inferential procedure may be subjected to error if there is a departure from this. In all, violation of the normality assumption may lead to the use of suboptimal estimators, invalid inferential statements and inaccurate predictions and for this reason test for normality has become an essential part of regression analysis.

Several methods have been suggested in the literature for assessing the assumption of normality. There are a considerable amount of written papers relating to the performance of various tests for normality in regression[5,8,9,13,18,20]. Among them the Jarque-Bera (JB) test[14] for normality (also known in statistics the Bowman-Shenton test) has become very popular with the statisticians. The JB test statistic is a

Corresponding Author: Sohel Rana, Laboratory of Applied and Computational Statistics, Institute for Mathematical Research,
University Putra Malaysia, 43400 Serdang, Selangor, Malaysia

sum of the sample coefficients of skewness and kurtosis and asymptotically follows a x^2 distribution with two degrees of freedom. But the main shortcoming of the JB test is that it possesses very poor power when the sample size is small or moderate[21]. To overcome this problem Imon[13] modified this test by suitably rescaling the moment estimators obtained from the least squares residuals. His proposed Rescaled Moment (RM) test improves the estimation of moments, coefficients of skewness and kurtosis and gives better power than the JB test for normality. However, both the JB and the RM test statistics rely on coefficients of sample skewness and kurtosis which are very sensitive to outliers. For this reason some robust tests of normality are being proposed by some authors[9,8]. Gel and Gastwirth[8] propose a modification of the JB test utilizing a robust estimate of spread, namely the average absolute deviation from the median (MAAD), in the denominators of skewness and kurtosis. This Robust Jarque-Bera (RJB) test statistic asymptotically follows a x^2 distribution[8] with two degrees of freedom and provides higher or similar power in detecting heavy-tailed alternatives compared to the JB. Since the RJB is designed as a general statistical test for normality, we suspect that it may not perform well in regression analysis. In this study we propose a robust rescaled moment test (RRM) for normality designed for regression models extending the idea of Imon[13] and Gel and Gastwirth[8]. The main objective of the proposed method is to suggest a normality test that is fairly robust in the presence of outliers and also performs best in general. The empirical evidence shows that the proposed RRM test offers substantial improvements over the existing tests of normality.

MATERIALS AND METHODS

The robust rescaled moments test: The simplest form of the general linear model is given by:

$$y_i = x_i^T \beta + \epsilon_i, \quad i = 1, 2, ..., n \tag{1}$$

Where:
y_i = The i-th observed response
x_i = A p×1 vector of predictors
β = A p×1 vector of unknown finite parameters
ϵ = Uncorrelated random errors with mean 0 and variance σ^2

Writing $Y = (y_1, y_2, ... y_n)^T$, $X = (x_1, x_2, ... x_n)^T$ and $\epsilon = (\epsilon_1, \epsilon_2, ... \epsilon_n)^T$ the model is:

$$Y = X \beta + \epsilon \tag{2}$$

where, $E(\epsilon) = 0$, $V(\epsilon) = \sigma^2 I$ and I is an identity matrix of order n.

Now we establish general expressions for the moments, coefficients of skewness and kurtosis of the true errors assuming only that the errors are uncorrelated, zero mean and identically distributed and the existence of their first four moments. Let us define the k-th moment about the origin of the i-th error by:

$$\mu_k = E(\epsilon_i^k), \; k = 1, 2, ... \tag{3}$$

As $E(\epsilon) = 0$, on the null model, we need not distinguish here between moments about the origin and those about the mean. Simple forms of the coefficients of skewness and kurtosis are given by:

$$\beta_1 = \frac{\mu_3}{\mu_2^{3/2}}, \qquad \beta_2 = \frac{\mu_4}{\mu_2^2} \tag{4}$$

In practice population moments are often estimated by sample moments. We generally define the k-th sample moment by:

$$m_k(\epsilon) = \frac{1}{n} \sum_i (\epsilon_i - \bar{\epsilon})^k, \quad k = 1, 2, ... \tag{5}$$

and hence the raw coefficients of skewness and kurtosis can be defined as:

$$S(\epsilon) = \frac{m_3(\epsilon)}{[m_2(\epsilon)]^{3/2}}, K(\epsilon) = \frac{m_4(\epsilon)}{[m_2(\epsilon)]^2} \tag{6}$$

Much work has been done in producing omnibus tests for normality, combining S and K in one test statistic. So a large deviation either of S from 0 or K from 3 should give a significant result, regardless of which one deviates from normal values. D'Agostino and Pearson[6] first suggested this kind of test. However, it requires an assumption of independence between S and K which is only asymptotically correct. Bowman and Shenton[2] suggested a normality test, popularly known as Jarque-Bera test, with the test statistic:

$$JB = \frac{S^2}{\sigma^2(S)} + \frac{(K-3)^2}{\sigma^2(K)} \tag{7}$$

Where:
$\sigma^2(S) = 6/n$ and $\sigma^2(K) = 24/n$ are the asymptotic variance of S and K respectively
S = Under normality asymptotically, is distributed as x^2 distribution with 2 degrees of freedom

Gel and Gastwirth[8] give a robust version of the Jarque-Bera test using a robust estimate of spread which is less influenced by outliers in the denominators of the sample estimates of skewness and kurtosis. They consider the average absolute deviation from the sample median (MAAD) proposed by Gastwirth[7] and is defined by:

$$J_n = \frac{A}{n}\sum_{i=1}^{n}|X_i - M| \tag{8}$$

where, $A = \sqrt{\pi/2}$. The robust sample estimates of skewness and kurtosis are $\hat{m}_3/J_n^3$ and $\hat{m}_4/J_n^4$, where $\hat{m}_3$ and $\hat{m}_4$ are the 3rd and 4th order of the estimated sample moments respectively, which lead the development of Gel and Gastwirth Robust Jarque-Bera (RJB) test statistic:

$$RJB = \frac{n}{B_1}\left(\frac{\hat{m}_3}{J_n^3}\right)^2 + \frac{n}{B_2}\left(\frac{\hat{m}_4}{J_n^4} - 3\right)^2 \tag{9}$$

To obtain the constants B_1 and B_2, we need to find the expressions for $E(J_n^3)$, $E(J_n^4)$, $E(J_n^6)$ and $E(J_n^8)$ for a finite sample size n, which can be calculated as suggested by Geary[10]. However, such calculations are quite tedious and are not of practical use since the convergence of estimators of kurtosis to the asymptotic normal distribution is very slow. We obtain B_1 and B_2 from the Monte Carlo simulation results given by Gel and Gastwirth[8]. In particular, if one desires to preserve the nominal level of 0.05, they recommend $B_1 = 6$ and $B_2 = 64$.

The above test procedures are designed in a fashion that we know the value of each observation and we can easily compute the value of a test statistic and can come to a conclusion. It is a common practice over the years to use the Ordinary Least Squares (OLS) residuals as substitutes of true errors but there is evidence that residuals have very different skewness and kurtosis than their corresponding true errors and hence normality tests on OLS residuals may perform poorly unless the test statistics are modified.

Let us assume that X is an n×p matrix of full column rank p, the OLS estimator of β is $\hat{\beta} = (X^TX)^{-1}X^TY$. The OLS residuals are given by:

$$\hat{\epsilon}_i = y_i - x_i^T\hat{\beta} \;,\; I = 1,2,...n \tag{10}$$

In matrix notation, the residual vector is:

$$\hat{\epsilon} = Y - X\hat{\beta} \tag{11}$$

The residual vector can also be expressed in terms of unobservable errors as:

$$\hat{\epsilon} = [I - X(X^TX)^{-1}X^T]\epsilon = H\epsilon \tag{12}$$

We shall follow Chatterjee and Hadi[3] by referring to $H = [I - X(X^TX)^{-1}X^T]$ as the residual hat matrix. The elements h_{ij} of this matrix will be termed as residual hat elements, which play a very important role in linear regression. The quantities $(1 - h_{ii})$ are often referred to as leverage values which measure how far the input vector x_i are from the rest of the data. Let the k-th moment of the i-th OLS residual be defined as:

$$\mu_k(\hat{\epsilon}_i) = E(\hat{\epsilon}_i^k) \;,\; k = 1,2,... \tag{13}$$

for i = 1,...,n, where we note that $E(\hat{\epsilon}_i) = 0$. Using Eq. 12, the i-th residual can be expressed in terms of true errors as:

$$\hat{\epsilon}_i = \sum_j h_{ij}\epsilon_j,\; i = 1,2,...,n \tag{14}$$

Hence the second, third and fourth order moments of OLS residuals can be expressed[11] as:

$$\mu_2(\hat{\epsilon}_i) = \mu_2 h_{ii} \tag{15}$$

$$\mu_3(\hat{\epsilon}_i) = \mu_3\sum_j h_{ij}^3 \tag{16}$$

$$\mu_4(\hat{\epsilon}_i) = [\mu_4 - 3\mu_2^2]\sum_j h_{ij}^4 + 3\mu_2^2 h_{ii}^2 \tag{17}$$

The coefficients of skewness and kurtosis of $\hat{\epsilon}_i$ are:

$$\beta_1(\hat{\epsilon}_i) = \beta_1\left\{\frac{\sum_j h_{ij}^3}{h_{ii}^{3/2}}\right\}, \beta_2(\hat{\epsilon}_i) = 3 + (\beta_2 - 3)\left\{\frac{\sum_j h_{ij}^4}{h_{ii}^2}\right\} \tag{18}$$

In the case of OLS residuals, their sample mean is zero and therefore, the k-th sample moment is defined as:

$$m_k(\hat{\epsilon}) = \frac{1}{n}\sum_i \hat{\epsilon}_i^k, \quad k = 1,2,... \tag{19}$$

The sample coefficients of skewness and kurtosis on OLS residuals are then directly obtained from (5) by

substituting $\in$ by $\hat{\in}$ in it and thus coefficients of skewness and kurtosis based on OLS residuals can be defined as:

$$S(\hat{\in}) = \frac{m_3(\hat{\in})}{[m_2(\hat{\in})]^{3/2}}, K(\hat{\in}) = \frac{m_4(\hat{\in})}{[m_2(\hat{\in})]^2} \qquad (20)$$

Imon[13] shows that the sample moments as defined in (19) are biased and he also suggests unbiased estimates of the second, third and fourth order moments defined as:

$$m_2^*(\hat{\in}) = \frac{n}{\sum_i \sum_j h_{ij}^2} m_2(\hat{\in}) = \frac{n}{\sum_i h_{ii}^2} m_2(\hat{\in}) \qquad (21)$$

$$m_3^*(\hat{\in}) = \frac{n}{\sum_i \sum_j h_{ij}^3} m_3(\hat{\in}) \qquad (22)$$

$$m_4^*(\hat{\in}) = \frac{n}{\sum_i \sum_j h_{ij}^4} m_4(\hat{\in}) - 3\left(\frac{n}{n-p}\right)^2$$
$$\left\{\frac{\sum_i h_{ii}^2}{\sum_i \sum_j h_{ij}^4} - 1\right\}\{m_2(\hat{\in})\}^2 \qquad (23)$$

Imon[13] also suggests some simple approximations for the scaling factors which avoid the necessity of additional computation with the h_{ij}'s after the regression. He shows that an expression such as $\sum_j h_{ij}^k$, where k>1, is often dominated by the term h_{ii}^k. When also summing over i, the replacement of $\sum_j h_{ij}^k$ by $\left(\frac{n-p}{n}\right)^k$ the k-th power of the average value of h_{ij} may be considered as possible approximations. Thus we might approximate:

$$\sum_i h_{ii}^2 \approx \frac{(n-p)^2}{n} \qquad (24)$$

and

$$\sum_i \sum_j h_{ij}^k \approx \frac{(n-p)^k}{n^{k-1}} \qquad (25)$$

Substituting the values (24) and (25) in (21-23), Imon[13] suggests Rescaled Moments (RM) of OLS residuals as:

$$m_2^*(\hat{\in}) = cm_2(\hat{\in}) \qquad (26)$$

$$m_3^*(\hat{\in}) \approx c^3 m_3(\hat{\in}) \qquad (27)$$

$$m_4^*(\hat{\in}) \approx c^4\left[m_4(\hat{\in}) - 3\left(1 - c^{-2}\right)\{m_2(\hat{\in})\}^2\right] \qquad (28)$$

where, c = n/(n-p). Using the above approximations, the rescaled coefficients of skewness and kurtosis become:

$$S^*(\hat{\in}) = c^{3/2} S(\hat{\in}), K^*(\hat{\in}) = c^2[K(\hat{\in}) - 3] \qquad (29)$$

and the Rescaled Moment (RM) normality test statistic is defined as:

$$RM = \left[nc^3/6\right]\left[S^2 + c(K-3)^2/4\right] \qquad (30)$$

Likewise the JB statistic, the RM statistic follows a chi-square distribution with 2 degrees of freedom. Now using the average absolute deviation from the sample median (MAAD) as a robust estimate of spread we can define the Robust Rescaled Moment (RRM) test statistic as:

$$RRM = \frac{nc^3}{B_1}\left(\frac{\hat{m}_3}{J_n^3}\right)^2 + \frac{nc^4}{B_2}\left(\frac{\hat{m}_4}{J_n^4} - 3\right)^2 \qquad (31)$$

Under the null hypothesis of normality, the RRM test statistic asymptotically follows a chi-square distribution with 2 degrees of freedom. B_1 and B_2 are computed similar to the RJB test statistic in Eq. 9 as suggested by Geary[10].

RESULTS

Numerical examples: We consider few real life data sets for testing normality assumption in the presence of outliers.

Belgian road accident data: Our first example presents the number of road accidents in Belgium recorded between 1975 and 1981 taken from Rousseeuw and Leroy[19] and shown in Table 1. It has been reported by many authors that the Belgian road accident data contains a single outlier (record of 1979) which must cause nonnormality of residuals. Table 2 exemplifies the power of normality tests of this data.

Shelf-stocking data: Next we consider the shelf-stocking data given by Montgomery *et al.*[17].

Table 1: Belgian road accident data

Year	No. of accident
1975	18031
1976	18273
1977	16660
1978	16660
1979	23805
1980	15241
1981	13295

Table 2: Power of normality tests for Belgian road accident data

Tests	Value of statistic	p-value
JB	4.379730	0.112
RM	12.83904	0.002
RJB	54.50434	0.000
RRM	177.7253	0.000

Table 3: Shelf- stocking data (original and modified)

Time (min)	Cases stocked	Time (min)	Cases stocked	Time (min)	Cases stocked
10.15	25	5.06	13	7.57	19
2.96	6	9.14	23	1.74	4
3.00	8	11.86	30	9.38	24
6.88	17	11.69	28	0.16 (1.70)	1
0.28	2	6.04	14	1.84	5

Table 4: Power of normality tests for original data modified Shelf-stocking data

	Original data		Modified data	
Tests	Value of statistic	p-value	Value of statistic	p-value
JB	1.2643	0.5314	2.1820	0.3359
RM	1.9700	0.3735	3.4524	0.1779
RJB	1.4632	0.4811	5.0890	0.0785
RRM	2.2477	0.3250	8.2475	0.0161

This data presents the time required for a merchandiser to stock a grocery store shelf with a soft drink product as well as the number of cases of product stocked. We deliberately change one data point (putted in parenthesis) to get an outlier and both the original and modified data are shown in Table 3. Table 4 shows the power of normality test of this original and modified data.

Power simulations: We carry out a simulation experiment to compare the performance of the newly proposed RRM test with the existing tests of normality, in particular, with the Jarque-Bera, the Rescaled Moment (RM) and the Robust Jarque-Bera (RJB) tests. The estimated powers of these tests under various distributions for different sample sizes are shown in Table 5-7. We consider simulated power for ten different distributions; the normal, the exponential, the t-distribution with 3 and 5 degree of freedom, the logistic, the Cauchy, the exponential, the log-normal and the contaminated normal distributions with shifts in location,

Table 5: Simulated power of different tests for normality for n = 20

Distribution	JB	RM	RJB	RRM
Normal	0.0256	0.0425	0.0619	0.0892
t_3	0.3126	0.3720	0.4400	0.4892
t_5	0.1629	0.2115	0.2570	0.3106
Logistic	0.0904	0.1305	0.1721	0.2148
Cauchy	0.8193	0.8532	0.9122	0.9288
Exponential	0.4842	0.5855	0.6019	0.6609
Log-normal	0.7233	0.7975	0.8109	0.8458
Cont. normal (scale shift)	0.3243	0.3740	0.4063	0.4473
Cont. normal (location shift)	0.8065	0.8850	0.9292	0.9552
Cont. normal (outlier)	0.7958	0.9151	0.9616	0.9805

Table 6: Simulated power of different tests for normality for n = 50

Distribution	JB	RM	RJB	RRM
Normal	0.0385	0.0497	0.0637	0.0754
t_3	0.6674	0.6893	0.7475	0.7638
t_5	0.3999	0.4286	0.4897	0.5143
Logistic	0.2197	0.2478	0.2992	0.3266
Cauchy	0.9932	0.9947	0.9985	0.9987
Exponential	0.9517	0.9706	0.9463	0.9590
Log-normal	0.9958	0.9976	0.9953	0.9968
Cont. normal (scale shift)	0.6487	0.6696	0.6919	0.7091
Cont. normal (location shift)	0.9977	0.9984	0.9995	0.9998
Cont. normal (outlier)	0.9989	0.9995	0.9999	0.9999

Table 7: Simulated power of different tests for normality for n = 100

Distribution	JB	RM	RJB	RRM
Normal	0.0395	0.0442	0.0540	0.0601
t_3	0.8988	0.9050	0.9353	0.9404
t_5	0.6308	0.6504	0.6959	0.7076
Logistic	0.3770	0.3934	0.4464	0.4613
Cauchy	1.0000	1.0000	1.0000	1.0000
Exponential	1.0000	1.0000	0.9999	1.0000
Log-normal	1.0000	1.0000	1.0000	1.0000
Cont. normal (scale shift)	0.8652	0.8712	0.8792	0.8845
Cont. normal (location shift)	0.5536	0.5702	0.5851	0.5976
Cont. normal (outlier)	1.0000	1.0000	1.0000	1.0000

scale and containing outliers. For contaminating normal distributions, 90% observations are generated from standard normal distribution and the remaining 10% observations come from a normal with either shifted mean or variance or containing outliers. All our results are given at the 5% level of significance and are based on 10,000 simulations.

DISCUSSION

Here we discuss the results that we have obtained in previous section using real data sets and Monte Carlo simulation experiments.

Figure 1 shows us a genuine picture that the residuals for this data do not follow normality pattern. We compute the coefficients of skewness (1.764) and kurtosis (4.601) for this data which are also far from normality.
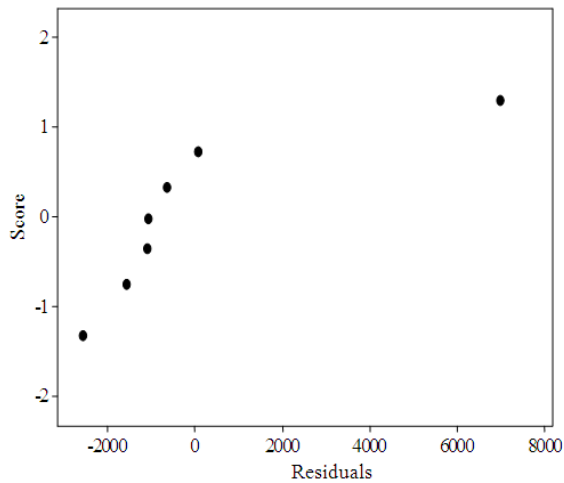
Fig. 1: Normal probability plot of the residuals of the Belgian road accident data

But it is interesting to observe from the results shown in Table 2 that the classical Jarque-Bera test fails to detect the nonnormality pattern of residuals for this data even at the 10% level of significance. The rescaled moments tests as suggested by Imon[13] can successfully detect nonnormality at the 5% level. Both the robust RJB and RRM rejects normality in a highly significant way, but the RRM possesses higher power than the RJB for this data.

The normal probability plots of the original and modified shelf-stocking data are shown in Fig. 2. It is not clear from these plots whether the residuals are normally distributed or not. We apply the classical and robust methods to check the normality and the results are shown in Table 4.

For the original data all the methods show that the residuals for this data are normally distributed. Since we have inserted an outlier in the modified data it is expected that the normality pattern of residuals will be affected. It is worth mentioning that all commonly used outlier detection techniques such as the Least Median of Squares (LMS), the Least Trimmed Squares (LTS), the Block Adaptive Computationally-Efficient Outlier Nominator (BACON)[4] can easily identify one observation as an outlier. The standard theory tells us that the normality should break down in the presence of outliers. But it is interesting to observe that both the Jarque-Bera and the RM test fail to detect non-normality here. The RJB test also fails to detect non-normality at the 5% level of significance. But the performance of the RRM test is quite satisfactory in this occasion. It can detect the problem of non-normality even at 1.6% level of significance.
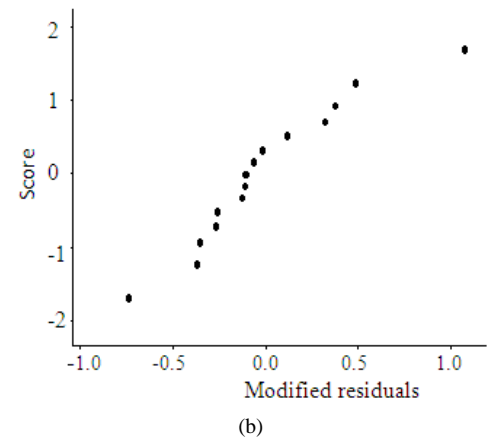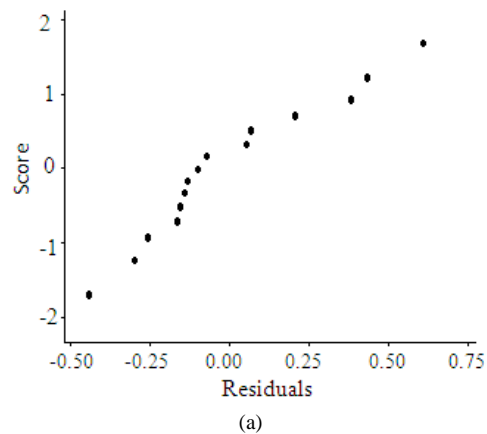

(a)

(b)

Fig. 2: Normal probability plot of the residuals of (a): The original and (b): The modified shelf-stocking data

Simulation results shown in Table 5-7 show that when the data come from normal distributions the performance of classical normality tests are good. Both the RJB and the RRM show a slightly higher size but the values are not that big and tend to decrease with the increase in sample size. But the classical tests perform poorly when the errors come from either heavier tailed or contaminated normal distributions. But the newly proposed RRM test performed best throughout and hence may be considered as the most powerful test for normality.

Figure 3 shows the merit of using the RRM test for normality. We observe that this test possesses much higher power than the Jarque-Bera, the RM and the robust Jarque-Bera tests under a variety of error distributions.
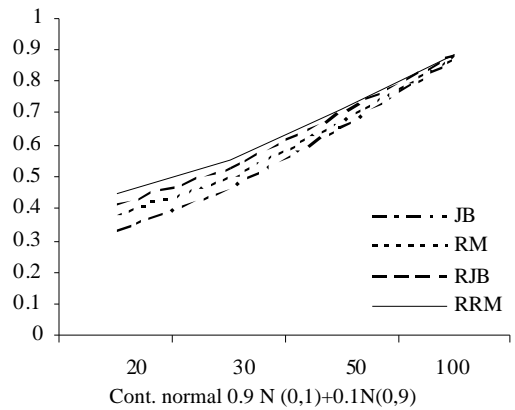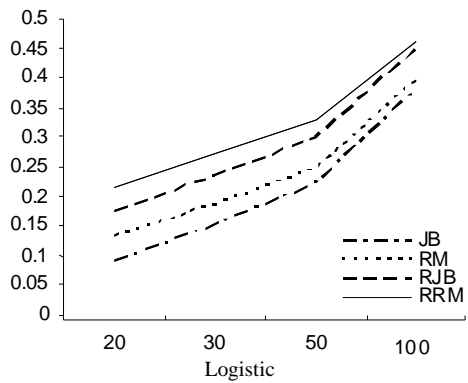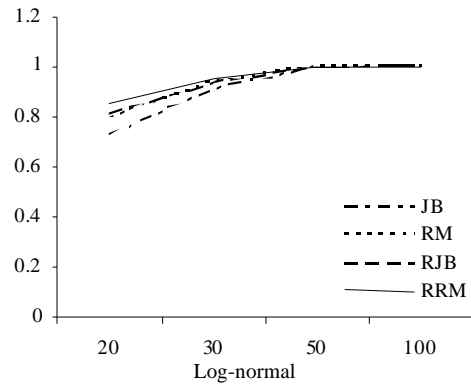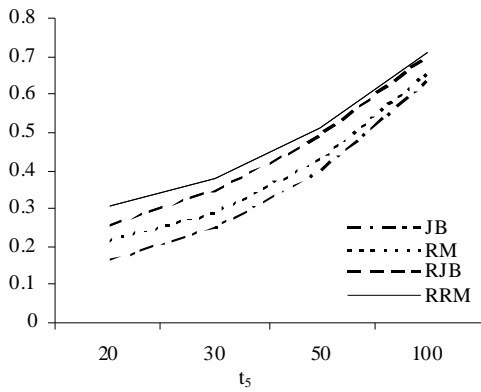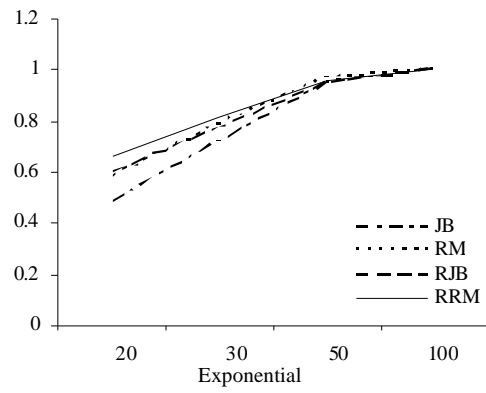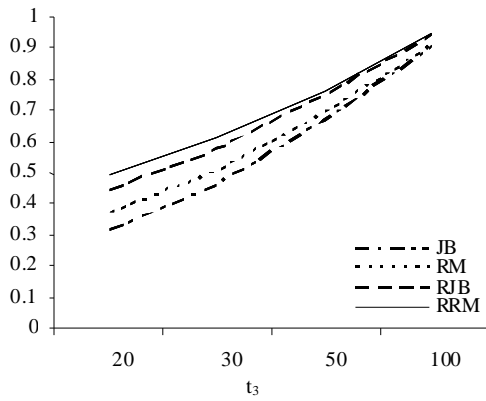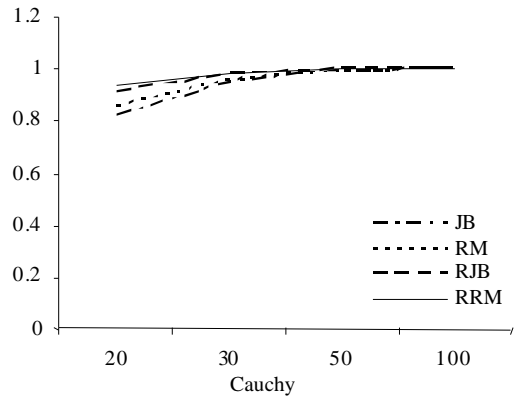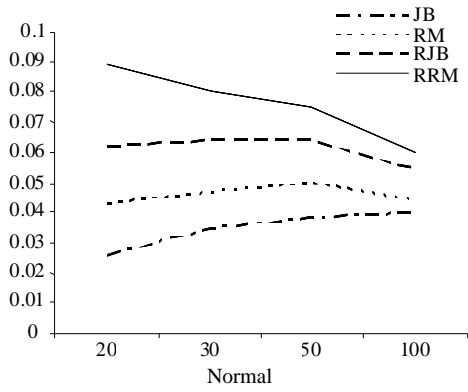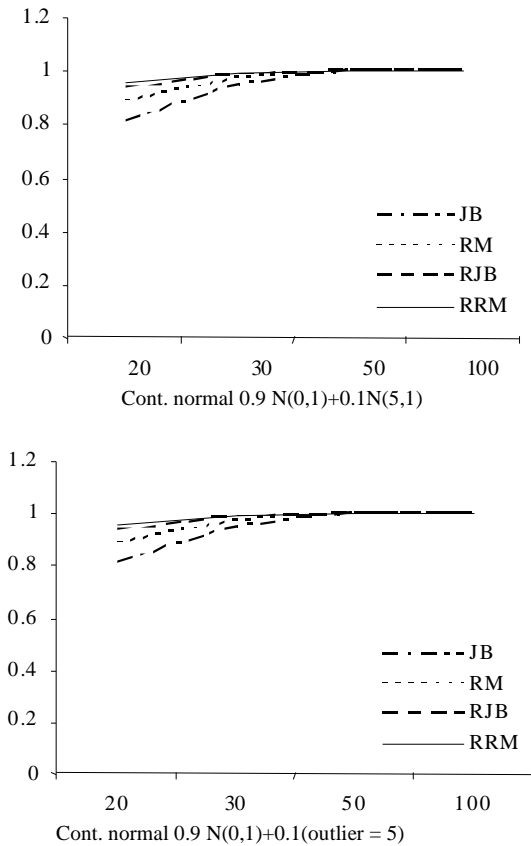
Fig. 3: Power of the normality tests for different distributions

CONCLUSION

In this study we develop a new test for normality in the presence of outliers, called the RRM, which is a robust modification of the rescaled moments test for normality. The real data sets and Monte Carlo simulation shows that the modified rescaled moment test offers substantial improvements over the existing tests and performs superbly to test the normality assumption of regression residuals.

REFERENCES

1. Bera, A.K. and C.M. Jarque, 1982. Model specification tests: A simultaneous approach. J. Econ., 20: 59-82. DOI: 10.1016/0304-4076(82)90103-8
2. Bowman, K. and L. Shenton, 1975. Omnibus test contours for departures from normality based on $\sqrt{b_1}$ and b_2. Biometrika, 64: 243-250. DOI: 10.1093/biomet/62.2.243
3. Chatterjee, S. and A.S. Hadi, 1988. Sensitivity Analysis in Linear Regression. 1st Edn., Wiley, New York, ISBN: 9780471822165, pp: 336.
4. Chatterjee, S. and A.S. Hadi, 2006. Regression Analysis by Examples. 4th Edn., Wiley, New York, ISBN: 9780471746966, pp: 375.
5. D'Agostino, R. and M. Stephens, 1986. Goodness-of-Fit Techniques. 1st Edn., Marcel Dekker, New York, ISBN: 9780824774875, pp: 576.
6. D'Agostino, R. and M. Pearson, 1973. Testing for departures from normality. Empirical results for distribution of b_2 and $\sqrt{b_1}$. Biometrika, 60: 613-622. DOI: 10.1093/biomet/60.3.613
7. Gastwirth, J.L., 1982. Statistical properties of a measure of tax assessment uniformity. J. Stat. Plan. Inform., 6: 1-12. DOI: 10.1016/0378-3758(82)90050-7
8. Gel, Y.R. and J.L. Gastwirth, 2007. A robust modification of the Jarque-Bera test of normality. Econ. Lett., 99: 30-32. DOI: 10.1016/j.econlet.2007.05.022
9. Gel, Y.R., W. Miro and J.L. Gastwirth, 2006. Robust direct test of normality against heavy-tailed alternatives. Commun. Stat. Data Anal., 51: 2734-2746. DOI: 10.1016/j.csda.2006.08.022
10. Geary, R.C., 1936. Moments of the ratio of the mean deviation to the standard deviation for normal sample. Biometrika, 28: 295-307. DOI: 10.1093/biomet/28.3-4.295
11. Huang, C.F. and B.W. Bolch, 1974. On the testing of regression disturbances for normality. J. Am. Stat. Assoc., 69: 330-335.
12. Huber, P.J., 1973. Robust regression: Asymptotics, conjectures and Monte Carlo. Ann. Stat., 1: 799-821. DOI: 10.1214/aos/1176342503
13. Imon, A.H.M.R., 2003. Regression residuals, moments and their use in tests for normality. Commun. Stat. Theor. Methods, 32: 1021-1034. DOI: 10.1081/sta-120019960
14. Jarque, C.M. and A.K. Bera, 1980. Efficient tests for normality, homoscedasticity and serial dependence of regression residuals. Econ. Lett., 6: 255-259. DOI: 10.1016/0165-1765(80)90024-5
15. Judge, G.G., W.E. Griffith, R.C. Hill, H. Lutkepohl and T. Lee, 1985. Theory and Practice of Econometrics. 2nd Edn., Wiley, New York, ISBN: 9780471895305, pp: 1019
16. Koenker, R.W., 1982. Robust methods in econometrics. Econ. Rev., 1: 213-90. DOI: 10.1080/07311768208800017
17. Montgomery, D.C., E.A. Peck and G.G. Vining, 2001. Introduction to Linear Regression Analysis. 3rd Edn., Wiley, New York, ISBN: 0-471-31565-6, pp: 641.

18. Pierce, D.A. and R.J. Gray, 1982. Testing normality of errors in regression models. Biometrika, 69: 233-236. DOI: 10.1093/biomet/69.1.233

19. Rousseeuw, P.J. and A. Leroy, 1987. Robust Regression and Outlier Detection. 1st Edn., Wiley, New York, ISBN: 0471852333, pp: 329.

20. Thadewald, T. and Buning, H., 2007. Jarque-Bera test and its competitors for testing a normality-power comparison. J. Applied Stat., 34: 87-105. DOI: 10.1080/02664760600994539

21. Wetherill, G.B., 1986. Regression Analysis with Application. 1st Edn., Chapman and Hall, London, ISBN: 9780412274909, pp: 311.