

Power of t -test for Simple Linear Regression Model with Non-normal Error Distribution: A Quantile Function Distribution Approach

S. Jahan^{1*} and A. Khan²

¹Department of Arts and Sciences, Ahsanullah University of Science and Technology, Dhaka-1208, Bangladesh

²Department of Applied Statistics, Institute of Statistical Research and Training (ISRT), University of Dhaka, Dhaka-1000, Bangladesh

Received 9 March 2012, accepted in final revised form 11 August 2012

Abstract

One of the major assumptions of the regression analysis is the normality assumption of the model error. We generally assume that the error term of the simple linear regression model is normally distributed. But in this paper g -and- k distribution is used as the underlying assumption for the distribution of error in simple linear regression model and a numerical study is conducted to see what extent of the deviation from normality causes what extent of effect on the size and power of t -test for simple linear regression model with the deviation being measured by a set of skewness and kurtosis parameters. The strength of t -test is evaluated by observing the power function of t -test. The simulation result shows that, the performance of the t -test for simple linear regression model with g -and- k error distribution is seen to be vastly affected in presence of excess kurtosis and small samples (i.e. $n < 100$). t -test is size robust under normal situation. Skewness and kurtosis parameter has a very little effect on the size of the t -test.

Keywords: g -and- k distribution; Robustness; Skewness; Kurtosis.

© 2012 JSR Publications. ISSN: 2070-0237 (Print); 2070-0245 (Online). All rights reserved.

doi: <http://dx.doi.org/10.3329/jsr.v4i3.9067>

J. Sci. Res. 4 (3), 609-622(2012)

1. Introduction

Most of the statistical procedures such as t -test, tests for regression coefficients, analysis of variance, and the F -test of homogeneity of variance have a fundamental assumption that the sampled data come from a normal distribution. The assumption of normality in a statistical procedure requires an effective test of whether the assumption holds, or a vigilant argument showing that violation of assumption does not invalidate the procedure used. Much statistical research has been concerned with evaluating the effect of the violations of assumption on the true significance level of a test or the efficiency of the

* Corresponding author: samsad.jahan@gmail.com

parameter estimates. In this manuscript, we focus on the power of regression tests by considering the error term of the simple linear regression model to follow the g -and- k distribution and observe the extent of the effect of deviation from normality on the size and power of regression tests for simple linear regression model with the deviation being measured by a set of skewness and kurtosis parameter.

There have been a wide range of studies on the influences of non-normality on several statistical test procedures during the last several decades. Swailowsky and Blair [1] evaluated the Type I and II error properties of the t -test by means of a Monte Carlo Study. Chen [2] investigated the upper-tailed test for the mean of the positively skewed distribution. He proposed a modified t -test. A Monte Carlo study investigated the new procedures for a variety of positively skewed distributions with small sample sizes and he found that the new test procedure is more accurate and more powerful than both Johnson modified t -test and Sutton's composite test. He also suggested that the new procedure is good choice for use as an upper tailed test of the mean of a positively skewed distribution with a slightly large or large skewness when the sample size is small. From the Monte Carlo studies he also found that the right tail of the sampling distribution of the new procedure is somewhat robust. Carolan and Rayner [3] derived a score test of location from a distorted normal distribution and compared the performance of the test to t -test and Wilcoxon test for symmetrical data from such a distribution by a simulation study. Carolan and Rayner [4] showed that standard tests of location such as t -test and Wilcoxon test, Which test for mean and median respectively, can perform poorly as tests when the data is other than unimodal and symmetric. Khan and Rayner [5] studied the effect of deviating from normal distribution assumption when considering the power of two many-sample location test procedures: ANOVA (parametric) and Kruskal-Wallis (non-parametric). Rasch and Gulard [6] performed a systematic research of the robustness of statistical procedures against non-normality. Yanagihara [7] presented the conditions for robustness to the non-normality on three test statistics for a general multivariate linear hypothesis, which were proposed under the normal assumption in a generalized multivariate analysis of variance (GMANOVA) model. Borhan and Khan [8] have studied about the sensitivity to non-normality of one sample t -test, Wilcoxon test and Z -test. Schmider *et al.* [9] carried out the empirical evidence to the robustness of the analysis of variance (ANOVA) concerning violation of the normality by means of Monte Carlo methods. Uddin *et al.* [10] examined the non-normality in the form of skewness and kurtosis in lot acceptance quality characteristics data. Their investigation showed that high skewness and kurtosis were associated with higher lot variability. This variability produced misleading results in regard to inflated Type-I error and low power of F -test. They also found the t -test to be quite robust for distinguishing mean differences. Very recently, Landau and stahl [11] showed how Monte Carlo simulation can be used for sample size, power or precision calculations when planning medical research studies.

It is clear that measuring power using g -and- k distribution has received great interest in recent years. But so far, a little study has been attempted to investigate the extent of effect of non-normality on the size and power of various regression tests for simple linear

regression model. Hence, in this paper the extent of effect of non-normality on the power is investigated by varying the skewness and kurtosis parameter of g -and- k distribution and measured numerically.

2. The g -and- k Distribution

Rayner and MacGillivray [13] examined the effect of non-normality on the distribution of (numerical) maximum likelihood estimators. The g -and- k distribution [12] can be defined in terms of its quantile function as:

$$Q_X(u \mid A, B, g, k) = A + Bz_u \left(1 + c \frac{1 - e^{-gz_u}}{1 + e^{-gz_u}} \right) \left(1 + z_u^2 \right)^k, \quad (1)$$

where A and $B > 0$ are location and scale parameters respectively, g measures skewness in the distribution, $k > -1/2$ measures kurtosis (in general sense of peakness/tailedness) in the distribution and $z_u = \phi^{-1}(u)$ is the u th quantile of a standard normal variate, and c is a constant chosen to help produce proper distributions. It can be clearly observed that for $g = k = 0$, the quantile function in (1) is just the quantile function of a standard normal variate.

The sign of the skewness parameter indicates the direction of skewness; $g < 0$ indicates the distribution is skewed to the left, and $g > 0$ indicates skewed to the right. Increasing/decreasing the unsigned value of increases/decreases the skewness in the indicated direction. When $g = 0$ the distribution is symmetric. The kurtosis parameter k , for the g -and- k distribution, behaves similarly.

Increasing k increases the level of kurtosis and vice versa. The value $k = 0$ corresponds to no extra kurtosis added to the standard normal base distribution. However, this distribution can represent less kurtosis than the normal distribution, as $k > -1/2$ can negative values. If curves with more kurtosis required then base distribution with less kurtosis than standardized normal distribution can be used.

For these distributions c is the value of 'overall symmetry' (MacGillivray). For an arbitrary distribution, theoretically the overall asymmetry can be as large as one, so it would appear that for $c < 1$, data or distribution could occur with skewness that cannot be matched by these distributions. However for $g \neq 0$, the larger the value chosen for c , the more restrictions on k are required to produce a completely proper distribution. Real data seldom produce overall asymmetry values greater than 0.8. We have used $c = 0.83$ throughout this paper. To examine extent of the effect of different level of non-normality on the test of conventional linear regression model we have considered that the random error belongs to the g -and- k distribution.

3. Robustness

Robustness occurs when the nominal and actual test sizes are not drastically different under slight model failure. When the validity of a certain test result is not extremely

affected by poorly structured data, then the test is considered robust. So the technical definition of robustness is: if the actual Type I error of a test is close to the proclaimed Type I rate, say 0.05, then the test is said to be robust. In other words it is resistant against the violation of assumptions.

4. Regression Model

We have considered a basic regression model where there is only one independent variable and the regression function is linear. The model [14] can be stated as follows:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \quad (2)$$

where Y_i is the value of the response variable in the i th trial, β_0 and β_1 are intercept and slope parameter, X_i is the value of the independent variable in the i th trial, ε_i is a random error term with mean $E(\varepsilon_i) = 0$ and variance $V(\varepsilon_i) = \sigma^2$, ε_i and ε_j are uncorrelated so that the covariance, $V(\varepsilon_i, \varepsilon_j) = 0$ for all $i, j; i \neq j; i = 1, 2, \dots, n$.

Model (2) is said to be simple, linear in the parameters, and linear in the independent variable. It is “simple” in a sense that there is only one independent variable, “linear in the parameters” because of the sense that no parameter appears as an exponent or is multiplied or divides by another parameter and “linear in independent variable” because this variable appears only in the first power. A model which is both linear in the parameter and the independent variable is also called first-order model.

The parameters β_0 and β_1 in regression model (2) are called regression coefficients. β_1 is the slope coefficient of the regression line. It measures the change in the mean of the probability distribution of Y per unit increase in X . The parameter β_0 is the intercept coefficient of the regression line. If the scope of the model includes, $X = 0$, β_0 gives the mean of the probability distribution of Y at $X = 0$. When the scope of the model does not cover $X = 0$, β_0 does not have any particular meaning as a separate term in the regression model.

5. Simulation Technique

In this paper, we have considered simple linear regression model with one explanatory variable. As we know, in simple linear regression model, the error term ε_i are normally distributed. But in this paper, we assumed the random error term ε_i to follow the g -and- k distribution. We have observed the extent of non-normality on the size and power of test $H_0: \text{slope} = 0$ by varying the skewness and the kurtosis parameter of the g -and- k distribution. Using the g -and- k distribution allows us to quantify how much the data depart from normality in terms of the values chosen for the g (skewness) and k (kurtosis) parameters. For $g = k = 0$, the quantile function for g -and- k distribution is just the quantile function of a normal variate.

In order to observe the power of the tests, expressions for the power curve are required. However, in practice it is not practical to obtain analytic expressions for these power

functions. Instead, we have conducted a simulation to estimate these power functions for various combinations of the g and k parameter values for the error distribution from the g -and- k distribution. In simulating size and power, we have considered a simple linear regression model of y with one explanatory variable x where the explanatory variable was generated from Uniform distribution. The values of the slope coefficient β_1 are to be varied from -4 to 4. The distribution of error is taken as g -and- k distribution where the values of skewness parameter g is taken from (-2, 2) and kurtosis parameter are -0.5, 0, 0.5, 1. While considering the distribution, the location parameter of the distribution A is taken to be 0, scale parameter B is taken as 1, constant c is considered as 0.83. The random variable z (0, 1) is a standard normal random variate. To get the power of t -test, we test $H_0: \beta_1 = 0$ against the alternative $H_1: \beta_1 \neq 0$. For each combination of g and k , we determine the number of rejections of the null hypotheses out of 10,000 times simulation for each value of β_1 other than zero and divide the total number of rejections by 10,000 which gives the power of the test. To get the size we generate data under the null hypotheses and repeat the above procedure. We have considered different combinations of g and k for different sample size $n = 10, 20, 30$ and 100. The level of significance is considered as 0.05 throughout the simulation. The above procedure is conducted by using statistical software R.

6. Size of t -test

The probability of rejecting a true null hypothesis is called type I error. It is also known as size of the test. First, we consider the effect of non-normality on the size of the test. For simulating, the size of t -test we generate explanatory variable x from uniform distribution and the random error ε from g -and- k distribution with location and scale parameters $A = 0$ and $B = 1$, respectively. Using statistical software R we generate data for sample size 10, 20, 30 and 100, and test

$$H_0: \beta_1 = 0.$$

Against the alternative,

$$H_1: \beta_1 \neq 0.$$

The test statistic is:

$$t = \frac{\hat{\beta}_1}{se(\hat{\beta}_1)} \sim t_{n-2} \quad \text{under } H_0 \quad \text{where, } se = \text{standard error}$$

To determine the size of the test, we generate data under the null hypothesis and repeat the test 10,000 times and divide the total number of times the hypothesis is rejected by 10,000. Tests are carried out using a nominal size of $\alpha = 0.05$. In Table 1 some simulation results are documented to see the effect of different level of non-normality on the size of t -test. For sample size 10, t -test is size robust under normal situation, but under non-normal

situation, if the contamination is small there is a little effect on the size of the test. For sample size 20 and 30, we see that both the skewness and kurtosis parameter has very little effect on the size of t -test. For sample size 100, even in the case of non-normal situation, t -test is almost size robust.

Table 1. Size of t -test for different combinations of (g,k) with varying sample sizes.

n	g	k	Size	n	g	k	Size
10	0.0	0.0	0.051	20	0.0	0.0	0.0508
	0.0	0.3	0.0466		0.0	0.3	0.0484
	0.0	0.5	0.0399		0.0	0.5	0.0443
	0.0	0.8	0.0386		0.0	0.8	0.041
	0.0	1.0	0.0372		0.0	1.0	0.0356
	0.0	-0.2	0.0492		0.0	-0.2	0.0534
	0.0	-0.3	0.0546		0.0	-0.3	0.0507
	0.0	-0.5	0.0513		0.0	-0.5	0.0515
	0.5	0.0	0.0479		0.5	0.0	0.0479
	0.7	0.0	0.0496		1.0	0.0	0.0522
	1.0	0.0	0.0481	1.5	0.0	0.0452	
30	0.0	0.0	0.0536	100	0.0	0.0	0.0522
	0.0	0.3	0.0454		0.0	0.3	0.0488
	0.0	0.5	0.0471		0.0	0.5	0.0492
	0.0	0.8	0.0477		0.0	0.8	0.0529
	0.0	1.0	0.0399		0.0	1.0	0.0438
	0.0	-0.2	0.0502		0.0	-0.2	0.0465
	0.0	-0.3	0.0510		0.0	-0.3	0.0505
	0.0	-0.5	0.0518		0.0	-0.5	0.0484
	0.5	0.0	0.0485		0.5	0.0	0.0495
	1.0	0.0	0.0486		1.0	0.0	0.051
	1.5	0.0	0.0454	1.5	0.0	0.0524	
	2.0	0.0	0.0478	2.0	0.0	0.0494	

7. Power of t -test

Power of the test is an important and considerable matter. Powerful test gives better conclusion, so it is desired to use powerful test. The probability of accepting a false null hypothesis is called Type II error. The power of the test is the probability that the Type II error will not occur.

To simulate power, we test $H_0 : \beta_1 = 0$.

Against the alternative, $H_1 : \beta_1 \neq 0$.

The test statistic is:

$$t = \frac{\hat{\beta}_1}{se(\hat{\beta}_1)} \sim t_{n-2} \text{ under } H_0$$

where, se =standard error.

We generate data using $\beta_1 \in (-4, -3.5, -3.0 \dots 3.0, 3.5, 4)$ and repeat the test procedure 10,000 times. We first determine the number of rejections of the test out of 10,000 times for each value in the mentioned set and divide the total number of rejections by 10,000, with the level of significance $\alpha = 0.05$. We generate power for sample size $n=10, 20, 30$, and 100 with various combinations of shape parameters g and k to see the extent of effect of skewness and kurtosis on the power of t -test.

Figs. 1 through 7 show the power curve of t -test with various combinations of (g,k) for sample size $n = 10, 20, 30$, and 100. In Fig. 1, for sample size 10, from the curves of $(g,k) = (0,0), (0,0.3), (0,0.5), (0,0.8), (0,1)$, we see that for fixed $g = 0$ and varying the kurtosis parameter k in positive direction the power of the test decreases vastly than that of the normal data $(g,k) = (0,0)$. The curves of $(g,k) = (0,0), (0.5,0), (0.7,0), (1,0)$ shows that for fixed $k = 0$, and varying the skewness parameter g , the power of the t -test decreases slightly than that of normal data. The curves of $(g,k) = (0,0), (0.5,0), (0,0.5), (0.5,0.5)$ and $(0,0), (1,0), (0,1), (1,1)$ shows that increasing kurtosis parameter reduces the power more than that of the skewness parameter. In Fig. 2, from the curves of $(g,k) = (0,0), (0,-0.2), (0,-0.3), (0,-0.5)$, we see the effect of negative kurtosis on the power and it is seen that when the kurtosis parameter increases in negative direction t -test gives better power than that of normal data. The curves of $(g,k) = (0,0), (0.5,0), (0,-0.3), (0,-0.5)$ shows that negative kurtosis increases the power than that of normal data whereas varying the skewness parameter g decreases the power slightly. Fig. 3 also shows the same pattern of change in power in varying the skewness and kurtosis parameter for sample size 20. In Fig. 4, the curve of $(g,k) = (0,0), (-0.5,0), (-1,0), (-1.5, 0), (-2,0)$ shows that when the values of g increases in negative direction the power turn down slightly than that of normal curve. Fig. 5 and Fig. 6 shows the change of power for different combination of (g,k) for sample size 30 in the same way. In Fig. 7, for sample size $n = 100$, the curves of $(g,k) = (0,0), (0,0.3), (0,0.5), (0,0.8), (0,1)$, shows that the loss of power increases when the kurtosis parameter is increased in positive direction whereas the curves of $(g,k) = (0,0), (0.5,0), (1,0), (1.5,0), (2,0)$ shows the t -test to some extent robust for varying the skewness parameter in positive direction. Fig. 7 shows that for large sample size i.e. $n = 100$ the difference between the power in normal and non-normal situation decreases to some extent.

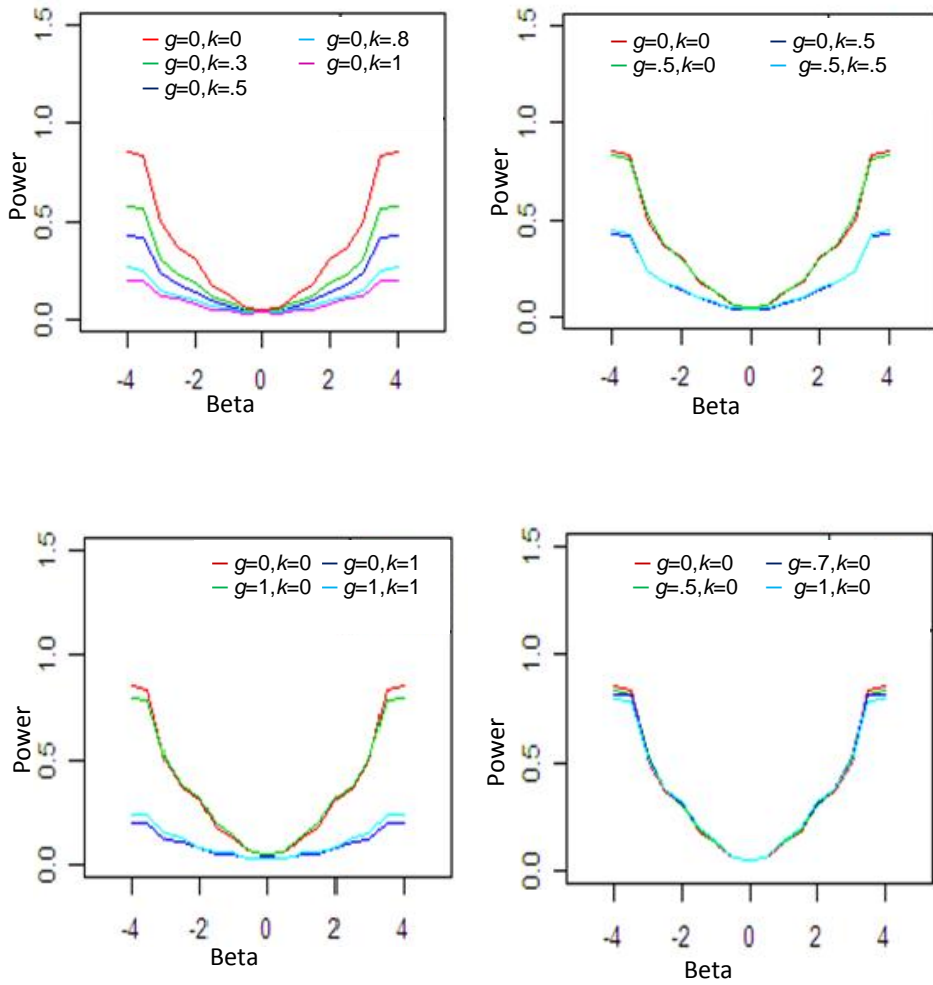
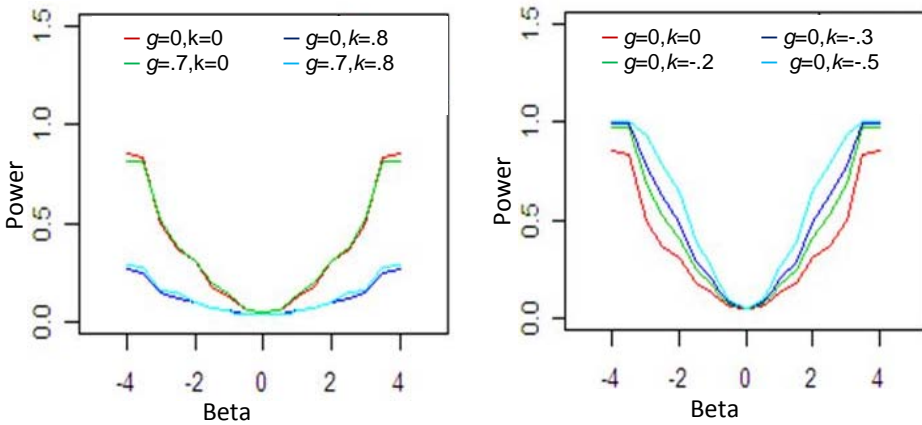


Fig. 1. Power curves of *t*-test with various combinations of (g, k) for sample size 10.



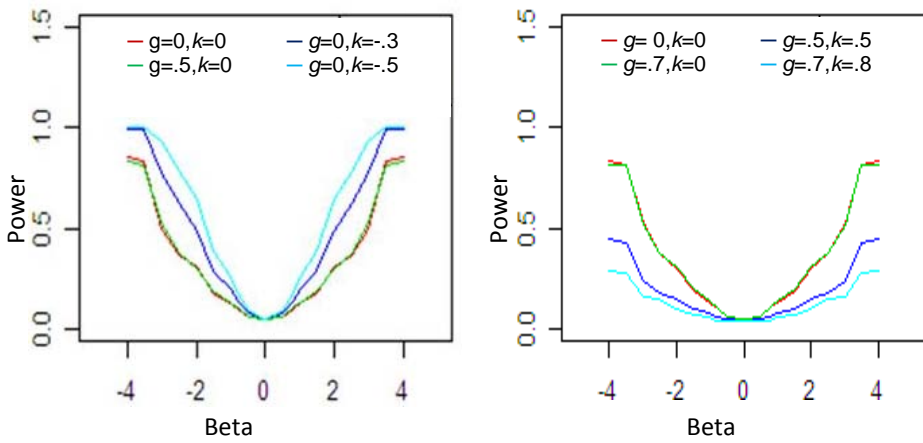


Fig. 2 . Power curves of t -test with various combinations of (g,k) for sample size 10.

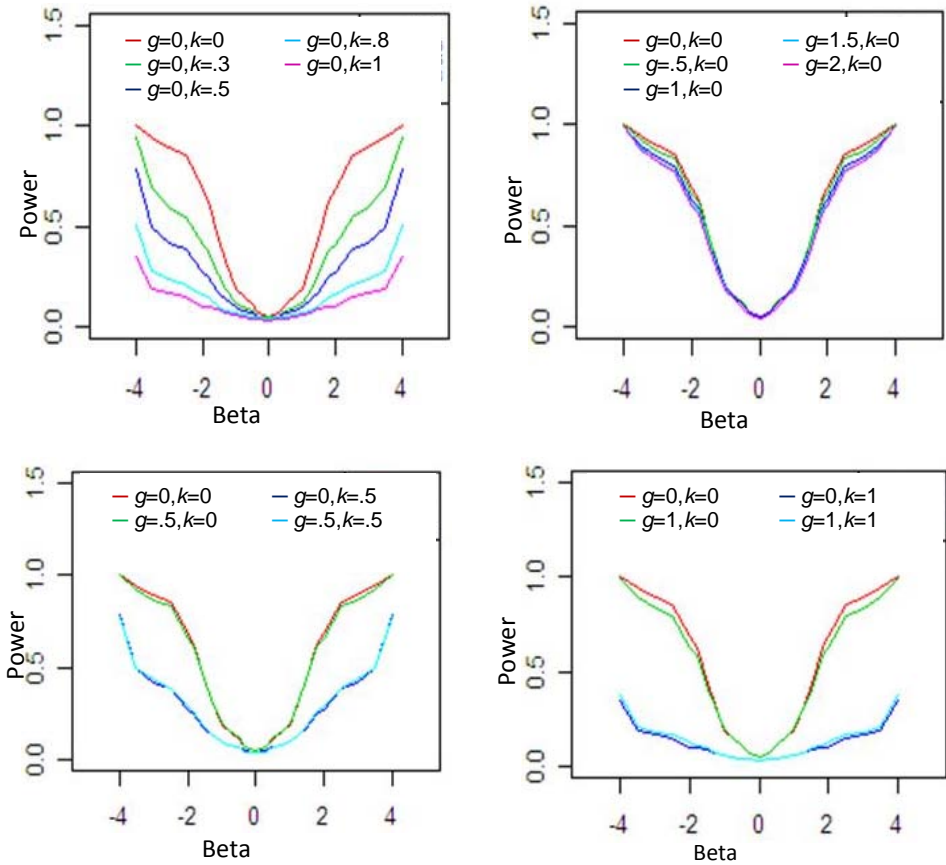


Fig. 3. Power curves of t -test with various combinations of (g,k) for sample size 20.

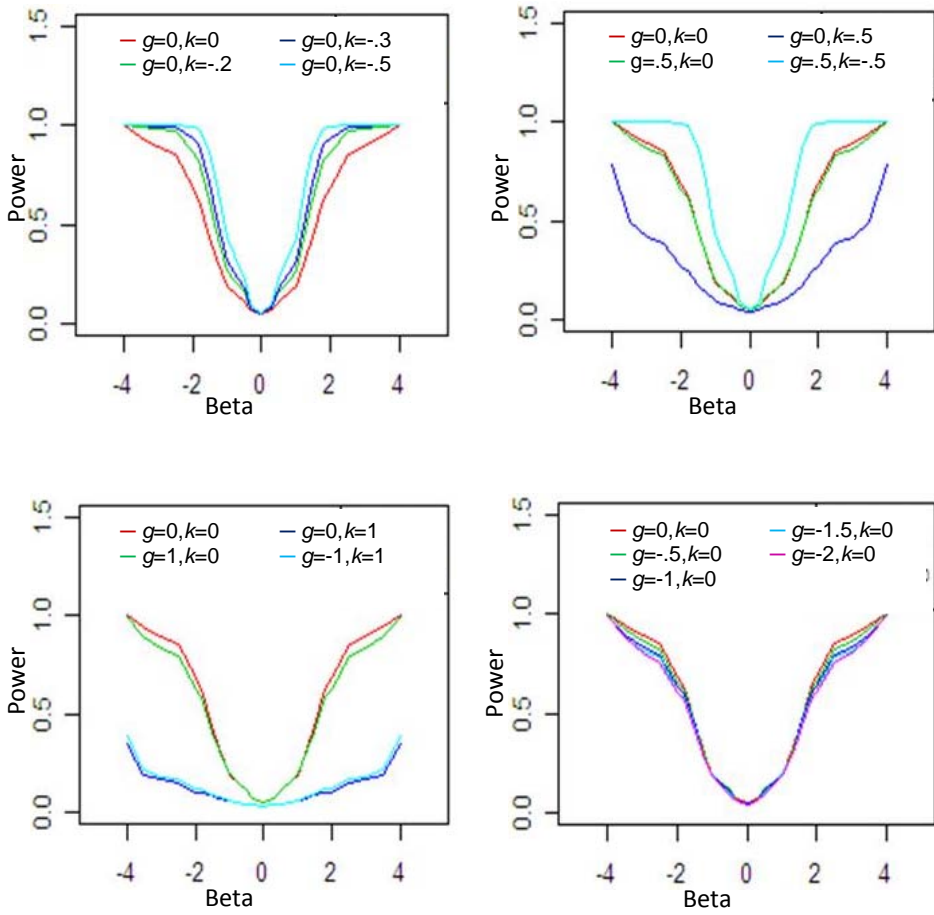
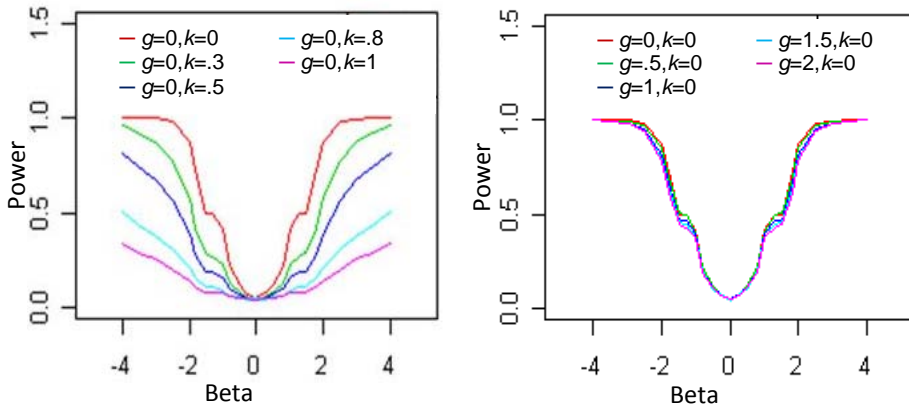


Fig. 4. Power curves of *t*-test with various combinations of (g, k) for sample size 20.



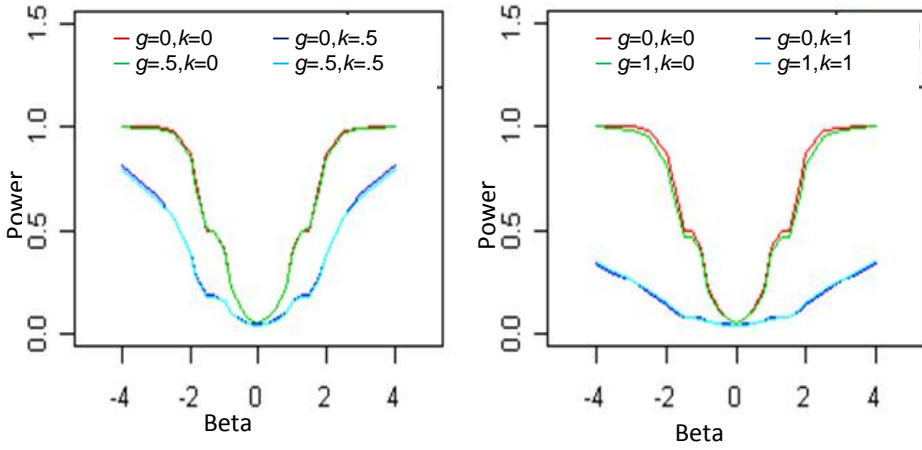


Fig. 5. Power curves of t -test with various combinations of (g,k) for sample size 30.

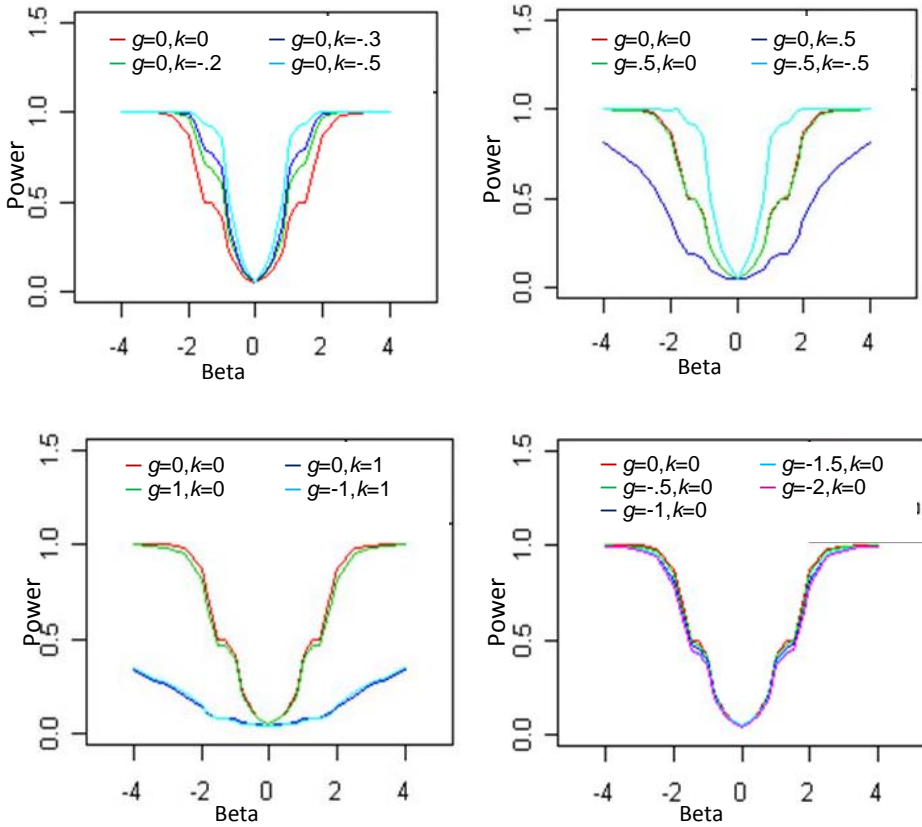


Fig. 6. Power curves of t -test with various combinations of (g,k) for sample size 30.

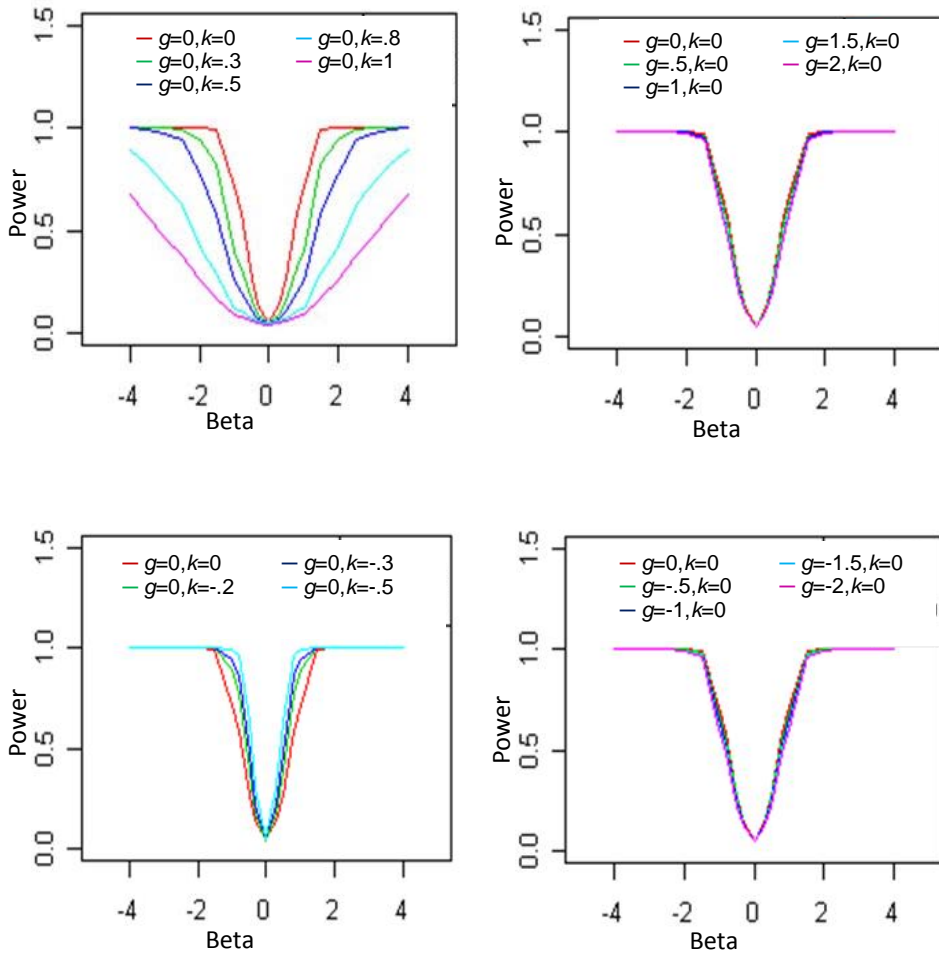


Fig. 7. Power curves of t -test with various combinations of (g, k) for sample size 100.

8. Results and Discussion

Observations from Figs. 1 through 7 indicate that t -test is neither optimally robust nor size robust under non-normal situation. The power of the test is badly affected by the sample skewness and kurtosis parameter. If the contamination is small, skewness and kurtosis parameter has little effect on the size of the test. The simulation results can be summarized in the following ways:

- a. Comparing Fig. 1, Fig. 3, Fig. 5, and Fig. 7, it is apparent that as the value of the kurtosis parameter k increases in positive direction, t -test gives less power than that of normal data.

- b. From the Fig 1 to Fig. 7 we can say that as the sample size increases, the difference between the powers in normal and non-normal situation decreases.
- c. From Fig. 1, Fig. 3, Fig. 5, and Fig. 7, the loss of power increases slightly than that of the normal data as the skewness parameter g increases. Skewness parameter has very little effect on the power of the test.
- d. From Fig. 2, Fig. 4, Fig. 6 and Fig. 7, it is clear that t -test gives better power than that of the normal situation when the data have negative kurtosis.
- e. From Fig. 4 and Fig. 6, it is clear that the power of the test decreases slightly than that of the normal data when the skewness parameter g increases in positive or negative direction.
- f. From Table 1 we can say that the skewness parameter g and the kurtosis parameter k have a little effect on the size of the t -test.

9. Conclusion

Extent of effect on power of t -test for simple linear regression model is measured numerically and power curves are shown graphically. We can distil our study into the following four observations:

- a. As the value of the parameter k increases in positive direction, t -test gives less power than that of normal data and in case of negative kurtosis it gives better power.
- b. Skewness parameter g has a very little effect on the power of the test.
- c. As the sample size increases the difference between the power in normal and non-normal situation decreases.
- d. Kurtosis parameter has more effect on the size and power of the test than that of skewness parameter for simple linear regression model.

References

1. S. S. Sawilowsky and R. C. Blair, *Psyc. Bul.* **111** (2), 352 (1992).
<http://dx.doi.org/10.1037/0033-2909.111.2.352>
2. L. Chen, *J. of American Statist. Asso.* **90** (430), 767(1995).
<http://dx.doi.org/10.1080/01621459.1995.10476571>
3. A. M. Carolan and J. C. W Rayner, *Commun.in Statist.* **29** (7), 1569 (2000).
<http://dx.doi.org/10.1080/03610920008832565>
4. A. M. Carolan and J. C. W Rayner, *J. Appl. Math. Deci. Sci.* **5** (1), 1 (2001).
<http://dx.doi.org/10.1155/S1173912601000013>
5. A. Khan and G. D. Rayner, *J. App. Math. Deci. Sci.* **7** (4), 187(2003).
<http://dx.doi.org/10.1155/S1173912603000178>
6. D. Rasch and V. Gulard, *Psyc. Sci.* **46**, 175 (2004).
7. H. Yanagihara, *J. Jpn. Statist. Soc.* **37** (1), 135 (2007).
8. A. S. M. Borhan and A. Khan, *Dhaka Univ. J. of Sci.* **57** (1), 61 (2009).

9. E. Schmider, M. Ziegler, E. Danny, L. Beyer, Buhner, *European J. Res. Method. Behav. Soc. Sci.* **6**(4), 147(2010). <http://dx.doi.org/10.1027/1614-2241/a000016>
10. M. Uddin, P. M. Goodrum, and K. C. Mahboub, *J. Trans. Res. Board* **2228**, 70 (2011). <http://dx.doi.org/10.3141/2228-09>
11. S. Landau and D. Stahl, *Stat. Methods Med. Res.* (2012). <http://dx.doi.org/10.1177/092280212439578>
12. H. L. MacGillivray and W. H. Cannon, Generalizations of the *g*-and-*h* distributions and their uses, Preprint (2002).
13. G. D. Rayner, and H. L. MacGillivray, *Stat. Comp.* **12**, 57 (2002). <http://dx.doi.org/10.1023/A:1013120305780>
14. J. Neter, W. Wasserman, and M. H. Kutner, *App. Linear Reg. Models* (Richard D. Irwin, INC., 1983).