M - 15

SELECTION OF THE BEST UNIVARIATE NORMALITY TEST ON THE CATEGORY OF MOMENTS USING MONTE CARLO SIMULATION

Sugiyanto and Etik Zukhronah
Mathematics Department, FMIPA UNS

Abstract

There are three univariate normality tests on the category of moment that are Geary, D'Agostino Pearson and lagrange multiplier tests. In this research, we compared them to determine which test has the highest sensitivity in the concluding test normality of the data. For the sample size n less than or equals to 30, the three tests can not be used to check for normality, whereas for sample size between 30 and 85, Geary's test was the best tests to check normality and for sample size more than or equals to 85, the three tests were can be used to check for normality. However, for samples that have a uniform distribution, the three tests were good to check normality when the sample size more than or equals to 130.

Key words: normality test, moments category, Geary, D'Agostino Pearson and lagrange multiplier tests

INTRODUCTION

Normality assumption of the data is widely used in many application. There are many tests that can be used to check for normality. Arshad et al. [1] revealed that the normality test can be categorized into four categories namely the chi-square test, moment-based tests, correlation-based tests, and tests based on empirical distribution functions. Meanwhile, according to Sumarjaya [6], the normality test can be grouped into several categories, namely: tests based on empirical distribution functions, moment-based tests, tests based on regression and correlation, entropy-based tests, tests based on the characteristics of Polya, Kernel-based tests, and test based on nonparametric methods.

The method used to construct these tests differ from each other, with the result of each test has its own advantages and disadvantages. For a group of specific data may correspond to a particular test, but not in accordance with other tests. These conditions are sometimes confusing to choose which test to use. In this study will be selected the best univariate tests on moments categories through Monte Carlo simulations. From this study, it can be used to choose a fitting test.

RESEARCH METHOD

In this research, data were randomly generated from exponential, gamma, uniform, Weibull, chi-square, student, Snedecor, beta, lognormal distributions with sample size 10 and repeat it as much as 10,000 times, and then calculate the probability of rejecting Ho. The process is repeated for sample size 15, 20, 25, ..., 100, then made the plot of probability of rejecting Ho.

RESULT AND DISCUSSION

Normality test based on moments categories include: D'Agostino Pearson test, Geary's test, and Lagrange multipliers test. The three tests will be compared through Monte Carlo simulations to determine the best test.

D'Agostino Pearson test procedure:

a. Hypothesis

Ho: a random sample is from the normally distributed population.

H1: a random sample is not from the normally distributed population.

b. The level of significance, $\alpha = 5\%$

c. Critical area: Ho is rejected when $K^2 > 5.991$

d. Statistical test

Statistical test of D'Agostino Pearson test is combining the coefficient of slope $\left(\sqrt{b_1} = \frac{m_3}{(m_2)^{3/2}}\right)$ and the coefficient of kurtosis $\left(b_2 = \frac{m_4}{(m_2)^2}\right)$, with $m_j = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^j$. D'Agostino Pearson statistic defined by the D'Agostino and Pearson [3] is $K^2 = Z^2(\sqrt{b_1} + Z^2(b_2)$ with

$$Z(\sqrt{b_1}) = \delta \ln\left(\frac{Y}{\alpha} + \left(\left\{\left(\frac{Y}{\alpha}\right)^2 + 1\right\}^{\frac{1}{2}}\right)\right)$$

$$Z(b_2) = \frac{\left(1 - \frac{2}{9A}\right) - \left[\frac{1 - \frac{2}{A}}{1 + x\sqrt{\frac{2}{A-4}}}\right]^{\frac{1}{3}}}{\sqrt{2/(9A)}}$$

According to D'Agostino and Stephens [4], the value of $\delta, Y, \alpha, W^2, A$ can be written by

$$Y = \sqrt{b_1}\left\{\frac{(n+1)(n+3)}{6(n-2)}\right\}^{1/2},$$

$$\delta = \frac{1}{\sqrt{\ln(W)}},$$

$$\alpha = 2/(W^2 - 1)^{1/2},$$

$$W^2 = -1 + \left\{2\left(\beta_2(\sqrt{b_1}) - 1\right)\right\}^{1/2}, \text{with } \beta_2(\sqrt{b_1}) = \frac{3(n^2+27n-70)(n+1)(n+3)}{(n-2)(n+5)(n+7)(n+9)}$$

$$A = 6 + \frac{8}{\sqrt{\beta_1(\sqrt{b_2})}}\left[\frac{2}{\sqrt{\beta_1(\sqrt{b_2})}} + \sqrt{1 + \frac{4}{\sqrt{\beta_1(\sqrt{b_2})}}}\right], \text{ with } \beta_1(\sqrt{b_2}) = \frac{6(n^2-5n+2)}{(n+7)(n+9)}\sqrt{\frac{6(n+3)(n+5)}{n(n-2)(n-3)}}.$$

Geary's test procedure:

a. Hypothesis

Ho: a random sample is from the normally distributed population.

H1: a random sample is not from the normally distributed population.

b. The level of significance. $\alpha = 5\%$
c. Critical area: Ho is rejected when G > 5.991
d. Statistical test

Statistical tests of Geary's test is using the two moments, namely G_w^2 and G_w^{2*} From the two moments, it can be defined the Geary's statistical test. Geary's statistical test is modified by Cho and Im [2] by taking the two test moments of first sample sm_j . The formulation of modified Geary's statistical test is $Z_w = (n+2)^{\frac{1}{2}}(\widehat{w} - 3)/3{,}54$

$$G = n\left[\frac{\left(a_1 - \left(\frac{2}{\pi}\right)^{\frac{1}{2}}\right)^2}{(1-(3/\pi))} + \frac{a_2^2}{(3-(8/\pi))}\right]$$

with

$$a_j = \frac{sm_j}{\hat{\sigma}^j}, j = 1, 2, ...$$

$$sm_j = (1/n)\sum_{i=1}^{n}(sgn(x-\bar{x})(x-\bar{x})^j)$$

$$\hat{\sigma} = n^{-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

where G is the modified Geary's statistical test, $n$is the number of random samples, π is 3.14 and $sgn(x-\bar{x})$ is the value of sign $(x-\bar{x})$ that can be expressed as

$$sgn(x-\bar{x}) = \begin{cases} -1, \text{jika}(x-\bar{x}) < 0 \\ 0, \text{jika}(x-\bar{x}) = 0 \\ 1, \text{jika } (x-\bar{x}) > 0 \end{cases}$$

Lagrange multiplier test procedure:

a. Hypothesis

Ho: a random sample is from the normally distributed population.

H1: a random sample is not from the normally distributed population.

b. The level of significance. $\alpha = 5\%$
c. Critical area: Ho is rejected when JBU > 5.991
d. Statistical test

Statistical test of Lagrange multiplier is using the coefficient of slope, $(\sqrt{b_1})$ and the coefficient of kurtosis(b_2) . Statistical test of Lagrange multiplier is defined by Jarque and Bera [5],

$$JB = n\left[\frac{\left(\sqrt{b_1}\right)^2}{6} + \frac{(b_2 - 3)^2}{24}\right]$$

Statistical test of Lagrange multiplier JB is then modified by Urzua in Sumarjaya [6]. Modification of statistical test of Lagrange multiplier is

$$JBU = \left[\frac{\left(\sqrt{b_1}\right)^2}{v_2} + \frac{(b_2 - v_2)^2}{v_3}\right]$$

with

$$v_1 = \frac{3(n-1)}{n+1}$$

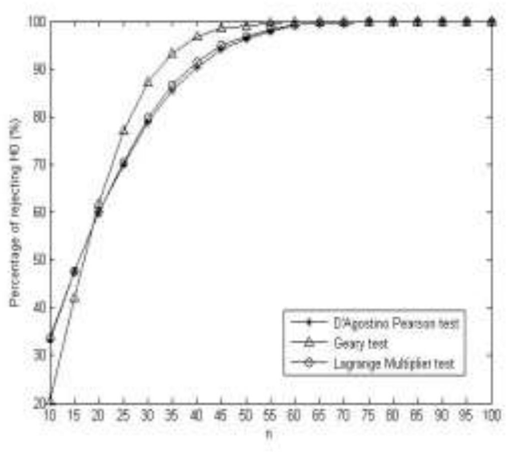
$$v_2 = \frac{6(n-2)}{(n+1)(n+3)}$$

$$v_3 = \frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}$$

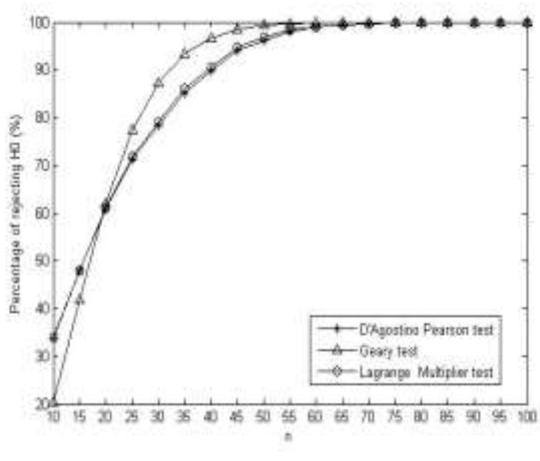
n is the number of random samples.

Monte Carlo Simulation

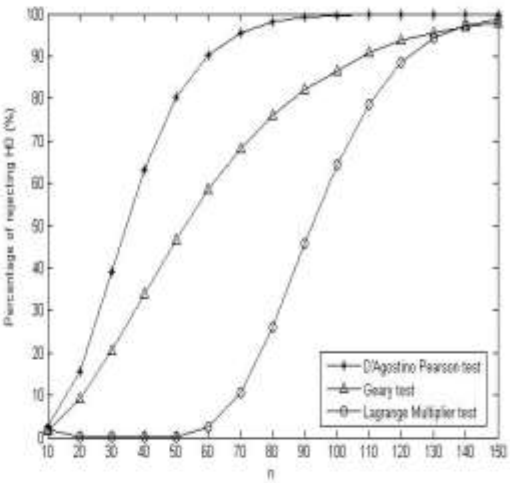
Simulation is done by generating random samples of various distributions that are not normally distributed, i.e. exponential, gamma, uniform, Weibull, chi-square, student, Snedecor, beta, and lognormal distributions. The sample size that is generated from each distribution are 10,15,20, ..., 100. Each sample sizes of each distributions is repeated 10,000 times and calculated the percentage of rejecting Ho. The result of simulation can be seen in figure (a)-(i).



(a)

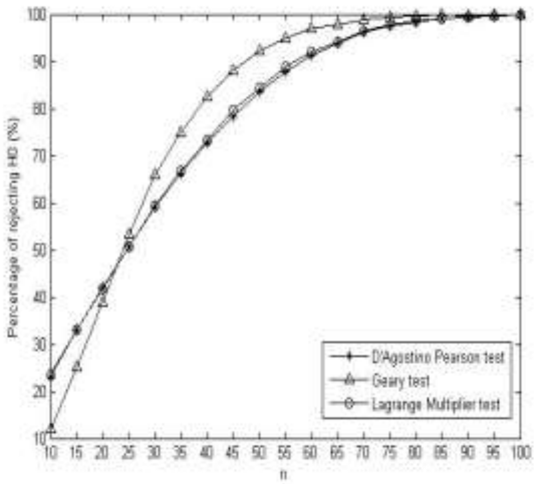


(b)



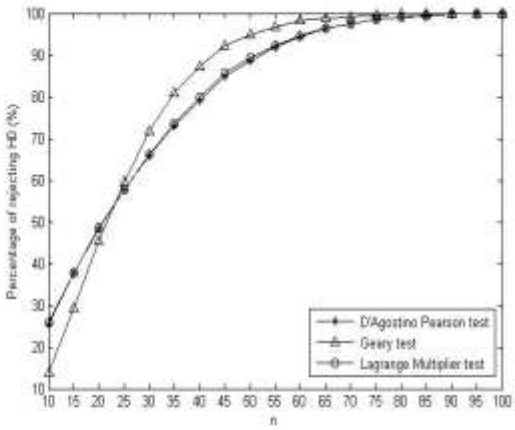
D'Agostino Pearson test
Geary test
Lagrange Multiplier test
n

(c)



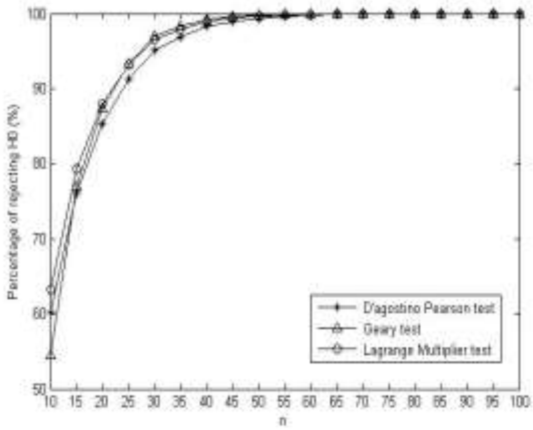
D'Agostino Pearson test
Geary test
Lagrange Multiplier test
n

(d)



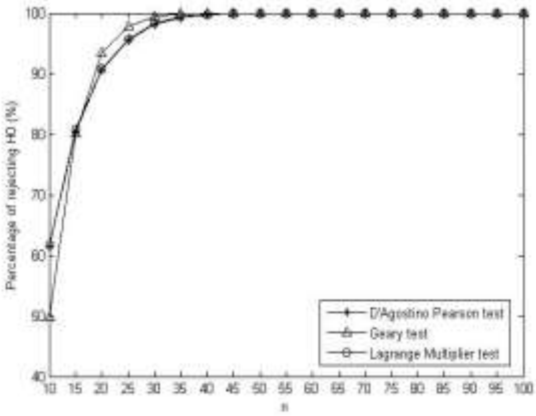
D'Agostino Pearson test
Geary test
Lagrange Multiplier test
n

(e)

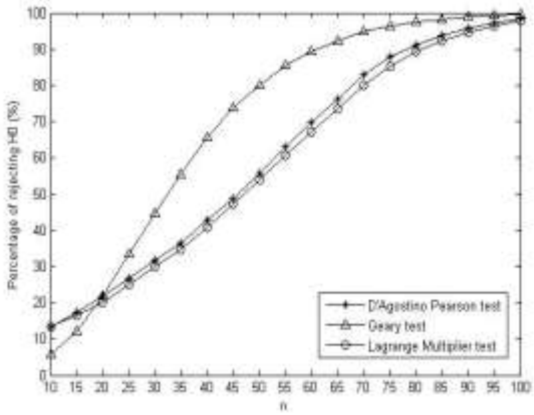


D'agostino Pearson test
Geary test
Lagrange Multiplier test
n

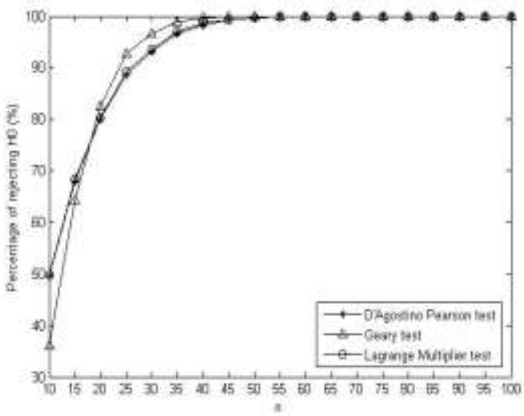
(f)



(g)



(h)



(i)

Figure (a) is the result of random sample simulation that generated from the exponential distribution EXP (3). For sample size n less than or equal to 25, the three tests were not able to distinguish whether the sample is normally distributed. For a sample size between 25 and 40, this sample size is not good to check normality, whereas for sample size more than or equals to 40, the three tests were very good to check normality.

Figure (b) is the result of random sample simulation that generated from a gamma distribution GAM (1.2).). For sample size n less than or equal to 25, the three tests were not able to distinguish whether the sample is normally distributed. For a sample size between 25 and 40, this sample size is not good to check normality, whereas for sample size more than or equals to 40, the three tests were very good to check normality.

Figure (c) is the result of random sample simulation that generated from a uniform distribution UNIF (-1.2). For sample size n less than or equal to 35, D'Agostino Pearson test was unable to distinguish whether the sample is normally distributed, For a sample size between 35 and 60, this sample size is not good to check normality, whereas for sample size more than or equals to 60, that test was very good to check normality. For sample size n less than or equal to 60, the Geary's test was not able to distinguish whether the sample is normally distributed, for sample size between 60 and 120, that test was not good to check normality, while for sample size more than or equals to 120, that test was very good to check normality. For the sample size less than or equals to 90, Lagrange multipliers test was not able to distinguish whether the sample is normally distributed, for the sample size between 90 and 130 that test wa not good to check normality, while the sample size more than or equals to 130 that test was very good to check normality.

Figure (d) is the result of random sample simulation that generated from weibull distribution WEI (3,1.25). For sample size n less than or equal to 25, the three tests were not able to distinguish whether the sample is normally distributed. For a sample size between 25 and 60, this sample size is not good to check normality, whereas for sample size more than or equals to 60, the three tests were very good to check normality.

Figure (e) is the result of random sample simulation that generated from the distribution of chi square $\chi^2(3,1.25)$. For sample size n less than or equal to 25, the three tests were not able to distinguish whether the sample is normally distributed. For a sample size between 25 and 55, this sample size is not good to check normality, whereas for sample size more than or equals to 55, the three tests were very good to check normality.

Figure (f) is the result of random sample simulation that generated from student distribution $t(1)$. For sample size n less than or equal to 10, the three tests were not able to distinguish whether the sample is normally distributed. For a sample size between 10 and 25, this sample size is not good to check normality, whereas for sample size more than or equals to 25, the three tests were very good to check normality.

Figure (g) is the result of random sample simulation that generated from the distribution snecdecor $F(2,3.5)$. For sample size n less than or equal to 10, the three tests were not able to distinguish whether the sample is normally distributed. For a sample size between 10 and 15, this sample size is not good to check normality, whereas for sample size more than or equals to 15, the three tests were very good to check normality.

Figure (h) is the result of random sample simulation that generated from the beta distribution $BETA(1,3)$. For sample size n less than or equal to 30, the three tests were not able to distinguish whether the sample is normally distributed. For a sample size between 30 and 65, Geary's test was not good to check normality, whereas for sample size more than or equals to 65, Geary's test was very good to check normality. For sample sizes between 35 and 85, D'Agostino Pearson and Lagrange multipliers tests were less well to check normality, while for the sample size more than or equals to 85, D'Agostino Pearson and Lagrange multipliers tests were very good to check normality.

Figure (i) is the result of random sample simulation that generated from the lognormal distribution $logN(3,1)$. For sample size n less than or equal to 10, the three tests were not able to distinguish whether the sample is normally distributed. For a sample size between 10 and 25, this sample size is not good to check normality, whereas for sample size more than or equals to 25, the three tests were very good to check normality.

CONCLUSION AND SUGGESTION

For the sample size n less than or equals to 30, the three tests can not be used to check for normality, whereas for sample size between 30 and 85, Geary's test was the best tests to check normality and for sample size more than or equals to 85, the three tests were can be used to check for normality. However, for samples that have a uniform distribution, the three tests were good to check normality when the sample size more than or equals to 130.

REFERENCES

[1] Arsad, M., Rsol, MT And M. I, Ahmad, *Anderson Darling and Modified Anderson Darling Test for the Generalized Pareto Distribution*, Pakistan Journal of Applied Sciences 3 (2003), no. 2, 85-88.

[2] Cho, D. W and KS Im, *A Test of normality Using Geary's Skewness and Kurtosis Statistics*, (2002).

[3] D'Agostino, R. B, and ES Pearson, *Testing for Departures From normality*, Biometrika 60 (1973), 613-622.

[4] D'Agostino, R. B, and MA Stephens, *Goodness of fit Techniques*, Marcel Dakker, New York, 1986.

[5] Jarque, CM And AK Bera, *A test of normality of Observations and Regression Residuals*, International Statistical Review 55 (1987), 163-172.

[6] Sumarjaya, IW, *Univariate Normality Test: A Literature Review*, Journal of Mathematics 1 (2010), 21-30.