Proceedings of the Pakistan Academy of Sciences:
A. Physical and Computational Sciences: 57 (1): 17-20 (2020)

ISSN: 2518-4245 (print), 2518-4253 (online)

Pakistan Academy of Sciences

Research Article

A Comparison of Non-parametric Modality Tests

Farrukh Jamal[1*], Gamze Ozel[2], Rana Muhammad Imran Arshad[3], and Shahina Imran[4]

[1]Department of Statistics, Government Sadiq Abbas Post Graduate College
Dera Nawab Sahib, Pakistan
[2]Department of Statistics, Hacettepe University, Ankara, Turkey
[3]Department of Statistics, Government Sadiq Egerton College, Bahawalpur, Pakistan
[4]Department of Economics, Government Sadiq Egerton College, Bahawalpur, Pakistan

Abstract: The non-parametric modality tests are widely and frequently used in finance, social, medicine, natural and biological sciences. In this study, we discuss the four non-parametric tests including Hartigan DIP (HD), Silverman's bandwidth (SB), proportional mass (PM), and excess mass (EM) tests for modality and multimodality. However, these tests have not been compared based on the size of the literature. Hence, this study compares these tests about the size and finds that which of them is the best test.

Keywords: SB test, PM test, ES test, Simulation, HD test, Non-parametric tests.

1. INTRODUCTION

It generally needs to estimate the number and location of modes for data coming from an unknown density [2]. A mode is known as the local maximum of the probability density function. When the density function with constant values at a peak, the points on the peak can be a single-mode.
Nonparametric methodology testing is a non-disseminated method for assessing proof of heterogeneity for a populace where potential subpopulations are very much isolated. Past research on this kind of system will be stretched out to assess the proof of heterogeneity of profits inside a populace [11]. The multimodality test will be considered as verification that the exhibitions of a particular variable shift as indicated by the various gatherings and periods [4]. Then again, in the SB test [11], a dismissal of the invalid speculation doesn't prompt the distinguishing proof of the reason for the multimodality. For instance, we don't have the foggiest idea of whether multimodality results from work. All we know is that multimodality is available. This is valuable as the methods and announcing quartiles of the parameter appraisals can cover significant data. Casual thickness review is additionally deficient because insignificant modes could be oddities owing to estimation blunder or other stochastic wonders [7]. Non-parametric density estimation is important for modeling the probabilistic or stochastic nature of the data. The present study compares four nonparametric tests including HD, SB, PM, and EM tests by estimating size via simulation and finds the best test for unimodality.

2. NON-PARAMETRIC TESTS

There have been several goals methods for multimodality tests. Many of them are about unimodality, bimodality or multimodality. Silverman [10-12] provided a test which depends critical bandwidths, the infimum of those smoothing parameters.

Data generating process (DGP) are obtained from standard normal distributted random variable. Density function for each test has a normal distribution with mean $\boldsymbol{\mu}$ and variance $\boldsymbol{\sigma^2}$ for estimating size and simulated critical value of the tests where X has the standard normal density function with μ= 0 and $\boldsymbol{\sigma^2=1}$. The size of a test

Received: March 2019; Accepted: March 2020
* Corresponding Author: Farrukh Jamal <drfarrukh1982@gmail.com>

called significance level of the test. This is type-I error which means probability of rejecting true hypothesis. α is size of test and (1- α) is level of confidence. In a two-tailed test, test size is sum of two symmetric areas at the tails of a probability distribution. These areas are called null hypothesis rejection regions in the sense that the rejection of the null hypothesis if a statistic falls into these regions.

For estimating the test size, we generate unimodal series from the standard normal variables by calculating the test statistics of each test and repeat this process 10,000 times by the Monte Carlo simulation technique. Then, we sort the calculated tests and find the critical value at 95th percentile and finally repeat steps from 1 to 3 for several choices of the parameters at sizes 50, 100, 200, 250 and 300. Steps for calculating size for the tests are summarized as follows:

- Step 1: Generate unimodal series from standard normal distributted variables.
- Step 2: Calculate test statistics of each test and repeat this process 10,000 times.
- Step 3: Sort the calculated tests and find the Critical value at the 95th percentile.
- Step 4: Repeat preivous steps for several choices of the parameters at sample sizes 50, 100 and 200, 250 and 300.

3. COMPARISON OF NON-PARAMETRIC MODALITY TESTS BY SIZE

3.1. DGP

The DGP of modality tests is obtained from the standard normal distributed variable. Probability density function to estimate simulated critical value and size for each test is calculated from X, where X is the standard normal distributed variable. The size of tests is done with the null hypothesis "H_0: There is modality".

3.2. Simulation Study

The procedure to make unimodal series, estimating size for finding the best and the worst test by using the Monte Carlo technique by repeating 10,000 times.

3.3. Calculating Critical Values

In this paper, a simulation study is conducted to obtain critical values. For a small sample size, the modality tests in the literature do not give true estimation values [8, 14]. Some of them are asymptotic and critical values need to work well in small samples.

3.4. Limitations

In this study, we only compare the modality tests for a univariate case for testing modality including SB test by Silverman [11], HD test proposed by Hartigan and Hartigan [6], PM test by Cavallo and Ringobon [1] and EM test by Muller and Sawitzki [9]. We only discuss the univariate case for testing modality on the other hand multimodality is out of scope for this paper.

3.5. Hypothesis

For the analysis of modality test size, we use following hypothesis at several choices and sizes:

H_0: It is evidence of unimodality

H_1:Otherwise (bimodality or multimodality)

Table 1. Four Test Size Distortion with Several Sample Sizes

	Modality Tests Sizes			
Sample Size	**HD**	**SB**	**PM**	**EM**
50	4.86	3.9	1.78	4.95
100	5.07	6.2	8.48	5.58
200	5.1	4.5	8	4.93
250	4.7	4.2	8.34	4.86
300	5.15	4.3	7.98	4.62

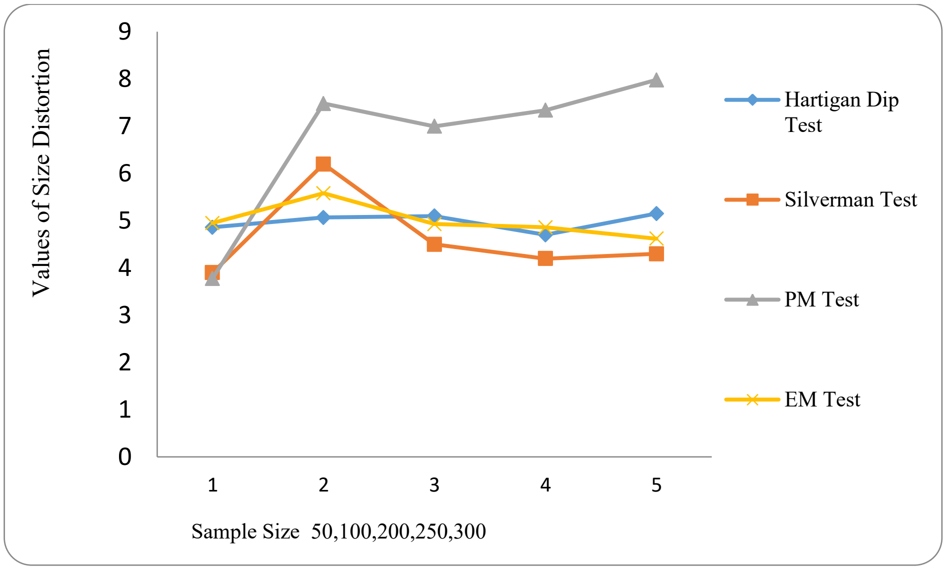

Fig. 1. Distortion of Size for Modality Tests

3.6. Computation of Size of Each Test

The present paper computes size for four modality tests for the simulation size 10000. For observing the minimum distortion in size from 5% of each test, we compute each test size for several sample sizes. The comparison of test sizes is presented in Table 1. As seen in Table 1, the distortion of HD and EM tests are close to 5% as compared to that of SB and PM tests. The PM test size distortion and this means that PM is the worst test. On the other hand, the HD test has the lowest distortion in size. Hence, it is the best test as compared to EM and SB tests.

3.7 Graphical Comparison for Size of Tests

Now, in Figure 1, we compare four modality tests graphically. Sample sizes values for 50, 100, 200, 250, and 300 are plotted. The resulting set of points is separately joined by a straight line for the tests.

The distortion size in HD, EM, SB tests is about 5% on the other hand distortion size for HD test is about 5% in Figure 1. Hence, the HD test is the best as regards size. The graph depicts that the distortion size for PM tests is high and hence, the worst test is PM.

The results show that HD and EM tests are close to 5% when we compare SB and PM tests. The distortion size of the PM test is large and hence, the worst test is based on size is PM. For this reason, the HD test has the lowest distortion concerning the size. Hence, it can be concluded that the HD test is the best.

Distortion sizes of EM and SB tests have been found near to 5% however the distortion size for HD test is about 5%. We concluded again that the HD test is the best and graphically distortion size of PM test is very high. For this reason, PM is the worst test.

4. DISCUSSION

This study aims to find the best modality test by size. It is concluded that regarding the 5% size distortion criteria, the HD test is the best to meet the criteria compared to EM and the SB tests, whereas a PM test s' is revealed far from complete. The size distortion of the PM test is very high, making this test the worst test based on size. On the other hand, the HD test has the smallest distortion size. So this is the best size-based test, after the EM and SB tests, respectively, with a small margin.

5. REFERENCES

1. A. Cavallo, and R. Rigobon, The Distribution of the size of price changes, NBER Working Paper No. 16760, 1-40 (1910).
2. N.I. Fisher, and J.S.Marron. Mode testing via the excess mass estimate. *Biometrica* 88(2) 499–517

(2001).

3. J.A. Hartigan. The span test of multimodality. In: *Classification and Related Methods of Data Analysis* H.H. Bock, (Ed.) North-Holland, Amsterdam, p. 229-236 (1988).
4. J.A. Hartigan, and P.M. Hartigan, The Hartigan Dip Test of unimodality. *The Annals of Statistics* 13(1): 70-84 (1985).
5. P. Hall, and M. York. On the calibration of Silverman's test for multimodality. *Statistica Sinica*,11: 515–536 (2001).
6. J.A. Hartigan, Algorithm AS217: Computation of the Hartigan Dip Test statistic to test for unimodality. *Journal of Applied Statistics* 34: 320–325 (1985).
7. M.C. Jones. Discretized and interpolated kernel density estimates. *Journal of American Statistical Association*, 84: 733-741 (1989).
8. M.C. Minnotte, and N.S. Jawhar, Adaptive kernel density estimation using mode information. Comput. *Sci. Statist.* 27: 545-549 (1995).
9. D.W. Muller, and G. Sawitzki. Excess mass estimates and tests for multimodality. *The Journal of the Acoustical Society of America* 86: 738 -746 (1991).
10. B.W. Silverman. Weak and strong uniform consistency of the kernel estimate of a density and its derivatives. *Annals of Statistics*. 6: 177-184 (1978).
11. B.W. Silverman. Using kernel density estimates to investigate multimodality. *Journal of Royal Statistical Society Series B*, 43: 97-99 (1981).
12. B.W. Silverman. Some properties of a test for multimodality based on kernel density estimates in J. F. C. Kingman and G. E. H. Reuter (Eds), Probability, Statistics, and Analysis, Cambridge University Press, Cambridge (1983).
13. B.W. Silverman. Density Estimation for Statistics and Data Analysis, Monographs on Statistics and Applied Probability. 26, Chapman and Hall, London (1986).
14. M.A. Wong, A bootstrap testing procedure for investigating the number of subpopulations. *Journal of Statistical Computation and Simulation*, 22: 99-112 (1985).