European Online Journal of Natural and Social Sciences 2019; www.european-science.com
Vol.8, No 4 (s) Special Issue: Perspectives of Economics and Management in Developing Countries
ISSN 1805-3602

Size of the Bimodality by using Trapezoidal Rule

Abdul Ghafar Shah*, Abdul Jabbar, Shahid Akbar
School of Economics, International Institute of Islamic Economics, International Islamic University, Islamabad, Pakistan
*Email: ghafarshahiiui@gmail.com
Tel.: +92-346-9291245

Abstract

This study investigates size of the bimodality to illustrate the gap between two modes/peaks in a bimodal distribution. To get this goal, one of the definite integral (i.e. Trapezoidal rule) is used on the mixture of normals (the mixture of one normal and the second one standard normal) with the selected parameter values of mean, mixing proportion and standard deviation. From the results, an increase of size of bimodality is identified if mean increases while all other parameters remain constant. A similar situation of increasing pattern of bimodality size has been observed when parametric values of mixing proportion increases (varies) while keeping other parameter values constant. However, a fraction changed is revealed in the size of bimodality if values of mean and mixing proportion are kept constant while standard deviation varies.

Keyword: Bimodality Size, Mixture of Normals, Mixing Proportion, Mean, Trapezoidal rule

Introduction

Researchers in a various areas are facing challenges i.e. not clear about the presentation of data where it is unimodal or bimodal e.g Knoll and Semikhatov (1998); Johnson and Yantis (1995); Roeder (1996); Sussman (1999); Volbrecht, Nergerand Randell (1997); etc.). Similarly, a lot of studies in social, health, numerical and natural sciences point out the same issue of bimodality. It is complicated to decide whether the data set follows a normal distribution (in unimodal case) or mixture of normals (in bimodal case) stated Frankland and Zumbo (2002). Mostly, it is anticipated that the data belongs to normal or mixture of normal distributions but this assumption is not necessary in each case (Yellot, 1971). The significance of the research for modality in statistics has been designated by Murphy (1964). Lindsay (1995) documented that theories about Mixture of distributions have been commonly used in social sciences. Robertson and Fryer (1969) described that, as two normal distributions joined collectively and formed another shape of a distribution which depends upon their parameters values. Schilling and Watkins (2002) documented that mixture of two normal distributions automatically lead to a bimodal distribution.

A likelihood ratio function for normal distribution and mixtures of normal distributions called coefficient of bimodality was used by Ashman and Bird (1994). In this connection Frankland and Zumbo (2002) introduced SPSS program for the distinction among unimodal normal distribution and a bimodal mixture of normal distributions. Choonpradub and McNeil (2005) detected the bimodality on the basis of augmentation in traditional boxplot by thickening the two ends of the box.

Using the same idea as an experiment this study considered mixture of normals (one standard normal and the second is normal) to find the size of bimodality with the help of its parameters. Also to check how these parameters effect the size of bimodality.

This study uses mixture of two normal distributions as data generating process "DGP" for checking bimodality size. Let x_1 is the random sample from the population i.e. first normal distribution with mean "μ_1" and variance σ_1^2 and x_2 is the random sample from the population i.e. second normal distribution with mean "μ_2" and variance σ_2^2 respectively.

$$Z = px_1 + (1-p)x_2 \tag{1}$$

Where "Z" is known as mixture of normal distributions with mixing probability "p" within the interval (0, 1).

Bimodality Size

In this study size of the bimodality illustrate the gap or area between two modes in a bimodal distribution. To calculate the [1]bimodality size we applied definite integrals i.e. Trapezoidal rule. On the other hand Riemann sums has low accuracy because it distributed the concern area in small rectangles as compare to Trapezoidal or Simpson's rules which distribute the area in various trapeziums. This study consider $\mu_1 = 0$, $\sigma_1^2 = 1$ (because of one standard normal distribution) and various values of other parameters i.e . p =(0.1, 0.2, 0.3,. . . . ., 0.9), μ_2 = (1, 2, 3, ... ,10) and σ_2^2 = (0.1, 0.2, 0.3, ... , 0.9).

Trapezoidal rule

It is not easy to estimate the integrals through analytical techniques. But, for the same approximate area we have used a numerical technique called Trapezoidal rule. The mathematical procedure of this technique is as follows;

$$\int_{\mu 1}^{\mu 2} f(z)dz = \frac{\Delta h}{2}[z_0 + 2(z_1 + z_2 + \ldots\ldots + z_{n-1}) + z_n)] \tag{2}$$

Where z_0, $z_m$are lower and upper limits respectively, here$z_0 = \mu_0$and$z_m = \mu_2$,

$$\Delta h = \frac{\mu_2 - \mu_1}{n}$$

"n" is the total number of trapeziums or sub-intervals of same size with (n+1) points. The accuracy is directly proportional to "n" and inversely proportional to"Δh".

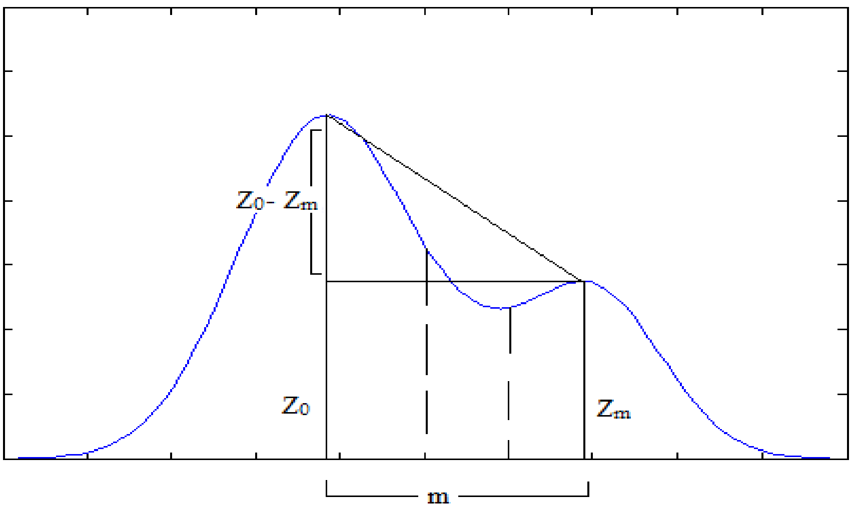

Figure 1. Size of the bimodality between two modes

Through any one of the definite integrals the bimodality size can be calculated. The surface which is above the bimodal distribution connected directly with two modes is the bimodality size (see Figure 1). Joining of the space which is calculated from numerical method and size makes a trapezoid. The heights of the two modes are denoted by "z_0 and z_m". Where the distance $(z_0 - z_m)$

[1]Bimodality size also describes distance or norm between two peaks/modes in a bimodal distribution.

is the vertical line of the right angle triangle. Similarly the distance *'m'* is the length and z_m is the width of the rectangle. In this study the bimodality size is estimated as the distinction between [2]area of the right angle triangle and the area through Trapezoidal rule. Figure 1 show both of the modes in a bimodal distribution where the area within these modes is the size of the same distribution. Following cases explains the bimodality size which influenced by the parameters of bimodal distribution.

Effect of mean "μ_2" in mixture of normal

As the mixture contains one standard normal and second one is normal, so we have changed the three parameters (i.e. μ_2, p and σ_2).This case possessed with varying values of "μ_2" while other two parameters have kept fixed.

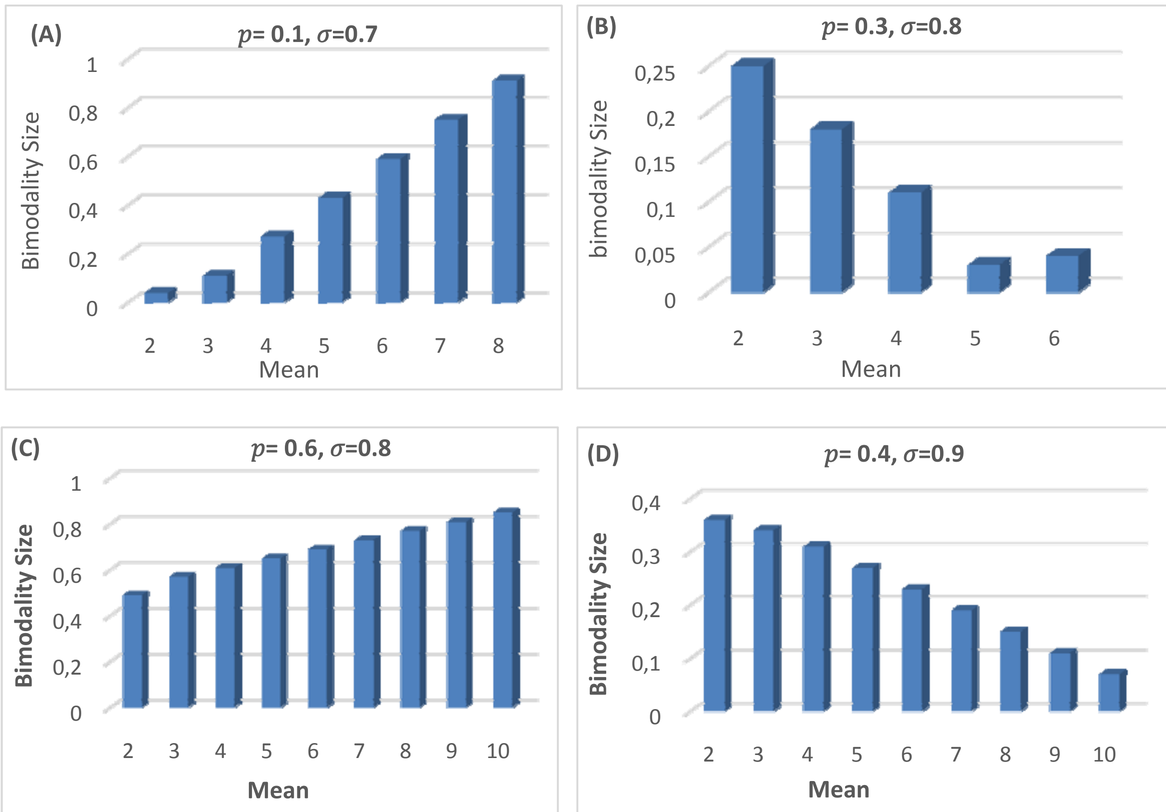

Figure 2. Size of the mixture of normal with various values of "μ_2"

Figure 2 (A) describes the bimodality size for different values of μ_2= (2, 3, ... , 8) and fixes p= *0.1*, $\sigma_2 = 0.7$. In this case the bimodality size increases as the value of μ_2 increases while keeping other parameter values as constant. For the same values of mean and combination of (p, σ_2)= {(0.1, 0.8), (0.2, 0.8), (0.1, 0.9)} result remains same.

Figure 2 (B) shows the bimodality size for different values of μ_2= (2, 3, ... , 6) and p= *0.3*, $\sigma_2 = 0.8$. Here, the bimodality size decreases as parametric values of 'μ_2' increases while keeping

[2] Area of the right angle triangle= $\frac{n(z_0 - z_m)}{2}$

other parameters (i.e. p,σ_2) constant. For the same values of mean and combination of (p, σ_2)= (0.2, 0.9) the result remains unchanged.

Similarly, Figure 2 (C) observes the bimodality size for the values of μ_2= (2, 3, ... , 10) and p= 0.4, $\sigma_2 = 0.9$. In this situation the bimodality size increases by increasing the ter $\mu_2$while keeping other parameters constant. At the same values of parameter mean, and combination of (p,σ_2)= {(0.2, 0.9), (0.6, 0.9)} the result remains same.

Figure 2 (D) identifies the bimodality size for the values of μ_2= (2, 3, ... , 6) and p= 0.3, $\sigma_2 = 0.8$. Here, the bimodality size decreases fractionally as the parameter mean values increases while keeping other parameters constant.

Effect of mixing proportion alpha "p" in mixture of normals

As we change the second parameter called mixing proportion 'p'and fix the remaining two parameters (i.e. μ_2, σ_2) then Figure 3 explains behavior of bimodality size.

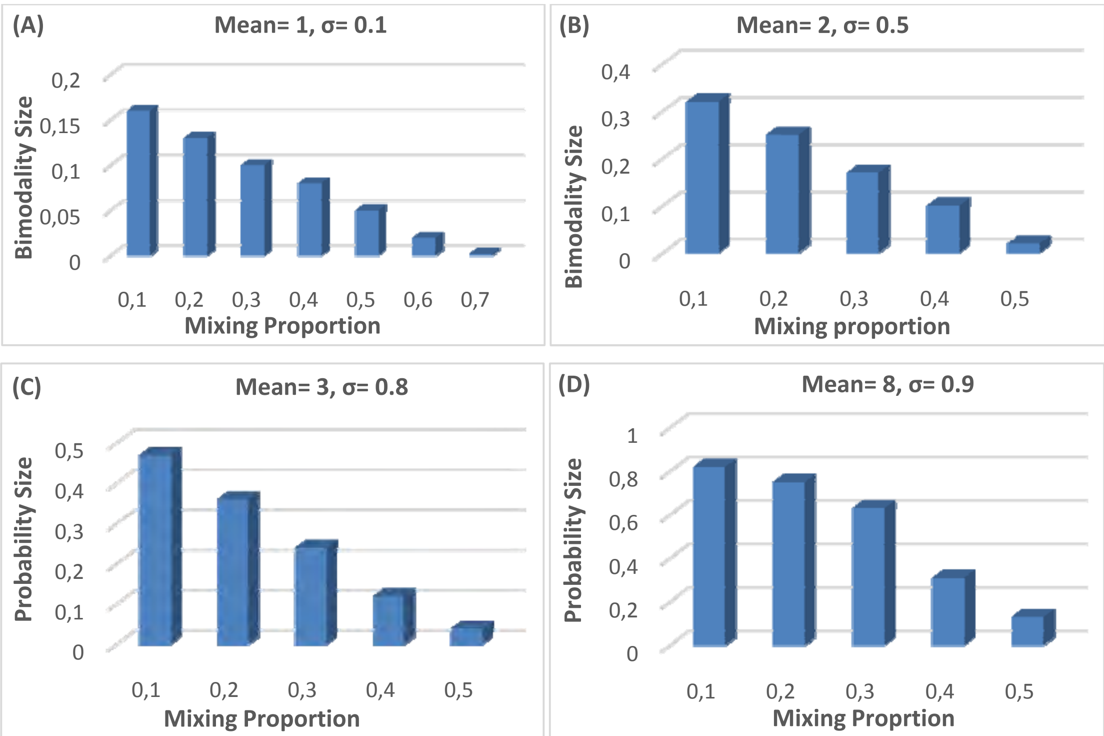

Figure 3. Size of the mixture of normal with various values of "p"

Figure 3 (A) shows the bimodality size for the values of p= (0.1, 0.2, 0.3,. . . . , 0.9) and μ_2= 1, σ_2= 0.1. In this case the bimodality size gradually decreases as the values of parameter "p" increases while keeping other parameters constant. For the same values of mixing proportion and combination of (μ_2, σ_2)= (1, 0.2) the results remain same. Figure 3 (B) describes the bimodality size for different values of p= (0.1, 0.2, 0.3, ... , 0.5) and μ_2= 2, $\sigma_2 = 0.5$. Here, the bimodality size decreases by increasing the parameter, mixing proportion'p' while keeping other parameters (i.e. μ_2, σ_2) constant. For the same values of "p" and combination of (μ_2,σ_2)= (0.2, 0.9) the result remains unchanged.

Next, Figure 3 (C) identifies the bimodality size for the values of p= (0.1, 0.2, ... , 0.5) and μ_2= 3, $\sigma_2 = 0.8$. In this situation the bimodality size decreases by increasing the parameter "p" while keeping other parameters constant. Figure 3 (D) shows the bimodality size for the values of p= (0.1, 0.2, ... , 0.5) and μ_2= 8, $\sigma_2 = 0.9$.At this time the bimodality size decreases by increasing the parameter "p" while keeping other parameters constant. For the same values of "p"and combination of (μ_2, σ_2)= (6, 0.9) the result remains same as Figure 3 (D).

Effect of standard deviation "σ_2" in mixture of normals

In this subsection the parametric values of standard deviation "σ_2" are changed while the values of remaining μ_2 and "p" are kept as constant to evaluate the bimodality size.

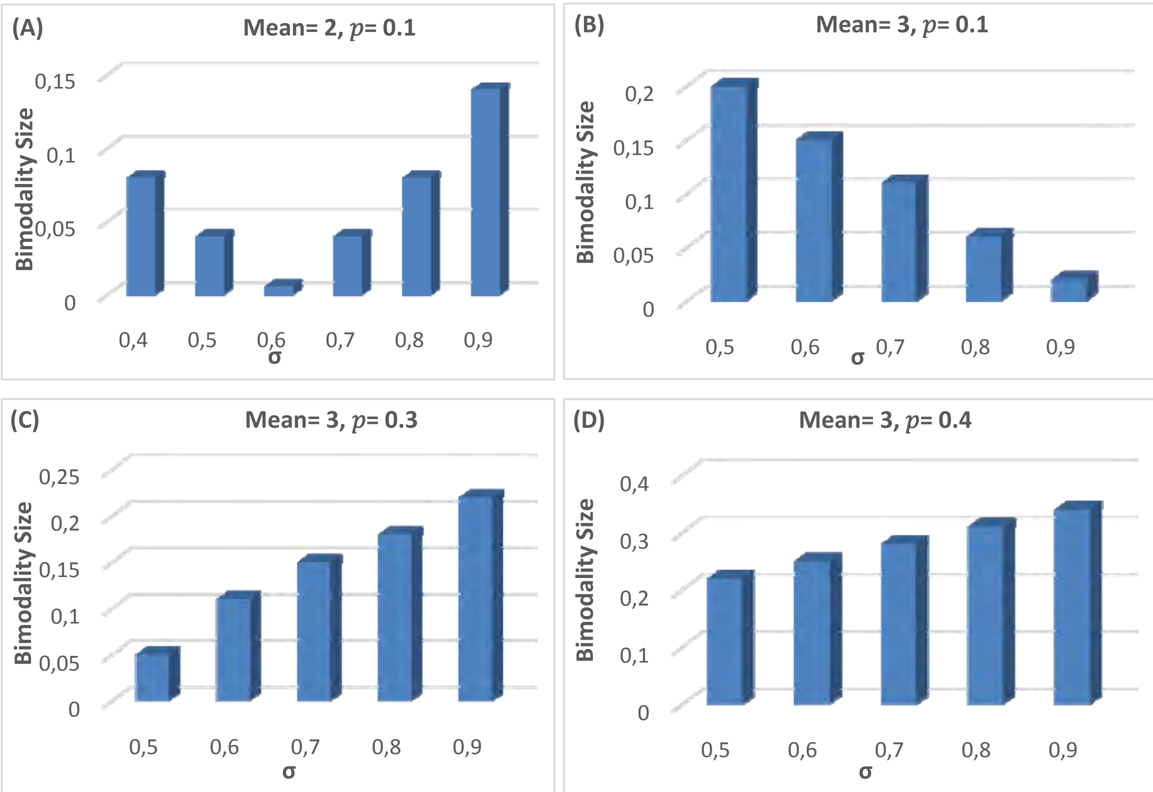

Figure 4. Size of the mixture of normal with various values of "σ_2"

Figure 4 (A) describes the bimodality size for different values of σ_2= (0.4, 0.5, 0.6, ... , 0.9) and fixed$p$= 0.1, μ_2 =2. In this case the bimodality size decreases for first three values of σ_2 then increases with minimum margin as the parametric value of σ_2 increases while keeping other parameter values constant. For the same values of σ_2 and combination of (μ_2, p)= (2, 0.2), (3, 0.2) the results remain same. Figure 4 (B) shows the bimodality size for different values of σ_2= (0.5, 0.6, ... , 0.9) and fixed p= 0.1, $\mu_2 = 3$. Here the bimodality size decreases fractionally by increasing the parameter "σ_2" while adjust other parameters (i.e.p, μ_2) constant.

Next Figure 4 (C) observes the bimodality size for the values of σ_2= (0.5, 0.6, ... , 0.9) and fixed p= 0.3, $\mu_2 = 3$. In this situation the bimodality size increases fractionally by increasing the parameter "σ_2" while keeping other parameters constant. At the same values of σ_2 and combination of (p, μ_2)= (0.2, 2) the result remains same. Figure 4 (D) identifies the bimodality size for the values of σ_2= (0.5, 0.6, ... , 0.9) and fixed$p$= *0.4*, $\mu_2 = 3$. Here the bimodality size increases slightly by increasing the parameterσ_2 while keeping other parameters constant. It means that the parameter "σ_2" has very low effect on the bimodality size as compare to other parameters. At the same values of σ_2 and combination of (p, μ_2)= (0.4, 2) the results remain same as have been observed for Figure 4 (C).

Conclusion

The analysis of this study determined that in mixture of normals, the parameter mean"μ_2" affected the bimodality size with the variation i.e. 7% to 10% approximately. For mixing probability i.e. 30%$\leq p \leq$ 50% the size decreases otherwise increases. Similarly, in the second case the parameter "p" also affected the size with the variation i.e. 2% to 5% approximately. The bimodality size decreases as the parameter mixing probability increases. In third case, the parameter standard deviation"σ_2" diverges the bimodality size fractionally about 0.3% to 0.5%. Mostly, the bimodality size decreases as the parameter mixing probability increases. So, it is investigated that the bimodality size changes often on the basis of two parameters of mixture of normals (i.e. mean and mixing proportion).

This study can be extended in future to use the other definite integrals for the same purpose. Also these integrals can be compared in case of bimodality.

References

Andrews, D. F., & Mallows, C. L. (1974). Scale mixtures of normal distributions. *Journal of the Royal Statistical Society. Series B (Methodological)*, 99-102.

Bertholon, H., Monfort, A., & Pegoraro, F. (2007). Pricing and inference with mixtures of conditionally normal processes.

Choonpradub, C., & McNeil, D. (2005). Can the box plot be improved. *Songklanakarin Journal of Science and Technology*, *27*(3), 649-657.

Day, N. E. (1969). Estimating the components of a mixture of normal distributions. *Biometrika*, *56*(3), 463-474.

Everitt, B. S. (1981). Bimodality and the nature of depression. *The British Journal of Psychiatry*, *138*(4), 336-339.

Ferguson, T. S. (1983). Bayesian density estimation by mixtures of normal distributions. In*Recent advances in statistics* (pp. 287-302).

Frankland, B. W., & Zumbo, B. D. (2002). Quantifying bimodality Part I: an easily implemented method using SPSS. *Journal of Modern Applied Statistical Methods*, *1*(1), 22.

Ghosal, S., & Van Der Vaart, A. W. (2001). Entropies and rates of convergence for maximum likelihood and Bayes estimation for mixtures of normal densities. *Annals of Statistics*, 1233-1263.

Haas, M., Mittnik, S., & Paolella, M. S. (2004). Mixed normal conditional heteroskedasticity. *Journal of Financial Econometrics*, *2*(2), 211-250.

Lindsay, B. G. (1995, January). Mixture models: theory, geometry and applications. In *NSF-CBMS regional conference series in probability and statistics* (pp. i-163). Institute of Mathematical Statistics and the American Statistical Association.

McLachlan, G. J., Peel, D., Basford, K. E., & Adams, P. (1999). The EMMIX software for the fitting of mixtures of normal and t-components. *Journal of Statistical Software*, *4*(2), 1-14.

Murphy, E. A. (1964). One cause? Many causes?: The argument from the bimodal distribution. *Journal of Chronic Diseases*, *17*(4), 301-324.

Robertson, C. A., & Fryer, J. G. (1969). Some descriptive properties of normal mixtures. *Scandinavian Actuarial Journal*, *1969*(3-4), 137-146.

Schilling, M. F., Watkins, A. E., & Watkins, W. (2002). Is human height bimodal?. *The American Statistician*, *56*(3), 223-229.

Shyu, C., & Ytreberg, F. M. (2008). Use of polynomial interpolation to reduce bias and uncertainty of free energy estimates via thermodynamic integration. *arXiv preprint arXiv:0809.0882*.

Sitek, G. (2016). The modes of a mixture of two normal distributions. *Silesian Journal of Pure and Applied Mathematics*, *6*(1), 59-67.

Suli, E.,&Mayers, D.F. (2003). An introduction to numerical analysis. *Cambridge UniversityPress, Cambridge, UK.*

Venkataraman, S. (1997). Value at risk for a mixture of normal distributions: the use of quasi-Bayesian estimation techniques. *Economic Perspectives-Federal Reserve Bank of Chicago*, *21*, 2-13.

West, M. (1987). On scale mixtures of normal distributions. *Biometrika, 74*(3), 646-648.