Alternative Chebyshev's Inequality

Mian Arif Shams Adnan[1], Silvia Irin Sharna[2]
[1,2]Bowling Green State University, Bowling Green, OH 43403

Abstract
Chebyshev's Inequality is being used to find what is the least percentage of observations fall within the specified sigma limits of the distribution. However, most of the distributions the industry experience with are asymmetric (or rarely symmetric) ones for which standard deviation may not be the proper measure of dispersion. Moreover, the presence of outliers may have huge amplification impact on standard deviation. In this paper a new form of Chebyshev's Inequality has been developed which is not based on sigma limits.

Key Words: Mean Deviation Measured from Mean.

1. Introduction

Chebyshev's Inequality helps us find the least percentage of observations that fall within the specified sigma limits of a symmetric or asymmetric distribution (Devore, 2016 and Mood, Graybil and Bose, 1974). Unfortunately sigma is not a proper variation measure when the distribution is asymmetric. Besides, for the presence of outlier(s), variance is seriously amplified. Mean is not as amplified as variance. We need a form of the Chebyshev's Inequality which is consisting of first order moment and one dimensional dispersion measure. For the presence of outlier, Means and Mean Deviation are not as affected as standard deviation.

2. Mean Deviation Version of Chebyshev's Inequality

The Mean Deviation version of Chebyshev's Inequality has been derived in the following theorem. It demonstrates that at least (1 - $\frac{25}{16r^2}$)100% observations fall within r Mean Deviation from Mean.

Theorem 2.1: If X be a random variable with location parameter μ and r be a positive number >0, then prove that

$$P(|X - \mu| > rE|x - \mu|) < \frac{25}{16r^2}.$$

Proof:

$$P(g(X) > k) < \frac{E(g(X))}{k}$$

$$\text{or, } P([X - \mu]^2 > k^2) < \frac{E[X - \mu]^2}{k^2}$$

$$\text{or, } P(|X - \mu| > rE|x - \mu|) < \frac{\left(\frac{5}{4}E|x-\mu|\right)^2}{(r\,E|x-\mu|)^2}$$

where, $E[X - \mu]^2 = \left(\frac{5}{4}E|x - \mu|\right)^2$, (Bluman, 2017).

or, $P(|X - \mu| > rE|x - \mu|) < \frac{25}{16r^2}$. This completes the prof■

2.1 Maximum Percentage Observations Fall Out for Different Number of Mean Deviations

If strictly, $r > 1.25$, then for different values of r, different percentage of observations fall out of 2 Mean Deviation Limit. As for example, for r = 2, 39% observations fall out of 2 Mean Deviation Limits for the following equation.

$$r = 2: P(|X - \mu| > 2E|x-\mu|) < \frac{25}{16(2)^2}100\% = 39\%.$$

Here, for r = 3: $P(|X - \mu| > 3E|x-\mu|) < \frac{25}{16(3)^2}100\% = 17\%$, r = 5: $P(|X - \mu| > 4E|x-\mu|) < \frac{25}{16(5)^2}100\% = 6\%$, r = 10: $P(|X - \mu| > 10E|x-\mu|) < \frac{25}{16(10)^2}100\% = 1.6\%$, r = 20: $P(|X - \mu| > 20E|x-\mu|) < \frac{25}{16(20)^2}100\% = 0.004\%$.

For r = 3.75, $P(|X - \mu| > 3.75E|x-\mu|) < \frac{25}{16(3.75)^2}100\% = 11\%$ and for r = 2.5: $P(|X - \mu| > 2.5E|x-\mu|) < \frac{25}{16(2.5)^2}100\% = 25\%$. That is, around 75% observations fall within 2.5 Mean Deviation Limits from Mean. Moreover, 89% observations fall within 3.75 Mean Deviation Limits.

Moreover, there is a relationship between mean deviation and standard deviation (Bluman, 2017). The relation is as below.

$$2.5 \text{ Mean Deviation} = 2.5\frac{4}{5} \text{ Standard Deviation} = 2 \text{ Standard Deviation.}$$

$$3.75 \text{ Mean Deviation} = 3.75\frac{4}{5} \text{ Standard Deviation} = 3 \text{ Standard Deviation.}$$

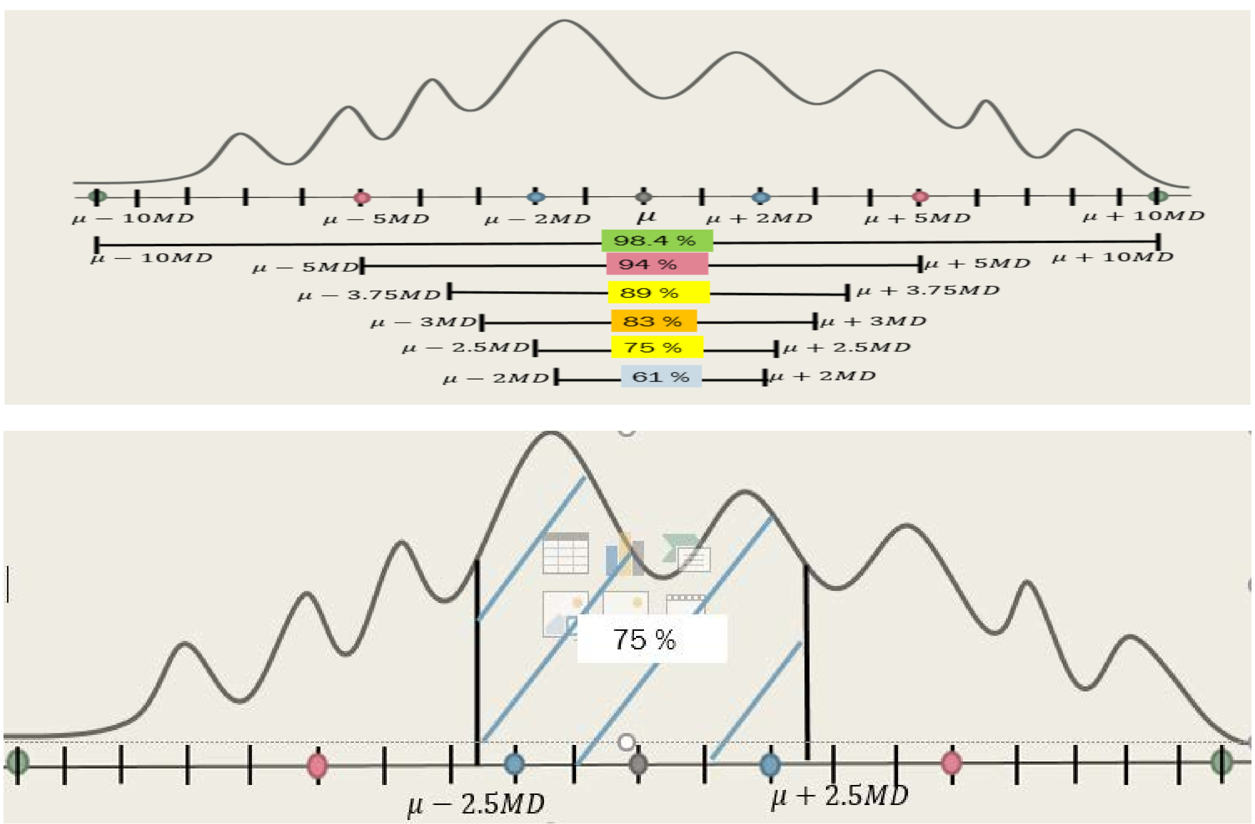

Figure 2: Comparison of Various Least Percentage of Observations Fall Within Mean Deviation Limits

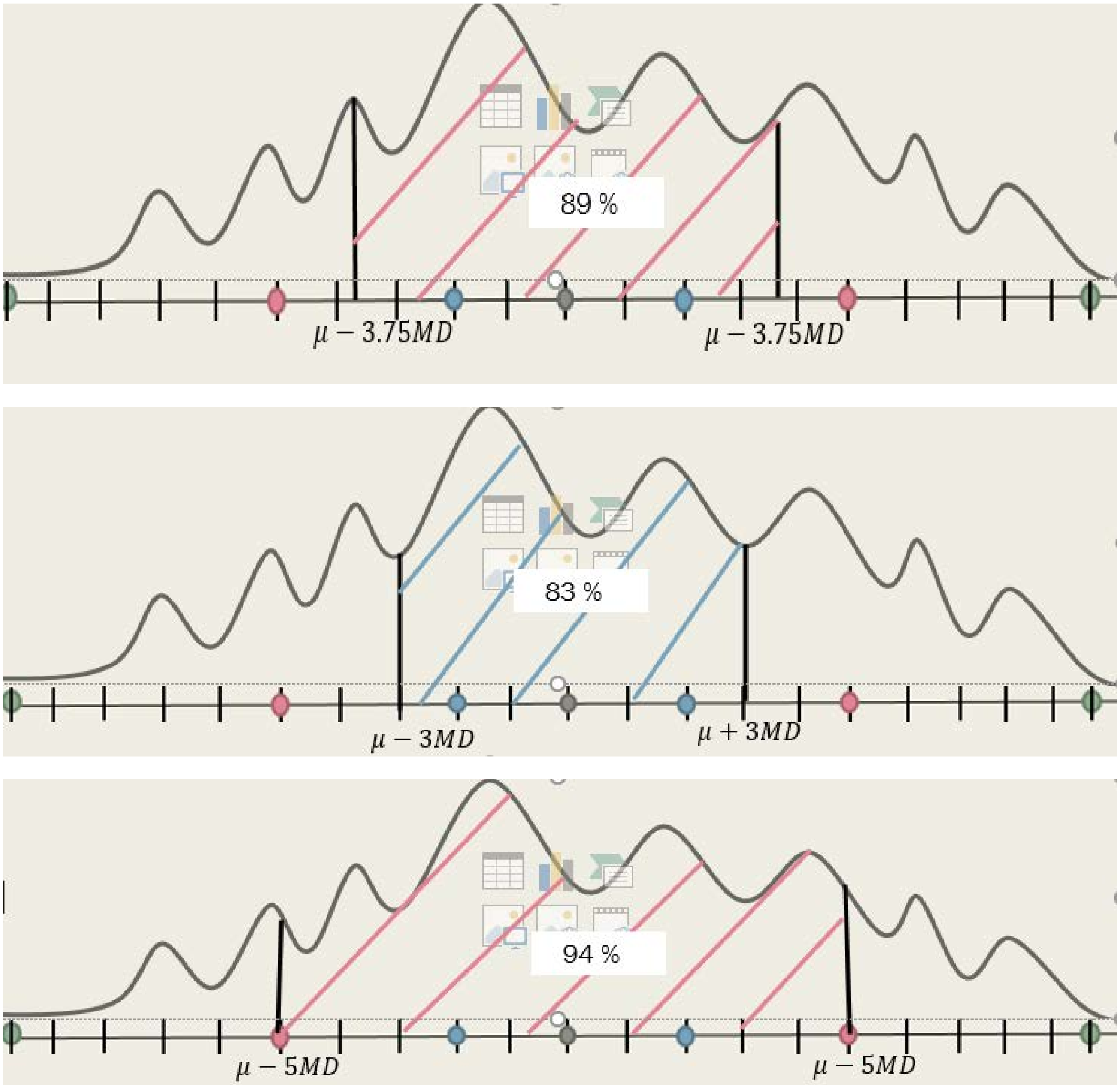

Figure 3: Comparison of Various Least Percentage of Observations Fall Within Mean Deviation Limits

Conclusion

Here, an alternative form of Chebyshev's Inequality has been developed which is based on mean deviation from mean limits. The future scope of the paper is to find the Chebyshev's Inequality using median and mean deviation from median.

Reference

Bluman, A. G. (2017). Elementary Statistics: A Step by Step Approach. McGraw-Hill.
Devore, J. L. (2016). Probability and Statistics for Engineering and the Sciences. Cole Publishing.
Mood, A. M., Graybill, F A. and Bose, D. C. (1974). Introduction to the Theory of Statistics. McGraw-Hill.