# Average deviation for measuring variation in data in small samples (n < 5)

Wenfa Ng

Department of Chemical and Biomolecular Engineering, National University of Singapore

Email: ngwenfa771@hotmail.com

## Abstract

Good control of experiment variability is critical to experiment success, and many methods are available for quantifying variation in data. Popular methods for measuring variability in data typically uses a statistical distribution such as a standard normal distribution, but these distributions are designed for large sample size with n > 30. However, experiments typically generate less than 5 replicates (n < 5). Thus, the key requirement for the use of standard normal distribution is not satisfied, which bring forth the need for the development of alternative ways of quantifying the variation in collected data for small sample size. This abstract describes a new statistic, average deviation, that aims to quantify the variation of repeated measurements of a variable. By taking an average of the sum of the differences between the mean and all measurements, average deviation provides a better representation of the variation in data around a mean, while capturing the impact of significant deviation from the mean by individual measurement. However, division of the sum of deviation of all measurements from the mean by the sample size meant that the presence of outlier measurement may not be fully represented by the calculated average deviation. Thus, the new statistic is better used with a small sample size of less than 5, which helps reduce the extent in which an outlier's influence on the average deviation would be diluted. In summary, for small sample size, average deviation better represents the deviation between each measurement and the mean compared to statistical distribution-based approaches such as standard error. However, desire to not dilute the impact of outlier measurement on the calculated average deviation meant that the new statistic is only suitable for sample size less than 5.

*Keywords:* average deviation, average, mean, statistical distribution, standard error, standard normal distribution, data variation, sample size, measurement,

*Subject areas:* statistics,

**Conflicts of interest**

The author declares no conflicts of interest.

**Author's contribution**

The author conceived the idea and developed the statistic. He wrote the manuscript.

**Funding**

No funding was used in this work.