

# Median absolute deviation

In statistics, the **median absolute deviation** (**MAD**) is a robust measure of the variability of a univariate sample of quantitative data. It can also refer to the population parameter that is estimated by the MAD calculated from a sample.

For a univariate data set  $X_1, X_2, \dots, X_n$ , the MAD is defined as the median of the absolute deviations from the data's median  $\tilde{X} = \text{median}(X)$ :

$$\text{MAD} = \text{median}(|X_i - \tilde{X}|)$$

that is, starting with the residuals (deviations) from the data's median, the MAD is the median of their absolute values.

## Contents

- Example
- Uses
- Relation to standard deviation
- Geometric median absolute deviation
- The population MAD
- See also
- Notes
- References

## Example

Consider the data (1, 1, 2, **2**, 4, 6, 9). It has a median value of 2. The absolute deviations about 2 are (1, 1, 0, 0, 2, 4, 7) which in turn have a median value of 1 (because the sorted absolute deviations are (0, 0, 1, **1**, 2, 4, 7)). So the median absolute deviation for this data is 1.

## Uses

The median absolute deviation is a measure of statistical dispersion. Moreover, the MAD is a robust statistic, being more resilient to outliers in a data set than the standard deviation. In the standard deviation, the distances from the mean are squared, so large deviations are weighted more heavily, and thus outliers can heavily influence it. In the MAD, the deviations of a small number of outliers are irrelevant.

Because the MAD is a more robust estimator of scale than the sample variance or standard deviation, it works better with distributions without a mean or variance, such as the Cauchy distribution.

## Relation to standard deviation

The MAD may be used similarly to how one would use the deviation for the average. In order to use the MAD as a consistent estimator for the estimation of the standard deviation  $\sigma$ , one takes

$$\hat{\sigma} = k \cdot \text{MAD},$$

where ***k*** is a constant scale factor, which depends on the distribution.<sup>[1]</sup>

For normally distributed data ***k*** is taken to be:

***k*** = 1/ (Φ<sup>−1</sup>(3/4)) ≈ 1.4826,

i.e., the reciprocal of the quantile function Φ<sup>−1</sup> (also known as the inverse of the cumulative distribution function) for the standard normal distribution ***Z*** = (***X*** − ***μ***)/***σ***.<sup>[2][3]</sup> The argument 3/4 is such that ±**MAD** covers 50% (between 1/4 and 3/4) of the standard normal cumulative distribution function, i.e.:

$\frac{1}{2} = P(|X - \mu| \leq \text{MAD}) = P\left(\left|\frac{X - \mu}{\sigma}\right| \leq \frac{\text{MAD}}{\sigma}\right) = P\left(|Z| \leq \frac{\text{MAD}}{\sigma}\right).$

Therefore, we must have that:

Φ (MAD /***σ***) − Φ (− MAD /***σ***) = 1/2.

Noticing that:

Φ (− MAD /***σ***) = 1 − Φ (MAD /***σ***)

we have that **MAD** /***σ*** = Φ<sup>−1</sup> (3/4) = 0.67449 from which we obtain the scale factor ***k*** = 1/Φ<sup>−1</sup>(3/4) = 1.4826.

Another way of establishing the relationship is noting that MAD equals the half-normal distribution median:

**MAD** = ***σ***√2erf<sup>−1</sup>(1/2) ≈ 0.67449***σ***.

This form is used in, e.g., the probable error.

Geometric median absolute deviation

Similarly to how the median generalizes to the geometric median in multivariate data, a geometric MAD can be constructed that generalizes the MAD. Given a 2 dimensional paired set of data (*X*<sub>1</sub>,*Y*<sub>1</sub>), (*X*<sub>2</sub>,*Y*<sub>2</sub>),..., (*X*<sub>*n*</sub>,*Y*<sub>*n*</sub>) and a suitably calculated geometric median (***X̂***,***Ŷ***), the geometric median absolute deviation is given by:

**MAD** = ( **median**(|*X*<sub>*i*</sub> − ***X̂***|)<sup>2</sup> + **median**(|*Y*<sub>*i*</sub> − ***Ŷ***|)<sup>2</sup> )<sup>1/2</sup>

This gives the identical result as the univariate MAD in 1 dimension and extends easily to higher dimensions. In the case of complex values (*X*+i*Y*), the relation of MAD to the standard deviation is unchanged for normally distributed data.

The population MAD

The population MAD is defined analogously to the sample MAD, but is based on the complete distribution rather than on a sample. For a symmetric distribution with zero mean, the population MAD is the 75th percentile of the distribution.

Unlike the variance, which may be infinite or undefined, the population MAD is always a finite number. For example, the standard Cauchy distribution has undefined variance, but its MAD is 1.

The earliest known mention of the concept of the MAD occurred in 1816, in a paper by Carl Friedrich Gauss on the determination of the accuracy of numerical observations.<sup>[4][5]</sup>

## See also

---

- Deviation (statistics)
- Interquartile range
- Probable error
- Robust measures of scale
- Relative mean absolute difference
- Average absolute deviation
- Least absolute deviations

## Notes

---

- Rousseeuw, P. J.; Croux, C. (1993). "Alternatives to the median absolute deviation". *Journal of the American Statistical Association*. **88** (424): 1273–1283. doi:10.1080/01621459.1993.10476408 (https://doi.org/10.1080%2F01621459.1993.10476408).
- Ruppert, D. (2010). *Statistics and Data Analysis for Financial Engineering* (https://books.google.co.uk/books?id=i2bD50PblikC&pg=PA118#v=onepage). Springer. p. 118. ISBN 9781441977878. Retrieved 2015-08-27.
- Leys, C.; et al. (2013). "Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median" (https://dipot.ulb.ac.be/dspace/bitstream/2013/139499/1/Leys\_MAD\_final-libre.pdf) (PDF). *Journal of Experimental Social Psychology*. **49** (4): 764–766. doi:10.1016/j.jesp.2013.03.013 (https://doi.org/10.1016%2Fj.jesp.2013.03.013).
- Gauss, Carl Friedrich (1816). "Bestimmung der Genauigkeit der Beobachtungen". *Zeitschrift für Astronomie und Verwandte Wissenschaften*. **1**: 187–197.
- Walker, Helen (1931). *Studies in the History of the Statistical Method*. Baltimore, MD: Williams & Wilkins Co. pp. 24–25.

## References

---

- Hoaglin, David C.; Frederick Mosteller; John W. Tukey (1983). *Understanding Robust and Exploratory Data Analysis*. John Wiley & Sons. pp. 404–414. ISBN 978-0-471-09777-8.
- Russell, Roberta S.; Bernard W. Taylor III. (2006). *Operations Management*. John Wiley & Sons. pp. 497–498. ISBN 978-0-471-69209-6.
- Venables, W.N.; B.D. Ripley (1999). *Modern Applied Statistics with S-PLUS*. Springer. p. 128. ISBN 978-0-387-98825-2.

---

Retrieved from "https://en.wikipedia.org/w/index.php?title=Median\_absolute\_deviation&oldid=905525974"

---

**This page was last edited on 9 July 2019, at 17:54 (UTC).**

Text is available under the Creative Commons Attribution-ShareAlike License; additional terms may apply. By using this site, you agree to the Terms of Use and Privacy Policy. Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc., a non-profit organization.