

How can the Median Absolute Deviation be calculated in R?

Authored by
stats writer

April 23, 2024

RECOMMENDED CITATION

stats writer (2024). *How can the Median Absolute Deviation be calculated in R?*.
PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=138434>

The Median Absolute Deviation (MAD) is a measure of variability that is less affected by extreme values compared to other measures such as standard deviation. In R, the MAD can be calculated by first finding the median of the data set. Then, each data point is subtracted from the median and the absolute value of the difference is taken. These absolute values are then sorted and the median is calculated again. This second median value is the MAD, representing the average distance of each data point from the median. This calculation can be easily completed using built-in functions in R, such as the `median()` and `abs()` functions.

Calculate Median Absolute Deviation in R

The median absolute deviation measures the spread of in a dataset.

It's a particularly useful metric because it's less affected by outliers than other like standard deviation and variance.

The formula to calculate median absolute deviation, often abbreviated MAD, is as follows:

$$\text{MAD} = \text{median}(|x_i - x_m|)$$

where:

x_i : The i th value in the dataset
 x_m : The median value in the dataset

The following examples shows how to calculate the median absolute deviation in R by using the built-in

mad() function.

Example 1: Calculate MAD for a Vector

The following code shows how to calculate the median absolute deviation for a single vector in R:

```
#define data
```

```
data <- c(1, 4, 4, 7, 12, 13, 16, 19, 22, 24)
```

```
#calculate MAD
```

```
mad(data)
```

```
11.1195
```

The median absolute deviation for the dataset turns out to be 11.1195.

Example 2: Calculate MAD for a Column in a Data Frame

The following code shows how to calculate MAD for a single column in a data frame:

```
#define data
```

```
data <- data.frame(x = c(1, 4, 4, 6, 7, 8, 12),
```

```
y = c(3, 4, 6, 8, 8, 9, 19),
```

```
z = c(2, 2, 2, 3, 5, 8, 11))
```

```
#calculate MAD for column y in data frame  
mad(data$y)
```

2.9652

The median absolute deviation for column y turns out to be 2.9652.

Example 3: Calculate MAD for Multiple Columns in a Data Frame

The following code shows how to calculate MAD for multiple columns in a data frame by using the `sapply()` function:

```
#define data  
data <- data.frame(x = c(1, 4, 4, 6, 7, 8, 12),  
y = c(3, 4, 6, 8, 8, 9, 19),  
z = c(2, 2, 2, 3, 5, 8, 11))
```

```
#calculate MAD for all columns in data frame  
sapply(data, mad)
```

```
x y z  
2.9652 2.9652 1.4826
```

The median absolute deviation is 2.9652 for column x,

2.9652 for column y, and 1.4826 for column z.

ARABPSYCHOLOGY.COM