

Project 2—CSC 433

Function Name: MAD Function

Author: Sarah Cummings

Description: The MAD (median absolute deviation) function in SAS finds the univariate scaled median absolute deviation of a data set. This value is a robust measure of variability of our data. 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 datasets median, given as follows:

$$MAD = \text{median}_i (|X_i - \text{median}_j(X_j)|)$$

Note that the MAD function treats the input data matrix (x) as univariate data by appending each row to the previous row to make a single row vector with elements $X_{11}, \dots, X_{1p}, X_{21}, \dots, X_{2p}, \dots, X_{n1}, \dots, X_{np}$.

Input Parameter: The function MAD (x , < method >) takes input x, an n x p input data matrix and an optional method string. Our choices for this method input are as follows:

“MAD”, the default which computes the median absolute deviation

“NMAD”, which computes the normalized version of the MAD:

$$MAD_n = b * \text{med}_i^n |x_i - \text{med}_j^n x_j|$$

where $b=1$ is the unscaled default and $b=1.4826$ is used for scaling, as is consistent with Gaussian distribution.

“SN”, which is an follows:

alternative to MAD given as

$$S_n = c_n * \text{med}_i \text{med}_{j \neq i} |x_i - x_j|$$

“QN”, which MAD given

is another alternative to as follows:

$$Q_n = d_n * \{ |x_i - x_j|; i < j \}_{(k)} \quad \text{with} \quad k \approx \binom{n}{2} / 4$$

Return Value: Numeric value as described above

SAS Test Script:

```
proc iml;
x = {14, 25, 42, 67, 89};
mad = mad(x, "MAD");
mad2= mad(x, "NMAD");
mad3= mad(x, "SN");
```

Output:

mad	mad2	mad3	mad4
25	37.065055	40.280065	41.256239

```
mad4= mad(x, "QN");  
print mad mad2 mad3 mad4;
```