

NAG Library Routine Document

G07DAF

Note: before using this routine, please read the Users' Note for your implementation to check the interpretation of *bold italicised* terms and other implementation-dependent details.

1 Purpose

G07DAF finds the median, median absolute deviation, and a robust estimate of the standard deviation for a set of ungrouped data.

2 Specification

```
SUBROUTINE G07DAF (N, X, Y, XME, XMD, XSD, IFAIL)
  INTEGER          N, IFAIL
  REAL (KIND=nag_wp) X(N), Y(N), XME, XMD, XSD
```

3 Description

The data consists of a sample of size n , denoted by x_1, x_2, \dots, x_n , drawn from a random variable X .

G07DAF first computes the median,

$$\theta_{\text{med}} = \text{med}_i \{x_i\},$$

and from this the median absolute deviation can be computed,

$$\sigma_{\text{med}} = \text{med}_i \{|x_i - \theta_{\text{med}}|\}.$$

Finally, a robust estimate of the standard deviation is computed,

$$\sigma'_{\text{med}} = \sigma_{\text{med}} / \Phi^{-1}(0.75)$$

where $\Phi^{-1}(0.75)$ is the value of the inverse standard Normal function at the point 0.75.

G07DAF is based upon subroutine LTMDVV within the ROBETH library, see Marazzi (1987).

4 References

Huber P J (1981) *Robust Statistics* Wiley

Marazzi A (1987) Subroutines for robust estimation of location and scale in ROBETH *Cah. Rech. Doc. IUMSP, No. 3 ROB 1* Institut Universitaire de Médecine Sociale et Préventive, Lausanne

5 Parameters

- | | | |
|----|---|---------------|
| 1: | N – INTEGER | <i>Input</i> |
| | <i>On entry:</i> n , the number of observations. | |
| | <i>Constraint:</i> $N > 1$. | |
| 2: | X(N) – REAL (KIND=nag_wp) array | <i>Input</i> |
| | <i>On entry:</i> the vector of observations, x_1, x_2, \dots, x_n . | |
| 3: | Y(N) – REAL (KIND=nag_wp) array | <i>Output</i> |
| | <i>On exit:</i> the observations sorted into ascending order. | |

- 4: XME – REAL (KIND=nag_wp) Output
On exit: the median, θ_{med} .
- 5: XMD – REAL (KIND=nag_wp) Output
On exit: the median absolute deviation, σ_{med} .
- 6: XSD – REAL (KIND=nag_wp) Output
On exit: the robust estimate of the standard deviation, σ'_{med} .
- 7: IFAIL – INTEGER Input/Output
On entry: IFAIL must be set to 0, -1 or 1. If you are unfamiliar with this parameter you should refer to Section 3.3 in the Essential Introduction for details.
- For environments where it might be inappropriate to halt program execution when an error is detected, the value -1 or 1 is recommended. If the output of error messages is undesirable, then the value 1 is recommended. Otherwise, if you are not familiar with this parameter, the recommended value is 0. **When the value -1 or 1 is used it is essential to test the value of IFAIL on exit.**
- On exit:* IFAIL = 0 unless the routine detects an error or a warning has been flagged (see Section 6).

6 Error Indicators and Warnings

If on entry IFAIL = 0 or -1, explanatory error messages are output on the current error message unit (as defined by X04AAF).

Errors or warnings detected by the routine:

IFAIL = 1

On entry, $N \leq 1$.

IFAIL = -99

An unexpected error has been triggered by this routine. Please contact NAG.

See Section 3.8 in the Essential Introduction for further information.

IFAIL = -399

Your licence key may have expired or may not have been installed correctly.

See Section 3.7 in the Essential Introduction for further information.

IFAIL = -999

Dynamic memory allocation failed.

See Section 3.6 in the Essential Introduction for further information.

7 Accuracy

The computations are believed to be stable.

8 Parallelism and Performance

G07DAF is threaded by NAG for parallel execution in multithreaded implementations of the NAG Library.

G07DAF makes calls to BLAS and/or LAPACK routines, which may be threaded within the vendor library used by this implementation. Consult the documentation for the vendor library for further information.

Please consult the X06 Chapter Introduction for information on how to control and interrogate the OpenMP environment used within this routine. Please also consult the Users' Note for your implementation for any additional implementation-specific information.

9 Further Comments

Unless otherwise stated in the Users' Note, the routine may be called with the same actual array supplied for parameters X and Y, in which case the sorted data values will overwrite the original contents of X. However this is not standard Fortran, and may not work on all systems.

10 Example

The following program reads in a set of data consisting of eleven observations of a variable X . The median, median absolute deviation and a robust estimate of the standard deviation are calculated and printed along with the sorted data in output array Y.

10.1 Program Text

```

Program g07dafa
!      G07DAF Example Program Text
!
!      Mark 25 Release. NAG Copyright 2014.
!
!      .. Use Statements ..
Use nag_library, Only: g07daf, nag_wp
!      .. Implicit None Statement ..
Implicit None
!      .. Parameters ..
Integer, Parameter      :: nin = 5, nout = 6
!      .. Local Scalars ..
Real (Kind=nag_wp)      :: xmd, xme, xsd
Integer                  :: ifail, n
!      .. Local Arrays ..
Real (Kind=nag_wp), Allocatable :: x(:), y(:)
!      .. Executable Statements ..
Write (nout,*) 'G07DAF Example Program Results'
Write (nout,*)

!      Skip heading in data file
Read (nin,*)

!      Read in problem size
Read (nin,*) n

Allocate (x(n),y(n))

!      Read in data
Read (nin,*) x(1:n)

!      Calculate robust summaries
ifail = 0
Call g07daf(n,x,y,xme,xmd,xsd,ifail)

!      Display results
Write (nout,*) 'Original Data'
Write (nout,99999) x(1:n)
Write (nout,*)
Write (nout,*) 'Sorted Data'
Write (nout,99999) y(1:n)
Write (nout,*)
Write (nout,99998) 'Median'

```

```
Write (nout,99998) 'Median absolute deviation          = ', xmd
Write (nout,99998) 'Robust estimate standard deviation = ', xsd

99999 Format (1X,11F7.3)
99998 Format (1X,A,F6.3)
End Program g07dafa
```

10.2 Program Data

```
G07DAF Example Program Data
11      : N, NUMBER OF OBSERVATIONS
13.0 11.0 16.0 5.0 3.0 18.0 9.0 8.0 6.0 27.0 7.0 : X, OBSERVATIONS
```

10.3 Program Results

G07DAF Example Program Results

Original Data

13.000 11.000 16.000 5.000 3.000 18.000 9.000 8.000 6.000 27.000 7.000

Sorted Data

3.000 5.000 6.000 7.000 8.000 9.000 11.000 13.000 16.000 18.000 27.000

Median = 9.000
Median absolute deviation = 4.000
Robust estimate standard deviation = 5.930
