# NAG Library Routine Document

# G01AAF

**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

G01AAF calculates the mean, standard deviation, coefficients of skewness and kurtosis, and the maximum and minimum values for a set of ungrouped data. Weighting may be used.

## 2 Specification

```
SUBROUTINE G01AAF (N, X, IWT, WT, XMEAN, S2, S3, S4, XMIN, XMAX, WTSUM,          &
                   IFAIL)

INTEGER             N, IWT, IFAIL
REAL (KIND=nag_wp) X(N), WT(N), XMEAN, S2, S3, S4, XMIN, XMAX, WTSUM
```

## 3 Description

The data consist of a single sample of $n$ observations, denoted by $x_i$, with corresponding weights, $w_i$, for $i = 1, 2, \ldots, n$.

If no specific weighting is required, then each $w_i$ is set to 1.

The quantities computed are:

(a) The sum of the weights

$$W = \sum_{i=1}^{n} w_i.$$

(b) Mean

$$\bar{x} = \frac{\sum_{i=1}^{n} w_i x_i}{W}.$$

(c) Standard deviation

$$s_2 = \sqrt{\frac{\sum_{i=1}^{n} w_i (x_i - \bar{x})^2}{d}}, \qquad \text{where} \qquad d = W - \frac{\sum_{i=1}^{n} w_i^2}{W}.$$

(d) Coefficient of skewness

$$s_3 = \frac{\sum_{i=1}^{n} w_i (x_i - \bar{x})^3}{d \times s_2^3}.$$

(e) Coefficient of kurtosis

$$s_4 = \frac{\sum_{i=1}^{n} w_i (x_i - \bar{x})^4}{d \times s_2^4} - 3.$$

(f) Maximum and minimum elements of the sample.

(g) The number of observations for which $w_i > 0$, i.e., the number of **valid** observations. Suppose $m$ observations are valid, then the quantities in (c), (d) and (e) will be computed if $m \geq 2$, and will be based on $m - 1$ degrees of freedom. The other quantities are evaluated provided $m \geq 1$.

## 4 References

None.

## 5 Parameters

1: N – INTEGER *Input*

*On entry*: $n$, the number of observations.

*Constraint*: N $\geq 1$.

2: X(N) – REAL (KIND=nag_wp) array *Input*

*On entry*: the sample observations, $x_i$, for $i = 1, 2, \ldots, n$.

3: IWT – INTEGER *Input/Output*

*On entry*: indicates whether weights are to be supplied by you or not. In the latter case, the weights will be assumed equal and assigned the value 1.0 in the routine.

IWT $= 0$
Indicates no user-supplied weights.

IWT $= 1$
Indicates user-supplied weights are required, and they will be supplied in the array WT.

*On exit*: IWT is used to indicate the number of valid observations, $m$; see (g) in Section 3 above.

4: WT(N) – REAL (KIND=nag_wp) array *Input/Output*

*On entry*: if IWT $= 1$, the elements of WT must contain the weights associated with the observations, $w_i$, for $i = 1, 2, \ldots, n$.

If IWT $= 0$, the elements of WT need not be set.

*On exit*: if IWT $= 1$, the elements of WT are unchanged.

If IWT $= 0$, each element of WT will be assigned the value 1.0.

5: XMEAN – REAL (KIND=nag_wp) *Output*

*On exit*: the mean, $\bar{x}$.

6: S2 – REAL (KIND=nag_wp) *Output*

*On exit*: the standard deviation, $s_2$.

7: S3 – REAL (KIND=nag_wp) *Output*

*On exit*: the coefficient of skewness, $s_3$.

8: S4 – REAL (KIND=nag_wp) *Output*

*On exit*: the coefficient of kurtosis, $s_4$.

9: XMIN – REAL (KIND=nag_wp) *Output*

*On exit*: the smallest value in the sample.

10: XMAX – REAL (KIND=nag_wp) *Output*

*On exit*: the largest value in the sample.

11: WTSUM – REAL (KIND=nag_wp) *Output*

*On exit*: the sum of the weights in the array WT, that is $\sum_{i=1}^{n} w_i$. This will be N if IWT was 0 on entry.

12: 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 $< 1$.

IFAIL $= 2$

The number of valid cases, $m$, is 1. In this case, standard deviation and coefficients of skewness and of kurtosis cannot be calculated.

IFAIL $= 3$

Either the number of valid cases is 0, or at least one weight is negative.

## 7 Accuracy

The method used is believed to be stable.

## 8 Further Comments

The time taken by G01AAF is approximately proportional to $n$.

## 9 Example

This example summarises an (optionally weighted) dataset and displays the results.

## 9.1 Program Text

```
    Program g01aafe

!     G01AAF Example Program Text

!     Mark 24 Release. NAG Copyright 2012.

!     .. Use Statements ..
      Use nag_library, Only: g01aaf, nag_wp
!     .. Implicit None Statement ..
      Implicit None
!     .. Parameters ..
      Integer, Parameter               :: nin = 5, nout = 6
!     .. Local Scalars ..
      Real (Kind=nag_wp)               :: s2, s3, s4, wtsum, xmax, xmean, xmin
      Integer                          :: ifail, iwt, n
!     .. Local Arrays ..
      Real (Kind=nag_wp), Allocatable  :: wt(:), wtin(:), x(:)
!     .. Executable Statements ..
      Write (nout,*) 'G01AAF Example Program Results'
      Write (nout,*)

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

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

      Allocate (wt(n),wtin(n),x(n))

!     Read in data
      Read (nin,*) x(1:n)
      If (iwt==1) Then
        Read (nin,*) wtin(1:n)
        wt(1:n) = wtin(1:n)
      End If

!     Display data
      Write (nout,99999) 'Number of cases ', n
      Write (nout,*) 'Data as input -'
      Write (nout,99998) x(1:n)
      If (iwt==1) Then
        Write (nout,*) 'Weights as input -'
        Write (nout,99998) wtin(1:n)
      End If
      Write (nout,*)

!     Calculate summary statistics
      ifail = -1
      Call g01aaf(n,x,iwt,wt,xmean,s2,s3,s4,xmin,xmax,wtsum,ifail)
      If (ifail/=0) Then
        If (ifail/=2) Then
          Go To 100
        End If
      End If

!     Display results
      Write (nout,99999) 'No. of valid cases    ', iwt
      Write (nout,99997) 'Mean          ', xmean
      Write (nout,99997) 'Minimum       ', xmin
      Write (nout,99997) 'Maximum       ', xmax
      Write (nout,99997) 'Sum of weights', wtsum

      If (ifail==0) Then
        Write (nout,99997) 'Std devn      ', s2
        Write (nout,99997) 'Skewness      ', s3
        Write (nout,99997) 'Kurtosis      ', s4
      Else
        Write (nout,*) 'Std devn and coeffts of skewness'
        Write (nout,*) 'and kurtosis not defined'
```

```
      End If

100   Continue

99999 Format (1X,A,I5)
99998 Format (1X,5F12.1)
99997 Format (1X,A,F13.1)
    End Program g01aafe
```

## 9.2 Program Data

```
G01AAF Example Program Data
24 0
193.0  216.0  112.0  161.0  92.0   140.0  38.0   33.0   279.0  249.0
473.0  339.0  60.0   130.0  20.0   50.0   257.0  284.0  447.0  52.0
67.0   61.0   150.0  2200.0
```

## 9.3 Program Results

```
 G01AAF Example Program Results

 Number of cases    24
 Data as input -
        193.0       216.0       112.0       161.0        92.0
        140.0        38.0        33.0       279.0       249.0
        473.0       339.0        60.0       130.0        20.0
         50.0       257.0       284.0       447.0        52.0
         67.0        61.0       150.0      2200.0

 No. of valid cases       24
 Mean                  254.3
 Minimum                20.0
 Maximum              2200.0
 Sum of weights         24.0
 Std devn              433.5
 Skewness                3.9
 Kurtosis               14.7
```