NAG Fortran 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,
1                  IFAIL)
 INTEGER           N, IWT, IFAIL
 real              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

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

(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: the number of observations, n.

Constraint: N ≥ 1.

2: X(N) – ***real*** 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 the user 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 3(g) above.

4: WT(N) – ***real*** array *Input/Output*

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

If IWT $= 0$, then 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*** *Output*

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

6: S2 – ***real*** *Output*

On exit: the standard deviation, s_2.

7: S3 – ***real*** *Output*

On exit: the coefficient of skewness, s_3.

8: S4 – ***real*** *Output*

On exit: the coefficient of kurtosis, s_4.

9: XMIN – ***real*** *Output*

On exit: the smallest value in the sample.

10: XMAX – ***real*** *Output*

On exit: the largest value in the sample.

11: WTSUM – ***real*** *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. Users who are unfamiliar with this parameter should refer to Chapter P01 for details.

On exit: IFAIL = 0 unless the routine detects an error (see Section 6).

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, for users 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.**

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 the routine is approximately proportional to n.

9 Example

In the example program, NPROB determines the number of data sets to be analysed. For each analysis, a set of observations and, optionally, weights is read and printed. After calling the routine, the calculated quantities are printed. In the example, there is one set of data with 24 unweighted data values.

9.1 Program Text

Note: the listing of the example program presented below uses ***bold italicised*** terms to denote precision-dependent details. Please read the Users' Note for your implementation to check the interpretation of these terms. As explained in the Essential Introduction to this manual, the results produced may not be identical for all implementations.

```
*     G01AAF Example Program Text
*     Mark 14 Revised.  NAG Copyright 1989.
*     .. Parameters ..
      INTEGER          NMAX
      PARAMETER        (NMAX=30)
      INTEGER          NIN, NOUT
      PARAMETER        (NIN=5,NOUT=6)
*     .. Local Scalars ..
      real             S2, S3, S4, WTSUM, XBAR, XMAX, XMIN
      INTEGER          I, IFAIL, IWT, J, N, NPROB
*     .. Local Arrays ..
      real             WT(NMAX), X(NMAX)
*     .. External Subroutines ..
```

```
      EXTERNAL         G01AAF
*     .. Executable Statements ..
      WRITE (NOUT,*) 'G01AAF Example Program Results'
*     Skip heading in data file
      READ (NIN,*)
      READ (NIN,*) NPROB
      DO 20 J = 1, NPROB
         READ (NIN,*) N, IWT
         WRITE (NOUT,*)
         WRITE (NOUT,99999) 'Problem ', J
         WRITE (NOUT,99999) 'Number of cases ', N
         IF (N.GE.1 .AND. N.LE.NMAX) THEN
            READ (NIN,*) (X(I),I=1,N)
            WRITE (NOUT,*) 'Data as input -'
            WRITE (NOUT,99998) (X(I),I=1,N)
            IF (IWT.EQ.1) THEN
               WRITE (NOUT,*) 'Weights as input -'
               READ (NIN,*) (WT(I),I=1,N)
               WRITE (NOUT,99998) (WT(I),I=1,N)
            END IF
            IFAIL = 1
*
            CALL G01AAF(N,X,IWT,WT,XBAR,S2,S3,S4,XMIN,XMAX,WTSUM,IFAIL)
*
            WRITE (NOUT,*)
            IF (IFAIL.EQ.0) THEN
               WRITE (NOUT,*) 'Successful call of G01AAF'
               WRITE (NOUT,99999) 'No. of valid cases    ', IWT
               WRITE (NOUT,99997) 'Mean          ', XBAR
               WRITE (NOUT,99997) 'Std devn      ', S2
               WRITE (NOUT,99997) 'Skewness      ', S3
               WRITE (NOUT,99997) 'Kurtosis      ', S4
               WRITE (NOUT,99997) 'Minimum       ', XMIN
               WRITE (NOUT,99997) 'Maximum       ', XMAX
               WRITE (NOUT,99997) 'Sum of weights', WTSUM
            ELSE
               WRITE (NOUT,*) 'Unsuccessful call of G01AAF'
               WRITE (NOUT,99999) 'IFAIL =', IFAIL
               IF (IFAIL.EQ.2) THEN
                  WRITE (NOUT,99999) 'No. of valid cases', IWT
                  WRITE (NOUT,99997) 'Mean          ', XBAR
                  WRITE (NOUT,99997) 'Minimum       ', XMIN
                  WRITE (NOUT,99997) 'Maximum       ', XMAX
                  WRITE (NOUT,99997) 'Sum of weights', WTSUM
                  WRITE (NOUT,*) 'Std devn and coeffts of skewness'
                  WRITE (NOUT,*) 'and kurtosis not defined'
               END IF
            END IF
         ELSE
            STOP
         END IF
   20 CONTINUE
      STOP
*
99999 FORMAT (1X,A,I5)
99998 FORMAT (1X,5F12.1)
99997 FORMAT (1X,A,F13.1)
      END
```

9.2 Program Data

```
G01AAF Example Program Data
 1
24 0
193.0  215.0  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

Problem     1
Number of cases    24
Data as input -
       193.0       215.0       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

Successful call of G01AAF
No. of valid cases       24
Mean                  254.2
Std devn              433.5
Skewness                3.9
Kurtosis               14.7
Minimum                20.0
Maximum              2200.0
Sum of weights         24.0
```