

# NAG Fortran Library Routine Document

## G01ABF

**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

G01ABF computes the means, standard deviations, corrected sums of squares and products, maximum and minimum values, and the product-moment correlation coefficient for two variables. Unequal weighting may be given.

### 2 Specification

```
SUBROUTINE G01ABF(N, X1, X2, IWT, WT, RES, IFAIL)
INTEGER          N, IWT, IFAIL
real           X1(N), X2(N), WT(N), RES(13)
```

### 3 Description

The data consist of two samples of  $n$  observations, denoted by  $x_i$ , and  $y_i$ , for  $i = 1, 2, \dots, n$ , with corresponding weights  $w_i$ , for  $i = 1, 2, \dots, n$ .

If no specific weighting is given, then each  $w_i$  is set to 1.0 in the routine.

The quantities calculated are:

- (a) The sum of weights,

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

- (b) The means,

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

- (c) The corrected sums of squares and products

$$c_{11} = \sum_{i=1}^n w_i (x_i - \bar{x})^2$$

$$c_{21} = c_{12} = \sum_{i=1}^n w_i (x_i - \bar{x})(y_i - \bar{y})$$

$$c_{22} = \sum_{i=1}^n w_i (y_i - \bar{y})^2.$$

- (d) The standard deviations

$$s_j = \sqrt{\frac{c_{jj}}{d}}, \quad \text{where } j = 1, 2 \quad \text{and} \quad d = W - \frac{\sum_{i=1}^n w_i^2}{W}.$$

- (e) The product-moment correlation coefficient

$$R = \frac{c_{12}}{\sqrt{c_{11}c_{22}}}.$$

- (f) The minimum and maximum elements in each of the two samples.

- (g) The number of pairs of observations,  $m$ , for which  $w_i > 0$ , i.e., the number of **valid** observations. The quantities in (d) and (e) above will only be computed if  $m \geq 2$ . All other items are computed if  $m \geq 1$ .

## 4 References

None.

## 5 Parameters

- 1: N – INTEGER *Input*  
*On entry:* the number of pairs of observations,  $n$ .  
*Constraint:*  $N \geq 1$ .
- 2: X1(N) – *real* array *Input*  
*On entry:* the observations from the first sample,  $x_i$ , for  $i = 1, 2, \dots, n$ .
- 3: X2(N) – *real* array *Input*  
*On entry:* the observations from the second sample,  $y_i$ , for  $i = 1, 2, \dots, n$ .
- 4: IWT – INTEGER *Input/Output*  
*On entry:* indicates whether user-supplied weights are provided by the user:  
 IWT = 1  
     Indicates that user-supplied weights are given in the array WT.  
 IWT  $\neq$  0  
     Indicates that user-supplied weights are not given. In this case the routine assigns the value 1.0 to each element of the weight array, WT.  
*On exit:* IWT is used to indicate the number of valid observations,  $m$ ; see Section 3(g), above.
- 5: WT(N) – *real* array *Input/Output*  
*On entry:* if IWT = 1, then the elements of WT must contain the weights,  $w_i$ , associated with the pairs of observations,  $x_i, y_i$ , for  $i = 1, 2, \dots, n$ .  
 If IWT = 0, then the elements of WT need not be set.  
*On exit:* if IWT = 1, then the elements of WT are unchanged.  
 If IWT = 0 each element of WT will be assigned the value 1.0.
- 6: RES(13) – *real* array *Output*  
*On exit:* the elements of RES contain the following results:  
 RES(1) mean of the first sample,  $\bar{x}$ ;  
 RES(2) mean of the second sample,  $\bar{y}$ ;  
 RES(3) standard deviation of the first sample,  $s_1$ ;  
 RES(4) standard deviation of the second sample,  $s_2$ ;  
 RES(5) corrected sum of squares of the first sample,  $c_{11}$ ;  
 RES(6) corrected sum of products of the two samples,  $c_{12}$ ;  
 RES(7) corrected sum of squares of the second sample,  $c_{22}$ ;  
 RES(8) product-moment correlation coefficient,  $R$ ;  
 RES(9) minimum of the first sample;  
 RES(10) maximum of the first sample;  
 RES(11) minimum of the second sample;  
 RES(12) maximum of the second sample;  
 RES(13) sum of weights,  $\sum_{i=1}^n w_i$  ( $= N$ , if IWT = 0, on entry).

## 7: 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, hence the standard deviation, 3(d), and the product-moment correlation coefficient, 3(e), cannot be calculated.

IFAIL = 3

The number of valid cases,  $m$ , is 0, or at least one of the weights is negative.

## 7 Accuracy

The method used is believed to be stable.

## 8 Further Comments

The time taken by the routine increases linearly with  $n$ .

## 9 Example

In the program below, 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, all the calculated quantities are printed. In the example, there is one set of data, with 29 (unweighted) pairs of observations.

### 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.

```
*      G01ABF 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 ..
      INTEGER          I, IFAIL, IWT, J, N, NPROB
*      .. Local Arrays ..
      real            ANS(13), WT(NMAX), X1(NMAX), X2(NMAX)
*      .. External Subroutines ..
      EXTERNAL        G01ABF
```

```

*      .. Executable Statements ..
WRITE (NOUT,*) 'G01ABF 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,*) (X1(I),X2(I),I=1,N)
    WRITE (NOUT,*) 'Data as input -'
    WRITE (NOUT,*)
+      Var 1      Var 2      Var 1      Var 2      Var 1      Va
+r      2'
    WRITE (NOUT,99995) (X1(I),X2(I),I=1,N)
    IF (IWT.EQ.1) THEN
      READ (NIN,*) (WT(I),I=1,N)
      WRITE (NOUT,*) 'Weights as input -'
      WRITE (NOUT,99994) (WT(I),I=1,N)
    END IF
    IFAIL = 1
*
    CALL G01ABF(N,X1,X2,IWT,WT,ANS,IFAIL)
*
    WRITE (NOUT,*)
    IF (IFAIL.EQ.0) THEN
      WRITE (NOUT,*) 'Successful call of G01ABF'
      WRITE (NOUT,99999) 'No. of valid cases', IWT
      WRITE (NOUT,*)
+      '          Variable 1          Variable 2'
      WRITE (NOUT,99998) 'Mean      ', ANS(1), ANS(2)
      WRITE (NOUT,99998) 'Std devn', ANS(3), ANS(4)
      WRITE (NOUT,99997) 'Corr SSP', ANS(5), ANS(6), ANS(7)
      WRITE (NOUT,99996) 'Correln ', ANS(8)
      WRITE (NOUT,99998) 'Minimum ', ANS(9), ANS(11)
      WRITE (NOUT,99998) 'Maximum ', ANS(10), ANS(12)
      WRITE (NOUT,99998) 'Sum of weights      ', ANS(13)
    ELSE
      WRITE (NOUT,*) 'Unsuccessful call of G01ABF'
      WRITE (NOUT,99999) 'IFAIL =', IFAIL
      IF (IFAIL.EQ.2) THEN
        WRITE (NOUT,99999) 'No. of valid cases', IWT
        WRITE (NOUT,*)
+      '          Variable 1          Variable 2'
        WRITE (NOUT,99998) 'Mean      ', ANS(1), ANS(2)
        WRITE (NOUT,99997) 'Corr SSP', ANS(5), ANS(6), ANS(7)
        WRITE (NOUT,99998) 'Minimum ', ANS(9), ANS(11)
        WRITE (NOUT,99998) 'Maximum ', ANS(10), ANS(12)
        WRITE (NOUT,99998) 'Sum of weights      ', ANS(13)
      END IF
    END IF
  ELSE
    STOP
  END IF
20 CONTINUE
STOP
*
99999 FORMAT (1X,A,I5)
99998 FORMAT (1X,A,F15.1,F30.1)
99997 FORMAT (1X,A,3E15.5)
99996 FORMAT (1X,A,F30.4)
99995 FORMAT (5X,6F11.1)
99994 FORMAT (13X,F9.3)
END

```

## 9.2 Program Data

G01ABF Example Program Data

```

1
29 0
  350   47   550   95   380   211   510   122  1270   530
  300   38  2630  278   810   309   140   75   450   43
2280  407   250  142   540   89   720  159   90   35
  480  103   180   78  3160  969  220  120  860  333
  300   73  1460  147   400   30   620  100  120   55
  780  145   230  101  1070  468  160   86

```

## 9.3 Program Results

G01ABF Example Program Results

```

Problem      1
Number of cases  29
Data as input -
  Var  1      Var  2      Var  1      Var  2      Var  1      Var  2
  350.0      47.0      550.0      95.0      380.0      211.0
  510.0      122.0     1270.0     530.0      300.0      38.0
 2630.0     278.0      810.0     309.0      140.0      75.0
  450.0       43.0     2280.0     407.0      250.0     142.0
  540.0       89.0      720.0     159.0       90.0      35.0
  480.0      103.0     3160.0     969.0     220.0     120.0
  220.0      120.0     860.0     333.0     300.0      73.0
1460.0     147.0      400.0      30.0     620.0     100.0
  120.0       55.0     780.0     145.0     230.0     101.0
1070.0     468.0     160.0      86.0

```

Successful call of G01ABF

```

No. of valid cases  29
  Variable 1      Variable 2
Mean              734.8      185.8
Std devn         765.2      201.1
Corr SSP         0.16396E+08  0.34830E+07  0.11319E+07
Correln          0.8085
Minimum          90.0      30.0
Maximum         3160.0     969.0
Sum of weights          29.0

```

---