NAG Fortran Library Routine Document

G01AHF

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

G01AHF performs a Normal probability plot on a character printing device, with a chosen number of character positions in each direction.

2 Specification

```
 SUBROUTINE G01AHF(X, NOBS, NSTEPX, NSTEPY, ISTAND, IWORK, WORK, LWORK,
1                  XSORT, XBAR, XSTD, IFAIL)
 INTEGER           NOBS, NSTEPX, NSTEPY, ISTAND, IWORK(NOBS), LWORK,
1                  IFAIL
 real              X(NOBS), WORK(LWORK), XSORT(NOBS), XBAR, XSTD
```

3 Description

In a Normal probability plot, the data (x) are plotted against Normal scores (y). The degree of linearity in the resultant plot provides a visual indication of the Normality of distribution of a set of residuals from some fitting process, such as multiple regression.

The data values are sorted into descending order prior to plotting, and may also be standardised to zero mean and unit standard deviation, if requested.

The plot is produced on a character printing device, using a chosen number of character positions in each direction. The output is directed to the current advisory message unit number (see the Users' Note for your implementation). This number may be changed by an appropriate call to X04ABF before calling G01AHF.

Axes are drawn and annotated and data points are plotted on the nearest character position. An appropriate step size for each axis is computed from the list

(0.1, 0.15, 0.2, 0.25, 0.4, 0.5, 0.6, 0.75, 0.8) $\times$ power of 10.

Points are plotted using the digits 1 to 9 to indicate the equivalent number of observations at a particular character position, a letter A to Z for 10 to 35 occurrences, or * if there are 36 or more coincident occurrences. Zero axes are marked if included in the plotting area.

4 References

None.

5 Parameters

1: X(NOBS) – ***real*** array *Input*

On entry: the vector of data values.

Constraint: all data values must not be equal.

2: NOBS – INTEGER *Input*

On entry: the number of data values.

Constraint: NOBS $\geq$ 2.

3: NSTEPX – INTEGER *Input*

On entry: the number of steps (character positions) to be plotted in the x-direction. If the supplied value of NSTEPX is less than 10, the value 10 will be used by the routine. The maximum value for NSTEPX is the number of character positions available on the chosen output device less 15, up to a maximum of 133. If NSTEPX exceeds 133 on input, the value 133 will be used by the routine.

4: NSTEPY – INTEGER *Input*

On entry: the number of steps (character positions) to be plotted in the y-direction. If the supplied value of NSTEPY is less than 10, the value 10 will be used by the routine. There is no maximum value for NSTEPY, but users should bear in mind that (NSTEPY + 5) records (lines) of output are generated by the routine.

5: ISTAND – INTEGER *Input*

On entry: indicates whether the residuals are to be standardised prior to plotting. If ISTAND > 0, the elements of X are standardised to zero mean and unit standard deviation.

6: IWORK(NOBS) – INTEGER array *Workspace*

7: WORK(LWORK) – ***real*** array *Output*

On exit: the first NOBS elements of WORK contain the Normal scores in ascending magnitude. The rest of the array is used as workspace.

8: LWORK – INTEGER *Input*

On entry: the dimension of the array WORK as declared in the (sub)program from which G01AHF is called.

Constraint: LWORK $\geq (5 \times \text{NOBS})/2$.

9: XSORT(NOBS) – ***real*** array *Output*

On exit: the data values, sorted into descending order, and standardised if ISTAND was positive on entry.

10: XBAR – ***real*** *Output*

On exit: the mean of the data values.

11: XSTD – ***real*** *Output*

On exit: the standard deviation of the data values.

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, NOBS < 2.

IFAIL $= 2$

All the supplied data values are equal.

IFAIL $= 3$

On entry, LWORK $< (5 \times \text{NOBS})/2$, i.e., the array WORK is too small.

7 Accuracy

Accuracy is limited by the number of plotting positions available.

8 Further Comments

For details of timing see G01AGF and G01DAF.

No blank records are output before or after the plot.

Users must make sure that it is permissible to write records containing NSTEPX characters to the current advisory message unit.

9 Example

The data are residuals from a linear regression. The 25 values are standardised and plotted against the Normal scores, and are seen to follow a straight line fairly closely, indicating that Normality assumptions are justified.

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.

```
*     G01AHF Example Program Text
*     Mark 15 Revised.  NAG Copyright 1991.
*     .. Parameters ..
      INTEGER          NOBS, LWORK
      PARAMETER        (NOBS=25,LWORK=5*NOBS/2)
      INTEGER          NIN, NOUT
      PARAMETER        (NIN=5,NOUT=6)
*     .. Local Scalars ..
      real             XBAR, XSTD
      INTEGER          I, IFAIL, ISTAND, NSTEPX, NSTEPY
*     .. Local Arrays ..
      real             WORK(LWORK), X(NOBS), XSORT(NOBS)
      INTEGER          IWORK(NOBS)
*     .. External Subroutines ..
      EXTERNAL         G01AHF, X04ABF
*     .. Executable Statements ..
      WRITE (NOUT,*) 'G01AHF Example Program Results'
*     Skip heading in data file
      READ (NIN,*)
      READ (NIN,*) (X(I),I=1,NOBS)
      WRITE (NOUT,*)
      WRITE (NOUT,*) '25 data values to be plotted'
```

```
      WRITE (NOUT,99997) (X(I),I=1,NOBS)
      WRITE (NOUT,*)
      WRITE (NOUT,*)
     +  'Plot of normal scores (Y) against standardised residuals (X)'
      WRITE (NOUT,*)
      CALL X04ABF(1,NOUT)
      NSTEPX = 50
      NSTEPY = 40
      ISTAND = 1
      IFAIL = 1
*
      CALL G01AHF(X,NOBS,NSTEPX,NSTEPY,ISTAND,IWORK,WORK,LWORK,XSORT,
     +            XBAR,XSTD,IFAIL)
*
      IF (IFAIL.NE.0) THEN
         WRITE (NOUT,99999) 'Failed in G01AHF. IFAIL = ', IFAIL
      ELSE
         WRITE (NOUT,*)
         WRITE (NOUT,99998) 'Mean of data values = ', XBAR
         WRITE (NOUT,99998) 'Standard deviation  = ', XSTD
         WRITE (NOUT,*)
         WRITE (NOUT,*) 'Sorted standardised data values'
         WRITE (NOUT,99997) (XSORT(I),I=1,NOBS)
      END IF
      STOP
*
99999 FORMAT (1X,A,I2)
99998 FORMAT (1X,A,F5.2)
99997 FORMAT (5X,5F7.2)
      END
```

9.2 Program Data

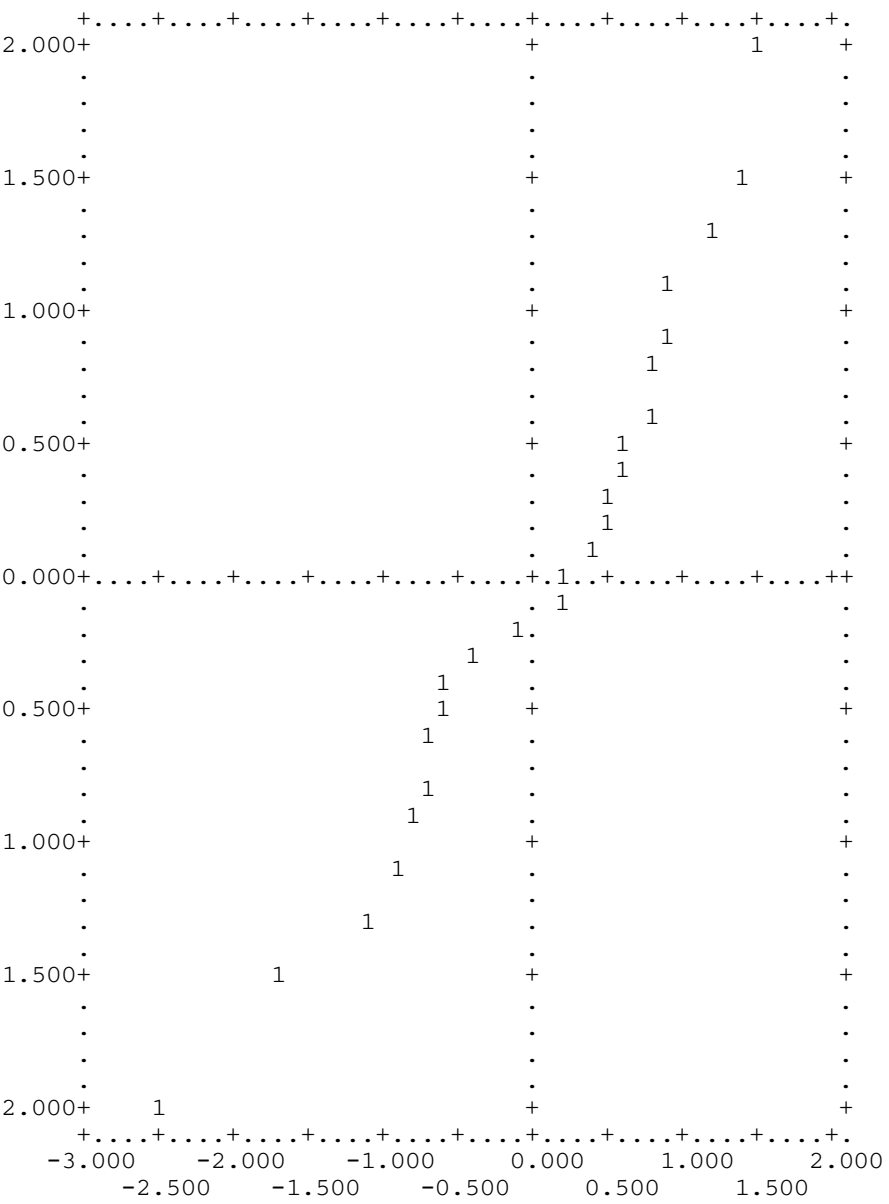
```
G01AHF Example Program Data
  0.35  0.10  0.95 -0.53  0.33
  0.30  0.39  0.26 -0.45  0.12
 -1.58  0.90  0.53 -0.58  0.54
 -0.09  0.79 -0.41  0.54  0.48
 -0.28 -0.71 -1.10 -0.41 -0.44
```

9.3 Program Results

```
G01AHF Example Program Results

25 data values to be plotted
       0.35   0.10   0.95  -0.53   0.33
       0.30   0.39   0.26  -0.45   0.12
      -1.58   0.90   0.53  -0.58   0.54
      -0.09   0.79  -0.41   0.54   0.48
      -0.28  -0.71  -1.10  -0.41  -0.44

Plot of normal scores (Y) against standardised residuals (X)
```




```
Mean of data values =  0.00
Standard deviation  =  0.63

Sorted standardised data values
       1.50   1.42   1.25   0.85   0.85
       0.84   0.76   0.62   0.55   0.52
       0.47   0.41   0.19   0.16  -0.14
      -0.44  -0.65  -0.65  -0.69  -0.71
      -0.84  -0.92  -1.12  -1.74  -2.49
```