NAG 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,     &
                   XSORT, XBAR, XSTD, IFAIL)
INTEGER            NOBS, NSTEPX, NSTEPY, ISTAND, IWORK(NOBS), LWORK,       &
                   IFAIL
REAL (KIND=nag_wp) 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 standardized 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

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

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

4 References

None.

5 Arguments

1: X(NOBS) – REAL (KIND=nag_wp) 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 ≥ 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 G01AHF. 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 G01AHF. There is no maximum value for NSTEPY, but you 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 standardized prior to plotting.

If ISTAND > 0, the elements of X are standardized to zero mean and unit standard deviation.

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

7: WORK(LWORK) – REAL (KIND=nag_wp) 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 (KIND=nag_wp) array *Output*

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

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

On exit: the mean of the data values.

11: XSTD – REAL (KIND=nag_wp) *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. If you are unfamiliar with this argument you should refer to Section 3.4 in How to Use the NAG Library and its Documentation 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 argument, 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, 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.

IFAIL $= -99$

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

See Section 3.9 in How to Use the NAG Library and its Documentation for further information.

IFAIL $= -399$

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

See Section 3.8 in How to Use the NAG Library and its Documentation for further information.

IFAIL $= -999$

Dynamic memory allocation failed.

See Section 3.7 in How to Use the NAG Library and its Documentation for further information.

7 Accuracy

Accuracy is limited by the number of plotting positions available.

8 Parallelism and Performance

G01AHF is threaded by NAG for parallel execution in multithreaded implementations of the NAG Library.

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

For details of timing see G01AGF and G01DAF.

No blank records are output before or after the plot.

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

10 Example

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

10.1 Program Text

```
    Program g01ahfe

!     G01AHF Example Program Text

!     Mark 26 Release. NAG Copyright 2016.

!     .. Use Statements ..
      Use nag_library, Only: g01ahf, nag_wp, x04abf
!     .. Implicit None Statement ..
      Implicit None
!     .. Parameters ..
      Integer, Parameter               :: iset = 1, nin = 5, nout = 6
!     .. Local Scalars ..
      Real (Kind=nag_wp)               :: xbar, xstd
      Integer                          :: ifail, istand, lwork, nobs, nstepx,  &
                                          nstepy, outchn
      Character (80)                   :: title
!     .. Local Arrays ..
      Real (Kind=nag_wp), Allocatable  :: work(:), x(:), xsort(:)
      Integer, Allocatable             :: iwork(:)
!     .. Intrinsic Procedures ..
      Intrinsic                        :: len_trim
!     .. Executable Statements ..
      Write (nout,*) 'G01AHF Example Program Results'
      Write (nout,*)

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

!     Read in the problem size
      Read (nin,*) nobs, nstepx, nstepy, istand

      lwork = 5*nobs/2
      Allocate (iwork(nobs),xsort(nobs),x(nobs),work(lwork))

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

!     Read in the title
      Read (nin,99997) title

!     Set advisory channel
      outchn = nout
      Call x04abf(iset,outchn)

!     Display data to be plotted
      Write (nout,*) 'Data values to be plotted'
      Write (nout,99998) x(1:nobs)
      Write (nout,*)
      Write (nout,*)

!     Display title
      Write (nout,*)
      Write (nout,*) title(1:len_trim(title))
      Write (nout,*)
      Flush (nout)

!     Produce the plot
      ifail = 0
      Call g01ahf(x,nobs,nstepx,nstepy,istand,iwork,work,lwork,xsort,xbar,     &
        xstd,ifail)

!     Display additional results
      Write (nout,*)
      Write (nout,99999) 'Mean of data values = ', xbar
      Write (nout,99999) 'Standard deviation  = ', xstd
      Write (nout,*)
      Write (nout,*) 'Sorted standardised data values'
      Write (nout,99998) xsort(1:nobs)
```

```
99999 Format (1X,A,F5.2)
99998 Format (5X,5F7.2)
99997 Format (A80)
    End Program g01ahfe
```

10.2 Program Data

```
G01AHF Example Program Data
25 50 40 1
  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)
```

10.3 Program Results

```
G01AHF Example Program Results

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)

           +....+....+....+....+....+....+....+....+....+....+....+.
      2.000+                              +              1     +
           .                              .                    .
           .                              .                    .
           .                              .                    .
           .                              .                    .
      1.500+                              +             1      +
           .                              .                    .
           .                              .           1        .
           .                              .                    .
           .                              .        1           .
      1.000+                              +                    +
           .                              .        1           .
           .                              .       1            .
           .                              .                    .
           .                              .       1            .
      0.500+                              +     1              +
           .                              .     1              .
           .                              .    1               .
           .                              .    1               .
           .                              .   1                .
      0.000+....+....+....+....+....+....+.1..+....+....+....++
           .                              . 1                  .
           .                             1.                    .
           .                          1   .                    .
           .                        1     .                    .
     -0.500+                        1     +                    +
           .                       1      .                    .
           .                              .                    .
           .                      1       .                    .
           .                     1        .                    .
     -1.000+                              +                    +
           .                    1         .                    .
           .                              .                    .
           .                  1           .                    .
           .                              .                    .
     -1.500+            1                +                    +
           .                              .                    .
           .                              .                    .
           .                              .                    .
           .                              .                    .
     -2.000+    1                        +                    +
           +....+....+....+....+....+....+....+....+....+....+....+.
         -3.000    -2.000    -1.000     0.000     1.000     2.000
             -2.500    -1.500    -0.500     0.500     1.500

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
```