

## NAG Toolbox

### nag\_stat\_plot\_scatter\_normal (g01ah)

## 1 Purpose

nag\_stat\_plot\_scatter\_normal (g01ah) performs a Normal probability plot on a character printing device, with a chosen number of character positions in each direction.

## 2 Syntax

```
[ifail, xbar, xstd, work, xsort] = nag_stat_plot_scatter_normal(x, nstepx,
nstepy, 'nobs', nobs, 'istand', istand)
[ifail, xbar, xstd, work, xsort] = g01ah(x, nstepx, nstepy, 'nobs', nobs,
'istand', istand)
```

**Note:** the interface to this routine has changed since earlier releases of the toolbox:

At Mark 23: only the first **nobs** elements of **work** are returned; **istand** was made optional (default 1); **lwork** is no longer an input parameter; output parameters were reordered.

## 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. This number may be changed by an appropriate call to nag\_file\_set\_unit\_advisory (x04ab) before calling nag\_stat\_plot\_scatter\_normal (g01ah).

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 Parameters

### 5.1 Compulsory Input Parameters

1: **x(nobs)** – REAL (KIND=nag\_wp) array

The vector of data values.

*Constraint:* all data values must not be equal.

2: **nstepx** – INTEGER

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 nag\_stat\_plot\_scatter\_normal (g01ah). 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 function.

3: **nstepy** – INTEGER

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 nag\_stat\_plot\_scatter\_normal (g01ah). There is no maximum value for **nstepy**, but you should bear in mind that (**nstepy** + 5) records (lines) of output are generated by the function.

## 5.2 Optional Input Parameters

1: **nobs** – INTEGER

*Default:* the dimension of the array **x**.

The number of data values.

*Constraint:* **nobs**  $\geq 2$ .

2: **istand** – INTEGER

*Default:* 1

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.

## 5.3 Output Parameters

1: **ifail** – INTEGER

**ifail** = 0 unless the function detects an error (see Section 5).

2: **xbar** – REAL (KIND=nag\_wp)

The mean of the data values.

3: **xstd** – REAL (KIND=nag\_wp)

The standard deviation of the data values.

4: **work(nobs)** – REAL (KIND=nag\_wp) array

The dimension of the array **work** will be **nobs**

$lwork = (5 \times \mathbf{nobs})/2$ .

the Normal scores in ascending magnitude.

5: **xsort(nobs)** – REAL (KIND=nag\_wp) array

The data values, sorted into descending order, and standardized if **istand** was positive on entry.

## 6 Error Indicators and Warnings

Errors or warnings detected by the function:

**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.

**ifail = -399**

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

**ifail = -999**

Dynamic memory allocation failed.

## 7 Accuracy

Accuracy is limited by the number of plotting positions available.

## 8 Further Comments

For details of timing see nag\_stat\_plot\_scatter\_2var (g01ag) and nag\_stat\_normal\_scores\_exact (g01da).

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.

## 9 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.

### 9.1 Program Text

```
function g01ah_example

fprintf('g01ah example results\n\n');

x = [ 0.35      0.1      0.95     -0.53      0.33      0.3      0.39      0.26 ...
       -0.45      0.12     -1.58      0.9       0.53     -0.58      0.54     -0.09 ...
       0.79     -0.41      0.54      0.48     -0.28     -0.71     -1.1     -0.41 ...
       -0.44];

nstepx = nag_int(50);
nstepy = nag_int(40);

fprintf('Plot of normal scores (Y) against standardised residuals (X)\n\n');

[ifail, xbar, xstd, work, xsort] = g01ah( ...
x, nstepx, nstepy);
```

## 9.2 Program Results

g01ah example results

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

