

NAG Toolbox

nag_stat_plot_box_whisker (g01as)

1 Purpose

nag_stat_plot_box_whisker (g01as) produces a specified number of box and whisker plots on a character printing device, with a chosen number of character positions in each direction.

2 Syntax

```
[ifail, plot, work] = nag_stat_plot_box_whisker(n, x, nstepx, nstepy, 'prt',
prt, 'm', m)
[ifail, plot, work] = g01as(n, x, nstepx, nstepy, 'prt', prt, 'm', m)
```

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

At Mark 23: **prt** was made optional (default 'P'); output parameters were reordered.

3 Description

nag_stat_plot_box_whisker (g01as) produces a series of box and whisker plots representing m data batches each of size n_i , for $i = 1, 2, \dots, m$. A box and whisker plot is a diagrammatic representation of the five-point summary of a data batch. The plot consists of a box spanning the hinges with the median indicated by a third line and two whiskers to represent the extreme values. The five-point summary is calculated internally and is returned in the workspace array.

The plot is returned in the character array **plot**. The size of the plot may be controlled using the arguments **nstepx** and **nstepy**. Optionally the plot can be output to an external file, in which case output is directed to the current advisory message unit as defined by nag_file_set_unit_advisory (x04ab).

An axis corresponding to the y axis is drawn and annotated and data points are plotted to the nearest character position.

4 References

Erickson B H and Nosanchuk T A (1985) *Understanding Data* Open University Press, Milton Keynes
 Tukey J W (1977) *Exploratory Data Analysis* Addison–Wesley

5 Parameters

5.1 Compulsory Input Parameters

1: **n(m)** – INTEGER array

n(i) contains the number of observations in the i th batch, n_i , for $i = 1, 2, \dots, m$.

If $n_i < 5$ the i th batch is omitted from the plot.

Constraint: at least one **n(i)** must be greater than or equal to 5, for $i = 1, 2, \dots, m$.

2: **x(ldx, m)** – REAL (KIND=nag_wp) array

ldx , the first dimension of the array, must satisfy the constraint $ldx \geq \max\{\mathbf{n}(i)\}$.

The i th column of **x** must contain the data for the i th batch, that is **x(j, i)** must contain the j th observation of the i th batch, for $i = 1, 2, \dots, m$ and $j = 1, 2, \dots, n_i$.

- 3: **nstepx** – INTEGER
The number of character positions to be plotted in the x -direction.
Constraint: $\mathbf{nstepx} \geq \max(19, (15 \times \mathbf{m}/4 + 9))$.
- 4: **nstepy** – INTEGER
The number of character positions to be plotted in the y -direction.
Constraint: $\mathbf{nstepy} \geq 9$.

5.2 Optional Input Parameters

- 1: **prt** – CHARACTER(1)
Default: 'P'
Indicates whether the box and whisker plot is to be output to an external file.
prt = 'N'
The box and whisker plot is not output to an external file.
prt = 'P'
The box and whisker plot is output to the current advisory message unit as defined by `nag_file_set_unit_advisory (x04ab)`.
Constraint: **prt** = 'P' or 'N'.
- 2: **m** – INTEGER
Default: the dimension of the array **n** and the second dimension of the array **x**. (An error is raised if these dimensions are not equal.)
 m , the number of data batches that are to be represented.
Constraint: $\mathbf{m} > 0$.

5.3 Output Parameters

- 1: **ifail** – INTEGER
ifail = 0 unless the function detects an error (see Section 5).
- 2: **plot**(*ldplot*, **nstepx**) – CHARACTER(1) array
Contains the box and whisker plots.
- 3: **work**($5 \times \mathbf{m}$) – REAL (KIND=nag_wp) array
work(j), for $j = (i - 1) \times 5 + 1, \dots, (i - 1) \times 5 + 5$, contains the five-point summary of the i th batch.

6 Error Indicators and Warnings

Errors or warnings detected by the function:

ifail = 1

On entry, $\mathbf{n}(i) < 5$ for some i , for $i = 1, 2, \dots, m$. For each batch where this occurs, 5 crosses are plotted in a vertical line to indicate that insufficient data was provided to produce a five-point summary and box-plot for that particular batch.

ifail = 2

On entry, **nstepx** < max(19, 15 × **m**/4 + 9). This indicates that the data region defined by **nstepx** is too small to produce the required plot.

ifail = 3

On entry, **nstepy** < 9.

ifail = 4

On entry, *ldplot* < **nstepy**.

ifail = 5

On entry, **prt** ≠ 'P' or 'N'.

ifail = 6

On entry, *ldx* < max(**n**(*i*)), for *i* = 1, 2, ..., *m*.

ifail = 7

The number of observations in all batches is less than 5.

ifail = 8

On entry, the data values are all identical.

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

If the range of observations in a particular batch is too small to allow each item of the five-point summary to be plotted separately, then a sequence of stars are plotted at the median point of the batch to indicate that the full box-plot could not be plotted.

8 Further Comments

The time taken by `nag_stat_plot_box_whisker` (g01as) increases with *m* and *n_i*, for *i* = 1, 2, ..., *m*.

9 Example

The following program produces a box and whisker plot for each one of 5 data batches of sizes 5, 6, 8, 8 and 7 respectively and prints the 5 box and whisker plots on the current advisory message unit.

9.1 Program Text

```
function g01as_example

fprintf('g01as example results\n\n');

n = nag_int([ 5; 6; 8; 8; 7]);
x = [ -9, -5.6, -9, 12, -5;
      -7.3, 12, 12, 10, 8;
      -4.9, -9, -6, 10, 1.5;
      -2.4, -3.9, -3, 8, -3.2;
      -0.6, -2.4, 0, 6, -3.2;
      0, -7.3, 3, 4, 6.2;
      0, 0, 6, 4, 6.2;
      0, 0, 9, -9, 0];

nstepx = nag_int(55);
nstepy = nag_int(21);
[ifail, plot, work] = g01as( ...
    n, x, nstepx, nstepy);
```

9.2 Program Results

```
g01as example results

0.1E+02+      ---      ---      ---
:             :         :         :
:             :         :         :-----
0.8E+01+      :         :         :         :
:             :         :-----:         :-----
:             :         :         :         :
0.4E+01+      :         :         :         :
:             :         :         :-----:         :
:             :         :         :         :
:             :         :-----:         :-----
-0.6E+00+ ---      :         :         :         :
:             :         :         :         :
:-----:         :         :         :         :-----
: : : :         : : : :         : : : :         :
-0.5E+01+ :---: :---: :---:         :         :
: : : :         : : : :         :         :
:-----:         :         :         :
: : : :         :         :         :
-0.9E+01+ ---      ---      ---      ---
```
