NAG Fortran Library Routine Document

G01ASF

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

G01ASF 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 Specification

```
SUBROUTINE G01ASF(PRT, M, N, X, LDX, NSTEPX, NSTEPY, PLOT, LDP, WORK,
1                 LWORK, IFAIL)
INTEGER           M, N(M), LDX, NSTEPX, NSTEPY, LDP, LWORK(LDX), IFAIL
real              X(LDX,M), WORK(5*M)
CHARACTER*1       PRT, PLOT(LDP,NSTEPX)
```

3 Description

G01ASF produces a series of box and whisker plots representing m data batches each of size n_i, for $i = 1, 2, \ldots, m$. 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 parameters 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 X04ABF.

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

4 References

Tukey J W (1977) *Exploratory Data Analysis* Addison-Wesley

Erickson B H and Nosanchuk T A (1985) *Understanding Data* Open University Press, Milton Keynes

5 Parameters

1: PRT – CHARACTER*1 *Input*

On entry: indicates whether the box and whisker plot is to be output to an external file.

If PRT = 'N', then the box and whisker plot is not output to an external file.

If PRT = 'P', then the box and whisker plot is output to the current advisory message unit as defined by X04ABF.

Constraint: PRT = 'P' or 'N'.

2: M – INTEGER *Input*

On entry: the number of data batches that are to be represented, m.

Constraint: M > 0.

3: N(M) – INTEGER array *Input*

On entry: N(i) contains the number of observations in the ith batch, n_i, for $i = 1, 2, \ldots, m$.

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

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

4: X(LDX,M) – ***real*** array *Input*

On entry: the ith column of X must contain the data for the ith batch, that is X(j, i) must contain the jth observation of the ith batch, for $j = 1, 2, \ldots, n_i$; $i = 1, 2, \ldots, m$.

5: LDX – INTEGER *Input*

On entry: the first dimension of the array X as declared in the (sub)program from which G01ASF is called.

Constraint: LDX $\geq$ max(N(i)).

6: NSTEPX – INTEGER *Input*

On entry: the number of character positions to be plotted in the x-direction.

Constraint: NSTEPX $\geq \max(19, (15 \times \mathrm{M}/4 + 9))$.

7: NSTEPY – INTEGER *Input*

On entry: the number of character positions to be plotted in the y-direction.

Constraint: NSTEPY ≥ 9.

8: PLOT(LDP,NSTEPX) – CHARACTER*1 array *Output*

On exit: PLOT contains the box and whisker plots.

9: LDP – INTEGER *Input*

On entry: the first dimension of the array PLOT as declared in the (sub)program from which G01ASF is called.

Constraint: LDP $\geq$ NSTEPY.

10: WORK(5*M) – ***real*** array *Output*

On exit: WORK(j), for $j = (i-1) \times 5 + 1, (i-1) \times 5 + 2, \ldots, (i-1) \times 5 + 5$, contains the five-point summary of the ith batch.

11: LWORK(LDX) – INTEGER array *Workspace*

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, N$(i) < 5$ for some i, for $i = 1, 2, \ldots, m$. 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$(19, 15 \times$ 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, LDP $<$ NSTEPY.

IFAIL $= 5$

On entry, PRT $\neq$ 'P' or 'N'.

IFAIL $= 6$

On entry, LDX $<$ max$(\text{N}(i))$, for $i = 1, 2, \ldots, m$.

IFAIL $= 7$

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

IFAIL $= 8$

On entry, the data values are all identical.

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 the routine increases with m and n_i, for $i = 1, 2, \ldots, m$.

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

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.

```
*     G01ASF Example Program Text
*     Mark 14 Release.  NAG Copyright 1989.
*     .. Parameters ..
      INTEGER          NIN, NOUT
      PARAMETER        (NIN=5,NOUT=6)
      INTEGER          NMAX, LDX
```

```
      PARAMETER        (NMAX=5,LDX=10)
*     .. Local Scalars ..
      INTEGER          I, IFAIL, J, LDP, M, NSTEPX, NSTEPY
*     .. Local Arrays ..
      real             WORK(5*NMAX), X(LDX,NMAX)
      INTEGER          IWORK(LDX), N(NMAX)
      CHARACTER        PLOT(60,132)
*     .. External Subroutines ..
      EXTERNAL         G01ASF, X04ABF
*     .. Executable Statements ..
      WRITE (NOUT,*) 'G01ASF Example Program Results'
*     Skip heading in data file
      READ (NIN,*)
*     Set advisory message unit for plot output to NOUT
      CALL X04ABF(1,NOUT)
      READ (NIN,*) M, (N(I),I=1,M)
      READ (NIN,*) NSTEPX, NSTEPY
      DO 20 J = 1, M
         READ (NIN,*) (X(I,J),I=1,N(J))
   20 CONTINUE
      LDP = NSTEPY
      IFAIL = 0
      WRITE (NOUT,*)
*
      CALL G01ASF('Print',M,N,X,LDX,NSTEPX,NSTEPY,PLOT,LDP,WORK,IWORK,
     +            IFAIL)
*
      STOP
      END
```

9.2 Program Data

```
G01ASF Example Program Data
  5   5   6   8   8   7                          :N(i)
  55  21                                         :PMIN PMAX NSTEPY
 -9.0 -7.3 -4.9 -2.4 -0.6
 -5.6 12.0 -9.0 -3.9 -2.4 -7.3
 -9.0 12.0 -6.0 -3.0  0.0  3.0  6.0  9.0
 12.0 10.0 10.0  8.0  6.0  4.0  4.0 -9.0
 -5.0  8.0  1.5 -3.2 -3.2  6.2  6.2              :X(i,j)
```

9.3 Program Results

```
 G01ASF Example Program Results

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