NAG Fortran Library Routine Document

G05EJF

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

G05EJF selects a pseudo-random sample without replacement from an integer vector.

2 Specification

```
SUBROUTINE G05EJF(IA, N, IB, M, IFAIL)
INTEGER           IA(N), N, IB(N), M, IFAIL
```

3 Description

The routine selects m elements from vector IA of length n and places them in vector IB. Their order in IA will be preserved in IB. Each of the $\binom{n}{m}$ possible combinations of elements of IA may be regarded as being equally probable.

For moderate or large values of n (greater than 75 say), it is theoretically impossible that all combinations of size m may occur, unless m is near 1 or near n. This is because $\binom{n}{m}$ exceeds the cycle length of G05CAF. For practical purposes this is irrelevant, as the time taken to generate all possible combinations is many millenia.

4 References

Knuth D E (1981) *The Art of Computer Programming (Volume 2)* (2nd Edition) Addison-Wesley

Kendall M G and Stuart A (1969) *The Advanced Theory of Statistics (Volume 1)* (3rd Edition) Griffin

5 Parameters

1: IA(N) – INTEGER array *Input*

On entry: the population to be sampled.

2: N – INTEGER *Input*

On entry: the number of elements in the vector to be sampled.

Constraint: N $\geq$ 1.

3: IB(N) – INTEGER array *Output*

On exit: the selected sample.

4: M – INTEGER *Input*

On entry: the sample size.

Constraint: 1 $\leq$ M $\leq$ N.

5: 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 < 1.

IFAIL $= 2$

On entry, M < 1,
or M $>$ N.

7 Accuracy

Not applicable.

8 Further Comments

The time taken by the routine is of order n.

In order to sample other kinds of vectors, or matrices of higher dimension, the following technique may be used:

(a) set IA$(i) = i$, for $i = 1, 2, \ldots, n$,

(b) use G05EJF to take a sample from IA and put it into IB,

(c) use the contents of IB as a set of indices to access the relevant vector or matrix.

In order to divide a population into several groups, G05EHF is more efficient.

9 Example

From a vector containing the first eight positive integers in ascending order, random samples of size 1,2,...,8 are selected and printed.

The generator mechanism used is selected by an initial call to G05ZAF.

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.

```
*     G05EJF Example Program Text
*     Mark 20 Revised. NAG Copyright 2001.
*     .. Parameters ..
      INTEGER          N
      PARAMETER        (N=8)
      INTEGER          NOUT
```

```
      PARAMETER        (NOUT=6)
*     .. Local Scalars ..
      INTEGER          I, IFAIL, K, M
*     .. Local Arrays ..
      INTEGER          IA(N), IB(N)
*     .. External Subroutines ..
      EXTERNAL         G05CBF, G05EJF, G05ZAF
*     .. Executable Statements ..
      CALL G05ZAF('O')
      WRITE (NOUT,*) 'G05EJF Example Program Results'
      WRITE (NOUT,*)
      CALL G05CBF(0)
      WRITE (NOUT,99999) 'Samples from the first ', N, ' integers'
      WRITE (NOUT,*)
      WRITE (NOUT,*) 'Sample size     Values'
      DO 20 I = 1, N
         IA(I) = I
   20 CONTINUE
      DO 40 M = 1, N
         IFAIL = 0
*
         CALL G05EJF(IA,N,IB,M,IFAIL)
*
         WRITE (NOUT,99998) M, (IB(K),K=1,M)
   40 CONTINUE
      STOP
*
99999 FORMAT (1X,A,I1,A)
99998 FORMAT (1X,I6,10X,8I3)
      END
```

9.2 Program Data

None.

9.3 Program Results

```
 G05EJF Example Program Results

 Samples from the first 8 integers

 Sample size     Values
      1            6
      2            1  7
      3            1  3  4
      4            1  2  6  8
      5            1  3  4  6  7
      6            1  2  3  4  5  6
      7            1  2  3  4  6  7  8
      8            1  2  3  4  5  6  7  8
```