

NAG Library Routine Document

G05MZF

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

G05MZF generates a vector of pseudorandom integers from a discrete distribution with a given PDF (probability density function) or CDF (cumulative distribution function) p .

2 Specification

```

SUBROUTINE G05MZF(MODE, P, NP, IP1, ITYPE, N, X, IGEN, ISEED, R, NR,
1                      IFAIL)
    INTEGER          MODE, NP, IP1, ITYPE, N, X(N), IGEN, ISEED(4), NR,
1                      IFAIL
    double precision P(NP), R(NR)

```

3 Description

G05MZF generates a sequence of n integers x_i , from a discrete distribution defined by information supplied in P. This may either be the PDF or CDF of the distribution. A reference vector is first set up to contain the CDF of the distribution in its higher elements, followed by an index. The full specifications of the reference vector are as follows.

$R(1)$ = the number of elements of index, k .

$R(2)$ = a check number to make sure that the values of IP1 and ITYPE haven't changed when calling G05MZF with $MODE = 1$.

$R(3)$ = the number of values the variates can take (i.e., the first value of k such that $CDF(k) = 1$).

$R(4) = IP1 - 1$.

$R(5)$ = the space available for indexing = $NR - (R(3) + 5)$.

$R(i + 5)$, for $i = 1, 2, \dots, R(3)$, the CDF.

$R(i) = \min\{j \mid CDF(j) > (i - 1)/k\}$, for $i = R(3) + 6, \dots, NR$.

Setting up the reference vector and subsequent generation of variates can each be performed by separate calls to G05MZF or may be combined in a single call.

One of the initialization routines G05KBF (for a repeatable sequence if computed sequentially) or G05KCF (for a non-repeatable sequence) must be called prior to the first call to G05MZF.

4 References

Kendall M G and Stuart A (1969) *The Advanced Theory of Statistics (Volume 1)* (3rd Edition) Griffin
 Knuth D E (1981) *The Art of Computer Programming (Volume 2)* (2nd Edition) Addison-Wesley

5 Parameters

1: MODE – INTEGER *Input*

On entry: a code for selecting the operation to be performed by the routine.

MODE = 0

Set up reference vector only.

MODE = 1

Generate variates using reference vector set up in a prior call to G05MZF.

MODE = 2

Set up reference vector and generate variates.

Constraint: MODE = 0, 1 or 2.

- 2: P(NP) – **double precision** array *Input*
On entry: the PDF or CDF of the distribution.

- 3: NP – INTEGER *Input*
On entry: the number of values supplied in P defining the PDF or CDF of the discrete distribution.
Constraint: NP > 0.

- 4: IP1 – INTEGER *Input*
On entry: the value of the variate, a whole number, to which the probability in P(1) corresponds.

- 5: ITYPE – INTEGER *Input*
On entry: indicates the type of information contained in P.
 ITYPE = 1
 P contains a probability distribution function (PDF).
 ITYPE = 2
 P contains a cumulative distribution function (CDF).
Constraint: ITYPE = 1 or 2.

- 6: N – INTEGER *Input*
On entry: n , the number of pseudorandom numbers to be generated.
Constraint: N ≥ 1.

- 7: X(N) – INTEGER array *Output*
On exit: contains n pseudorandom numbers from the specified discrete distribution.

- 8: IGEN – INTEGER *Input*
On entry: must contain the identification number for the generator to be used to return a pseudorandom number and should remain unchanged following initialization by a prior call to G05KBF or G05KCF.

- 9: ISEED(4) – INTEGER array *Communication Array*
On entry: contains values which define the current state of the selected generator.
On exit: contains updated values defining the new state of the selected generator.

- 10: R(NR) – **double precision** array *Communication Array*
On entry: if MODE = 1, the reference vector from the previous call to G05MZF.
On exit: the reference vector.

- 11: NR – INTEGER *Input*
On entry: the dimension of the array R as declared in the (sub)program from which G05MZF is called.

Suggested value: $NR = 8 + 1.4 \times NP$ approximately (for optimum efficiency in generating variates).

Constraint: $NR \geq NP + 6$.

12: IFAIL – INTEGER

Input/Output

On entry: IFAIL must be set to 0, -1 or 1. If you are unfamiliar with this parameter you should refer to Section 3.3 in the Essential Introduction 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, if you are 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, NR is too small when $MODE = 0$ or 2 (see Section 5).

IFAIL = 3

On entry, $NP < 1$.

IFAIL = 4

On entry, $ITYPE \neq 1$ or 2 .

IFAIL = 5

On entry, $MODE \neq 0, 1$ or 2 .

IFAIL = 6

With $ITYPE = 1$, $P(j) < 0$ for at least one value of j .

IFAIL = 7

With $ITYPE = 1$, the sum of $P(j)$, for $j = 1, 2, \dots, NP$, does not equal 1.

IFAIL = 8

With $ITYPE = 2$, the values of $P(j)$ are not all in non-descending order.

IFAIL = 9

The value of NP, ITYPE or IP1 is not the same as when R was set up in a previous call with $MODE = 0$ or 2 .

7 Accuracy

Not applicable.

8 Further Comments

None.

9 Example

This example prints 20 pseudorandom variates from a discrete distribution whose PDF, $f(n)$, is defined as follows:

n	$f(n)$
-5	0.01
-4	0.02
-3	0.04
-2	0.08
-1	0.20
0	0.30
1	0.20
2	0.08
3	0.04
4	0.02
5	0.01

The reference vector is set up and the variates are generated by a single call to G05MZF, after initialization by G05KBF.

9.1 Program Text

```
*      G05MZF Example Program Text
*      Mark 20 Release. NAG Copyright 2001.
*      .. Parameters ..
      INTEGER          NP, NOUT, N, NR
      PARAMETER        (NP=11,NOUT=6,N=20,NR=60)
*      .. Local Scalars ..
      INTEGER          I, IFAIL, IGEN, IP1, ITYPE
*      .. Local Arrays ..
      DOUBLE PRECISION P(NP), R(NR)
      INTEGER          ISEED(4), X(N)
*      .. External Subroutines ..
      EXTERNAL         G05KBF, G05MZF
*      .. Data statements ..
      DATA             P/0.01D0, 0.02D0, 0.04D0, 0.08D0, 0.2D0, 0.3D0,
+      0.2D0, 0.08D0, 0.04D0, 0.02D0, 0.01D0/
*      .. Executable Statements ..
      WRITE (NOUT,*) 'G05MZF Example Program Results'
      WRITE (NOUT,*)
*      Set the distribution parameters P and M
      IP1 = -5
      ITYPE = 1
*      Initialize the seed to a repeatable sequence
      ISEED(1) = 1762543
      ISEED(2) = 9324783
      ISEED(3) = 42344
      ISEED(4) = 742355
*      IGEN identifies the stream.
      IGEN = 1
      CALL G05KBF(IGEN,ISEED)
*      Choose MODE = 2
      IFAIL = 1
      CALL G05MZF(2,P,NP,IP1,ITYPE,N,X,IGEN,ISEED,R,NR,IFAIL)
*
*      IF (IFAIL.EQ.0) THEN
*
*          WRITE (NOUT,99999) (X(I),I=1,N)
*
*      ELSE
*          WRITE (NOUT,99998) IFAIL
```

```
      END IF
*
99999 FORMAT (1X,I12)
99998 FORMAT (1X,' ** G05MZF returned with IFAIL = ',I5)
      END
```

9.2 Program Data

None.

9.3 Program Results

G05MZF Example Program Results

```
-2
 3
 0
 1
 3
 0
 0
 0
-1
 2
 0
 0
 0
 5
-3
 0
 1
 0
-3
 1
```
