NAG Toolbox for MATLAB

g05mz

1   Purpose

g05mz generates a vector of pseudo-random integers from a discrete distribution with a given PDF (probability density function) or CDF (cumulative distribution function) p.

2   Syntax

```
[x, iseed, r, ifail] = g05mz(mode, p, ip1, itype, n, igen, iseed, r,
'np', np)
```

3   Description

g05mz 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.

$\mathbf{r}(1) = $ the number of elements of index, k.

$\mathbf{r}(2) = $ a check number to make sure that the values of **ip1** and **itype** haven't changed when calling g05mz with **mode** $= 1$.

$\mathbf{r}(3) = $ the number of values the variates can take (i.e., the first value of k such that $\mathrm{CDF}(k) = 1$).

$\mathbf{r}(4) = \mathbf{ip1} - 1$.

$\mathbf{r}(5) = $ the space available for indexing $= \mathbf{nr} - (\mathbf{r}(3) + 5)$.

$\mathbf{r}(i + 5)$, for $i = 1, 2, \ldots, \mathbf{r}(3)$, the CDF.

$\mathbf{r}(i) = \min\{j \mid \mathrm{CDF}(j) > (i - 1)/k\}$, for $i = \mathbf{r}(3) + 6, \ldots, \mathbf{nr}$.

Setting up the reference vector and subsequent generation of variates can each be performed by separate calls to g05mz or may be combined in a single call.

One of the initialization functions g05kb (for a repeatable sequence if computed sequentially) or g05kc (for a non-repeatable sequence) must be called prior to the first call to g05mz.

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

5.1   Compulsory Input Parameters

1:    **mode – int32 scalar**

A code for selecting the operation to be performed by the function:

mode $= 0$

Set up reference vector only.

mode $= 1$

Generate variates using reference vector set up in a prior call to g05mz.

mode $= 2$

Set up reference vector and generate variates.

Constraint: $0 \leq \textbf{mode} \leq 2$.

2:    **p**(**np**) − **double array**

The PDF or CDF of the distribution.

3:    **ip1** − **int32 scalar**

The value of the variate, assumed to be a whole number, to which the probability in **p**(1) corresponds.

4:    **itype** − **int32 scalar**

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.

5:    **n** − **int32 scalar**

n, the number of pseudo-random numbers to be generated.

Constraint: $\textbf{n} \geq 1$.

6:    **igen** − **int32 scalar**

Must contain the identification number for the generator to be used to return a pseudo-random number and should remain unchanged following initialization by a prior call to g05kb or g05kc.

7:    **iseed**(4) − **int32 array**

Contains values which define the current state of the selected generator.

8:    **r**(**nr**) − **double array**

If **mode** $= 1$, the reference vector from the previous call to g05mz.

5.2 Optional Input Parameters

1:    **np** − **int32 scalar**

Default: The dimension of the array **p**.

the number of values supplied in **p** defining the PDF or CDF of the discrete distribution.

Constraint: $\textbf{np} > 0$.

5.3 Input Parameters Omitted from the MATLAB Interface

nr

5.4 Output Parameters

1: **x(n) – int32 array**

Contains n pseudo-random numbers from the specified discrete distribution.

2: **iseed(4) – int32 array**

Contains updated values defining the new state of the selected generator.

3: **r(nr) – double array**

The reference vector.

4: **ifail – int32 scalar**

0 unless the function detects an error (see Section 6).

6 Error Indicators and Warnings

Errors or warnings detected by the function:

ifail = 1

On entry, $\mathbf{n} < 1$.

ifail = 2

On entry, **nr** is too small when **mode** $= 0$ or 2 (see Section 5).

ifail = 3

On entry, $\mathbf{np} < 1$.

ifail = 4

On entry, **itype** < 1,
or **itype** > 2.

ifail = 5

On entry, **mode** < 0,
or **mode** > 2.

ifail = 6

With **itype** $= 1$, $\mathbf{p}(j) < 0$ for at least one value of j.

ifail = 7

With **itype** $= 1$, the sum of $\mathbf{p}(j)$, for $j = 1, 2, \ldots, \mathbf{np}$, does not equal 1.

ifail = 8

With **itype** $= 2$, the values of $\mathbf{p}(j)$ 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

```
mode = int32(2);
p = [0.01;
     0.02;
     0.04;
     0.08;
     0.2;
     0.3;
     0.2;
     0.08;
     0.04;
     0.02;
     0.01];
ip1 = int32(-5);
itype = int32(1);
n = int32(20);
igen = int32(1);
iseed = [int32(1762543);
     int32(9324783);
     int32(42344);
     int32(742355)];
r = zeros(60, 1);
[x, iseedOut, rOut, ifail] = g05mz(mode, p, ip1, itype, n, igen, iseed,
r)

x =
          -2
           3
           0
           1
           3
           0
           0
           0
          -1
           2
           0
           0
           0
           5
          -3
           0
           1
           0
          -3
           1
iseedOut =
     7198201
     6860924
     8806996
     4852644
rOut =
     array elided
ifail =
           0
```