

NAG Toolbox for MATLAB

g05ef

1 Purpose

g05ef sets up a reference vector **r** for a hypergeometric distribution of the number of specified items in a sample of size l , taken from a population of size n with m specified items in it.

2 Syntax

```
[r, ifail] = g05ef(l, m, n, nr)
```

3 Description

g05ef sets up a reference vector for use in g05ey. Together these functions produce random numbers from the distribution defined by:

$$P(I = i) = \frac{l!m!(n-l)!(n-m)!}{i!(l-i)!(m-i)!(n-m-l+i)!n!} \quad \text{if} \quad i = \max(0, m+l-n), \dots, \min(l, m),$$

$$P(I = i) = 0 \quad \text{otherwise.}$$

The reference array is generated by a recurrence relation if $lm(n-l)(n-m) < 50n^3$, otherwise Stirling's approximation is used.

4 References

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

5 Parameters

5.1 Compulsory Input Parameters

1: **l – int32 scalar**

The parameter (sample size) l of the distribution.

Constraint: $0 \leq l \leq n$.

2: **m – int32 scalar**

The parameter (number of specified items) m of the distribution.

Constraint: $0 \leq m \leq n$.

3: **n – int32 scalar**

The parameter (population size) n of the distribution.

Constraint: $n \geq 0$.

4: **nr – int32 scalar**

Suggested value: $nr = 20 + \sqrt{(l \times m \times (n-m) \times (n-l))/n^3}$ approximately (for optimum efficiency in g05ey).

Constraint: **nr** must not be too small, but the limit is too complicated to specify. The suggested value will always be big enough.

5.2 Optional Input Parameters

None.

5.3 Input Parameters Omitted from the MATLAB Interface

None.

5.4 Output Parameters

1: **r(nr)** – double array

The reference vector.

2: **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, **n** is negative.

ifail = 2

On entry, **l** < 0,
or **l** > **n**.

ifail = 3

On entry, **m** < 0,
or **m** > **n**.

ifail = 4

On entry, **nr** is too small (see Section 5).

7 Accuracy

Not applicable.

8 Further Comments

The time taken by g05ef increases with **nr**.

9 Example

```
l = int32(100);
m = int32(50);
n = int32(1000);
nr = int32(62);
g05za('O');
g05cb(int32(0));
[r, ifail] = g05ef(l, m, n, nr)
```

```
r =
    array elided
ifail =
```

0
