# NAG Library Routine Document

# G01BLF

**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

G01BLF returns the lower tail, upper tail and point probabilities associated with a hypergeometric distribution.

## 2 Specification

```
SUBROUTINE G01BLF (N, L, M, K, PLEK, PGTK, PEQK, IFAIL)

INTEGER            N, L, M, K, IFAIL
REAL (KIND=nag_wp) PLEK, PGTK, PEQK
```

## 3 Description

Let $X$ denote a random variable having a hypergeometric distribution with parameters $n$, $l$ and $m$ ($n \geq l \geq 0$, $n \geq m \geq 0$). Then

$$\text{Prob}\{X = k\} = \frac{\binom{m}{k}\binom{n-m}{l-k}}{\binom{n}{l}},$$

where $\max(0, l-(n-m)) \leq k \leq \min(l, m)$, $0 \leq l \leq n$ and $0 \leq m \leq n$.

The hypergeometric distribution may arise if in a population of size $n$ a number $m$ are marked. From this population a sample of size $l$ is drawn and of these $k$ are observed to be marked.

The mean of the distribution $= \dfrac{lm}{n}$, and the variance $= \dfrac{lm(n-l)(n-m)}{n^2(n-1)}$.

G01BLF computes for given $n$, $l$, $m$ and $k$ the probabilities:

$$\begin{aligned} \text{PLEK} &= \text{Prob}\{X \leq k\} \\ \text{PGTK} &= \text{Prob}\{X > k\} \\ \text{PEQK} &= \text{Prob}\{X = k\}. \end{aligned}$$

The method is similar to the method for the Poisson distribution described in Knüsel (1986).

## 4 References

Knüsel L (1986) Computation of the chi-square and Poisson distribution *SIAM J. Sci. Statist. Comput.* **7** 1022–1036

## 5 Parameters

1: N – INTEGER *Input*

*On entry*: the parameter $n$ of the hypergeometric distribution.

*Constraint*: N $\geq 0$.

2: L – INTEGER *Input*

*On entry*: the parameter $l$ of the hypergeometric distribution.

*Constraint*: $0 \leq \mathrm{L} \leq \mathrm{N}$.

3: M – INTEGER *Input*

*On entry*: the parameter $m$ of the hypergeometric distribution.

*Constraint*: $0 \leq \mathrm{M} \leq \mathrm{N}$.

4: K – INTEGER *Input*

*On entry*: the integer $k$ which defines the required probabilities.

*Constraint*: $\max(0, \mathrm{L} - (\mathrm{N} - \mathrm{M})) \leq \mathrm{K} \leq \min(\mathrm{L}, \mathrm{M})$.

5: PLEK – REAL (KIND=nag_wp) *Output*

*On exit*: the lower tail probability, $\mathrm{Prob}\{X \leq k\}$.

6: PGTK – REAL (KIND=nag_wp) *Output*

*On exit*: the upper tail probability, $\mathrm{Prob}\{X > k\}$.

7: PEQK – REAL (KIND=nag_wp) *Output*

*On exit*: the point probability, $\mathrm{Prob}\{X = k\}$.

8: 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.

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.**

*On exit*: $\mathrm{IFAIL} = 0$ unless the routine detects an error or a warning has been flagged (see Section 6).

## 6 Error Indicators and Warnings

If on entry $\mathrm{IFAIL} = 0$ or $-1$, explanatory error messages are output on the current error message unit (as defined by X04AAF).

Errors or warnings detected by the routine:

$\mathrm{IFAIL} = 1$

On entry, $\mathrm{N} < 0$.

$\mathrm{IFAIL} = 2$

On entry, $\mathrm{L} < 0$,
or $\mathrm{L} > \mathrm{N}$.

$\mathrm{IFAIL} = 3$

On entry, $\mathrm{M} < 0$,
or $\mathrm{M} > \mathrm{N}$.

IFAIL = 4

On entry, K < 0,
or K > L,
or K > M,
or K < L + M − N.

IFAIL = 5

On entry, N is too large to be represented exactly as a real number.

IFAIL = 6

On entry, the variance (see Section 3) exceeds $10^6$.

## 7 Accuracy

Results are correct to a relative accuracy of at least $10^{-6}$ on machines with a precision of 9 or more decimal digits, and to a relative accuracy of at least $10^{-3}$ on machines of lower precision (provided that the results do not underflow to zero).

## 8 Further Comments

The time taken by G01BLF depends on the variance (see Section 3) and on $k$. For given variance, the time is greatest when $k \approx lm/n$ (= the mean), and is then approximately proportional to the square-root of the variance.

## 9 Example

This example reads values of $n$, $l$, $m$ and $k$ from a data file until end-of-file is reached, and prints the corresponding probabilities.

### 9.1 Program Text

```
    Program g01blfe

!     G01BLF Example Program Text

!     Mark 24 Release. NAG Copyright 2012.

!     .. Use Statements ..
      Use nag_library, Only: g01blf, nag_wp
!     .. Implicit None Statement ..
      Implicit None
!     .. Parameters ..
      Integer, Parameter               :: nin = 5, nout = 6
!     .. Local Scalars ..
      Real (Kind=nag_wp)               :: peqk, pgtk, plek
      Integer                          :: ifail, k, l, m, n
!     .. Executable Statements ..
      Write (nout,*) 'G01BLF Example Program Results'
      Write (nout,*)

!     Skip heading in data file
      Read (nin,*)

!     Display titles
      Write (nout,*) '   N   L   M   K     PLEK      PGTK      PEQK'
      Write (nout,*)

!     Loop over all dta
d_lp: Do
        Read (nin,*,Iostat=ifail) n, l, m, k
        If (ifail/=0) Then
```

```
!         All data processed
          Exit d_lp
        End If

!       Calculate probability
        ifail = 0
        Call g01blf(n,l,m,k,plek,pgtk,peqk,ifail)

!       Display results
        Write (nout,99999) n, l, m, k, plek, pgtk, peqk
      End Do d_lp

99999 Format (1X,4I4,3F10.5)
    End Program g01blfe
```

## 9.2 Program Data

```
G01BLF Example Program Data
  10   2   5   1    : N, L, M, K
  40  10   3   2
 155  35 122  22
1000 444 500 220
```

## 9.3 Program Results

```
 G01BLF Example Program Results

    N   L   M   K     PLEK      PGTK      PEQK

   10   2   5   1   0.77778   0.22222   0.55556
   40  10   3   2   0.98785   0.01215   0.13664
  155  35 122  22   0.01101   0.98899   0.00779
 1000 444 500 220   0.42429   0.57571   0.04913
```