NAG Library Routine Document

G01SJF

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

G01SJF returns a number of the lower tail, upper tail and point probabilities for the binomial distribution.

2 Specification

```
SUBROUTINE G01SJF (LN, N, LP, P, LK, K, PLEK, PGTK, PEQK, IVALID, IFAIL)
INTEGER            LN, N(LN), LP, LK, K(LK), IVALID(*), IFAIL
REAL (KIND=nag_wp) P(LP), PLEK(*), PGTK(*), PEQK(*)
```

3 Description

Let $X = \{X_i : i = 1, 2, \ldots, m\}$ denote a vector of random variables each having a binomial distribution with parameters n_i and p_i ($n_i \geq 0$ and $0 < p_i < 1$). Then

$$\text{Prob}\{X_i = k_i\} = \binom{n_i}{k_i} {p_i}^{k_i}(1-p_i)^{n_i-k_i}, \quad k_i = 0, 1, \ldots, n_i.$$

The mean of the each distribution is given by $n_i p_i$ and the variance by $n_i p_i(1-p_i)$.

G01SJF computes, for given n_i, p_i and k_i, the probabilities: Prob$\{X_i \leq k_i\}$, Prob$\{X_i > k_i\}$ and Prob$\{X_i = k_i\}$ using an algorithm similar to that described in KnÏsel (1986) for the Poisson distribution.

The input arrays to this routine are designed to allow maximum flexibility in the supply of vector arguments by re-using elements of any arrays that are shorter than the total number of evaluations required. See Section 2.6 in the G01 Chapter Introduction for further information.

4 References

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

5 Arguments

1: LN – INTEGER *Input*

On entry: the length of the array N

Constraint: LN > 0.

2: N(LN) – INTEGER array *Input*

On entry: n_i, the first parameter of the binomial distribution with $n_i = \text{N}(j)$, $j = ((i-1) \bmod \text{LN}) + 1$, for $i = 1, 2, \ldots, \max(\text{LN}, \text{LP}, \text{LK})$.

Constraint: N$(j) \geq 0$, for $j = 1, 2, \ldots, \text{LN}$.

3: LP – INTEGER *Input*

On entry: the length of the array P

Constraint: LP > 0.

4: P(LP) – REAL (KIND=nag_wp) array *Input*

On entry: p_i, the second parameter of the binomial distribution with $p_i = \mathrm{P}(j)$, $j = ((i-1) \bmod \mathrm{LP}) + 1$.

Constraint: $0.0 < \mathrm{P}(j) < 1.0$, for $j = 1, 2, \ldots, \mathrm{LP}$.

5: LK – INTEGER *Input*

On entry: the length of the array K

Constraint: $\mathrm{LK} > 0$.

6: K(LK) – INTEGER array *Input*

On entry: k_i, the integer which defines the required probabilities with $k_i = \mathrm{K}(j)$, $j = ((i-1) \bmod \mathrm{LK}) + 1$.

Constraint: $0 \leq k_i \leq n_i$.

7: PLEK(*) – REAL (KIND=nag_wp) array *Output*

Note: the dimension of the array PLEK must be at least $\max(\mathrm{LN}, \mathrm{LP}, \mathrm{LK})$.

On exit: $\mathrm{Prob}\{X_i \leq k_i\}$, the lower tail probabilities.

8: PGTK(*) – REAL (KIND=nag_wp) array *Output*

Note: the dimension of the array PGTK must be at least $\max(\mathrm{LN}, \mathrm{LP}, \mathrm{LK})$.

On exit: $\mathrm{Prob}\{X_i > k_i\}$, the upper tail probabilities.

9: PEQK(*) – REAL (KIND=nag_wp) array *Output*

Note: the dimension of the array PEQK must be at least $\max(\mathrm{LN}, \mathrm{LP}, \mathrm{LK})$.

On exit: $\mathrm{Prob}\{X_i = k_i\}$, the point probabilities.

10: IVALID(*) – INTEGER array *Output*

Note: the dimension of the array IVALID must be at least $\max(\mathrm{LN}, \mathrm{LP}, \mathrm{LK})$.

On exit: IVALID(i) indicates any errors with the input arguments, with

IVALID(i) $= 0$
No error.

IVALID(i) $= 1$
On entry, $n_i < 0$.

IVALID(i) $= 2$
On entry, $p_i \leq 0.0$,
or $p_i \geq 1.0$.

IVALID(i) $= 3$
On entry, $k_i < 0$,
or $k_i > n_i$.

IVALID(i) $= 4$
On entry, n_i is too large to be represented exactly as a real number.

IVALID(i) $= 5$
On entry, the variance ($= n_i p_i (1 - p_i)$) exceeds 10^6.

11: IFAIL – INTEGER *Input/Output*

On entry: IFAIL must be set to 0, -1 or 1. If you are unfamiliar with this argument you should refer to Section 3.4 in How to Use the NAG Library and its Documentation 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 argument, 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: 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 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, at least one value of N, P or K was invalid.
Check IVALID for more information.

IFAIL $= 2$

On entry, array size $= \langle value \rangle$.
Constraint: LN > 0.

IFAIL $= 3$

On entry, array size $= \langle value \rangle$.
Constraint: LP > 0.

IFAIL $= 4$

On entry, array size $= \langle value \rangle$.
Constraint: LK > 0.

IFAIL $= -99$

An unexpected error has been triggered by this routine. Please contact NAG.

See Section 3.9 in How to Use the NAG Library and its Documentation for further information.

IFAIL $= -399$

Your licence key may have expired or may not have been installed correctly.

See Section 3.8 in How to Use the NAG Library and its Documentation for further information.

IFAIL $= -999$

Dynamic memory allocation failed.

See Section 3.7 in How to Use the NAG Library and its Documentation for further information.

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 Parallelism and Performance

G01SJF is not threaded in any implementation.

9 Further Comments

The time taken by G01SJF to calculate each probability depends on the variance ($= n_i p_i(1-p_i)$) and on k_i. For given variance, the time is greatest when $k_i \approx n_i p_i$ (= the mean), and is then approximately proportional to the square-root of the variance.

10 Example

This example reads a vector of values for n, p and k, and prints the corresponding probabilities.

10.1 Program Text

```
    Program g01sjfe
!     G01SJF Example Program Text

!     Mark 26 Release. NAG Copyright 2016.

!     .. Use Statements ..
      Use nag_library, Only: g01sjf, nag_wp
!     .. Implicit None Statement ..
      Implicit None
!     .. Parameters ..
      Integer, Parameter               :: nin = 5, nout = 6
!     .. Local Scalars ..
      Integer                          :: i, ifail, lk, ln, lout, lp
!     .. Local Arrays ..
      Real (Kind=nag_wp), Allocatable  :: p(:), peqk(:), pgtk(:), plek(:)
      Integer, Allocatable             :: ivalid(:), k(:), n(:)
!     .. Intrinsic Procedures ..
      Intrinsic                        :: max, mod, repeat
!     .. Executable Statements ..
      Write (nout,*) 'G01SJF Example Program Results'
      Write (nout,*)

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

!     Read in the input vectors
      Read (nin,*) ln
      Allocate (n(ln))
      Read (nin,*) n(1:ln)

      Read (nin,*) lp
      Allocate (p(lp))
      Read (nin,*) p(1:lp)

      Read (nin,*) lk
      Allocate (k(lk))
      Read (nin,*) k(1:lk)

!     Allocate memory for output
      lout = max(ln,lp,lk)
      Allocate (peqk(lout),pgtk(lout),plek(lout),ivalid(lout))

!     Calculate probability
      ifail = -1
      Call g01sjf(ln,n,lp,p,lk,k,plek,pgtk,peqk,ivalid,ifail)

      If (ifail==0 .Or. ifail==1) Then
!       Display titles
        Write (nout,*)                                                  &
          '    N         P         K       PLEK      PGTK      PEQK    IVALID'
        Write (nout,*) repeat('-',68)
```

```
!       Display results
        Do i = 1, lout
          Write (nout,99999) n(mod(i-1,ln)+1), p(mod(i-1,lp)+1),              &
            k(mod(i-1,lk)+1), plek(i), pgtk(i), peqk(i), ivalid(i)
        End Do
      End If

99999 Format (1X,I6,4X,F6.2,4X,I6,3(4X,F6.3),4X,I3)
    End Program g01sjfe
```

10.2 Program Data

```
G01SJF Example Program Data
4                         :: LN
4 19 100 2000             :: N
4                         :: LP
0.500 0.440 0.750 0.330 :: P
4                         :: LK
2 13 67 700               :: K
```

10.3 Program Results

```
 G01SJF Example Program Results

     N         P         K       PLEK      PGTK      PEQK     IVALID
 ------------------------------------------------------------------
      4      0.50         2      0.688     0.312     0.375          0
     19      0.44        13      0.991     0.009     0.019          0
    100      0.75        67      0.045     0.955     0.017          0
   2000      0.33       700      0.973     0.027     0.003          0
```