

# NAG Library Routine Document

## G01HBF

**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

G01HBF returns the upper tail, lower tail or central probability associated with a multivariate Normal distribution of up to ten dimensions.

### 2 Specification

```
FUNCTION G01HBF (TAIL, N, A, B, XMU, SIG, LDSIG, TOL, WK, LWK, IFAIL)
REAL (KIND=nag_wp) G01HBF
INTEGER                N, LDSIG, LWK, IFAIL
REAL (KIND=nag_wp) A(N), B(N), XMU(N), SIG(LDSIG,N), TOL, WK(LWK)
CHARACTER(1)          TAIL
```

### 3 Description

Let the vector random variable  $X = (X_1, X_2, \dots, X_n)^T$  follow an  $n$ -dimensional multivariate Normal distribution with mean vector  $\mu$  and  $n$  by  $n$  variance-covariance matrix  $\Sigma$ , then the probability density function,  $f(X : \mu, \Sigma)$ , is given by

$$f(X : \mu, \Sigma) = (2\pi)^{-(1/2)n} |\Sigma|^{-1/2} \exp\left(-\frac{1}{2}(X - \mu)^T \Sigma^{-1}(X - \mu)\right).$$

The lower tail probability is defined by:

$$P(X_1 \leq b_1, \dots, X_n \leq b_n : \mu, \Sigma) = \int_{-\infty}^{b_1} \cdots \int_{-\infty}^{b_n} f(X : \mu, \Sigma) dX_n \cdots dX_1.$$

The upper tail probability is defined by:

$$P(X_1 \geq a_1, \dots, X_n \geq a_n : \mu, \Sigma) = \int_{a_1}^{\infty} \cdots \int_{a_n}^{\infty} f(X : \mu, \Sigma) dX_n \cdots dX_1.$$

The central probability is defined by:

$$P(a_1 \leq X_1 \leq b_1, \dots, a_n \leq X_n \leq b_n : \mu, \Sigma) = \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} f(X : \mu, \Sigma) dX_n \cdots dX_1.$$

To evaluate the probability for  $n \geq 3$ , the probability density function of  $X_1, X_2, \dots, X_n$  is considered as the product of the conditional probability of  $X_1, X_2, \dots, X_{n-2}$  given  $X_{n-1}$  and  $X_n$  and the marginal bivariate Normal distribution of  $X_{n-1}$  and  $X_n$ . The bivariate Normal probability can be evaluated as described in G01HAF and numerical integration is then used over the remaining  $n - 2$  dimensions. In the case of  $n = 3$ , D01AJF is used and for  $n > 3$  D01FCF is used.

To evaluate the probability for  $n = 1$  a direct call to G01EAF is made and for  $n = 2$  calls to G01HAF are made.

### 4 References

Kendall M G and Stuart A (1969) *The Advanced Theory of Statistics (Volume 1)* (3rd Edition) Griffin

## 5 Arguments

- 1: TAIL – CHARACTER(1) *Input*  
*On entry:* indicates which probability is to be returned.  
 TAIL = 'L'  
     The lower tail probability is returned.  
 TAIL = 'U'  
     The upper tail probability is returned.  
 TAIL = 'C'  
     The central probability is returned.  
*Constraint:* TAIL = 'L', 'U' or 'C'.
- 2: N – INTEGER *Input*  
*On entry:*  $n$ , the number of dimensions.  
*Constraint:*  $1 \leq N \leq 10$ .
- 3: A(N) – REAL (KIND=nag\_wp) array *Input*  
*On entry:* if TAIL = 'C' or 'U', the lower bounds,  $a_i$ , for  $i = 1, 2, \dots, n$ .  
 If TAIL = 'L', A is not referenced.
- 4: B(N) – REAL (KIND=nag\_wp) array *Input*  
*On entry:* if TAIL = 'C' or 'L', the upper bounds,  $b_i$ , for  $i = 1, 2, \dots, n$ .  
 If TAIL = 'U', B is not referenced.  
*Constraint:* if TAIL = 'C',  $A(i) < B(i)$ , for  $i = 1, 2, \dots, n$ .
- 5: XMU(N) – REAL (KIND=nag\_wp) array *Input*  
*On entry:*  $\mu$ , the mean vector of the multivariate Normal distribution.
- 6: SIG(LDSIG, N) – REAL (KIND=nag\_wp) array *Input*  
*On entry:*  $\Sigma$ , the variance-covariance matrix of the multivariate Normal distribution. Only the lower triangle is referenced.  
*Constraint:*  $\Sigma$  must be positive definite.
- 7: LDSIG – INTEGER *Input*  
*On entry:* the first dimension of the array SIG as declared in the (sub)program from which G01HBF is called.  
*Constraint:* LDSIG  $\geq$  N.
- 8: TOL – REAL (KIND=nag\_wp) *Input*  
*On entry:* if  $n > 2$  the relative accuracy required for the probability, and if the upper or the lower tail probability is requested then TOL is also used to determine the cut-off points, see Section 7.  
 If  $n = 1$ , TOL is not referenced.  
*Suggested value:* TOL = 0.0001.  
*Constraint:* if  $N > 1$ , TOL  $> 0.0$ .

- 9: WK(LWK) – REAL (KIND=nag\_wp) array Workspace  
 10: LWK – INTEGER Input

*On entry:* the length of workspace provided in array WK. This workspace is used by the numerical integration routines D01AJF for  $n = 3$  and D01FCF for  $n > 3$ .

If  $n = 3$ , then the maximum number of sub-intervals used by D01AJF is  $LWK/4$ . Note, however, increasing LWK above 1000 will not increase the maximum number of sub-intervals above 250.

If  $n > 3$  the maximum number of integrand evaluations used by D01FCF is  $\alpha(LWK/n - 1)$ , where  $\alpha = 2^{n-2} + 2(n - 2)^2 + 2(n - 2) + 1$ .

If  $n = 1$  or  $2$ , then WK will not be used.

*Suggested value:* 2000 if  $n > 3$  and 1000 if  $n = 3$ .

*Constraints:*

- if  $N \leq 2$ ,  $LWK \geq 1$ ;
- if  $N \geq 3$ ,  $LWK \geq 4 \times N$ .

- 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, because for this routine the values of the output arguments may be useful even if  $IFAIL \neq 0$  on exit, the recommended value is  $-1$ . **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).

**Note:** G01HBF may return useful information for one or more of the following detected errors or warnings.

Errors or warnings detected by the routine:

If on exit  $IFAIL = 1, 2$  or  $3$ , then G01HBF returns zero.

$IFAIL = 1$

On entry,  $LDSIG = \langle value \rangle$  and  $N = \langle value \rangle$ .  
 Constraint:  $LDSIG \geq N$ .

On entry,  $LWK = \langle value \rangle$ .  
 Constraint: if  $N \leq 2$ ,  $LWK \geq 1$ .

On entry,  $LWK = \langle value \rangle$ .  
 Constraint: if  $N \geq 3$ ,  $LWK \geq 4 \times N$ .

On entry,  $N = \langle value \rangle$ .  
 Constraint:  $1 \leq N \leq 10$ .

On entry, TAIL is not valid:  $TAIL = \langle value \rangle$ .

On entry,  $TOL = \langle value \rangle$ .  
 Constraint:  $TOL > 0.0$ .

IFAIL = 2

On entry, the  $\langle value \rangle$  value in B is less than or equal to the corresponding value in A.

IFAIL = 3

On entry, SIG is not positive definite.

IFAIL = 4

Full accuracy not achieved, relative accuracy =  $\langle value \rangle$ . A larger value of TOL can be tried or the length of the workspace increased. The returned value is an approximation to the required result.

IFAIL = 5

Accuracy requested by TOL is too strict: TOL =  $\langle value \rangle$ . Round-off error has prevented the requested accuracy from being achieved; a larger value of TOL can be tried. The returned value will be an approximation to the required result.

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

The accuracy should be as specified by TOL. When on exit IFAIL = 4 the approximate accuracy achieved is given in the error message. For the upper and lower tail probabilities the infinite limits are approximated by cut-off points for the  $n - 2$  dimensions over which the numerical integration takes place; these cut-off points are given by  $\Phi^{-1}(\text{TOL}/(10 \times n))$ , where  $\Phi^{-1}$  is the inverse univariate Normal distribution function.

## 8 Parallelism and Performance

G01HBF is not threaded in any implementation.

## 9 Further Comments

The time taken is related to the number of dimensions, the range over which the integration takes place ( $b_i - a_i$ , for  $i = 1, 2, \dots, n$ ) and the value of  $\Sigma$  as well as the accuracy required. As the numerical integration does not take place over the last two dimensions speed may be improved by arranging  $X$  so that the largest ranges of integration are for  $X_{n-1}$  and  $X_n$ .

## 10 Example

This example reads in the mean and covariance matrix for a multivariate Normal distribution and computes and prints the associated central probability.

**10.1 Program Text**

```

Program g01hbfe

!      G01HBF Example Program Text

!      Mark 26 Release. NAG Copyright 2016.

!      .. Use Statements ..
Use nag_library, Only: g01hbf, nag_wp
!      .. Implicit None Statement ..
Implicit None
!      .. Parameters ..
Integer, Parameter          :: nin = 5, nout = 6
!      .. Local Scalars ..
Real (Kind=nag_wp)         :: prob, tol
Integer                    :: i, ifail, ldsig, lwk, n
Character (1)              :: tail
!      .. Local Arrays ..
Real (Kind=nag_wp), Allocatable :: a(:), b(:), sig(:, :), wk(:), xmu(:)
!      .. Intrinsic Procedures ..
Intrinsic                  :: max
!      .. Executable Statements ..
Write (nout,*) 'G01HBF Example Program Results'
Write (nout,*)

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

!      Read in the problem size
Read (nin,*) n, tol, tail

!      Use more workspace, unless N is very large
!      as LWK also defines the number of sub-intervals
lwk = max(2000,4*n)

      ldsig = n
      Allocate (a(n),b(n),xmu(n),sig(ldsig,n),wk(lwk))

!      Read in the means
Read (nin,*) xmu(1:n)

!      Read in the variance covariance matrix
Read (nin,*)(sig(i,1:n),i=1,n)

!      Read in bounds
If (tail=='C' .Or. tail=='c' .Or. tail=='U' .Or. tail=='u') Then
  Read (nin,*) a(1:n)
End If
If (tail=='L' .Or. tail=='l' .Or. tail=='L' .Or. tail=='l') Then
  Read (nin,*) b(1:n)
End If

!      Calculate probability
ifail = -1
prob = g01hbf(tail,n,a,b,xmu,sig,ldsig,tol,lwk,lwk,ifail)
If (ifail/=0) Then
  If (ifail<=3) Then
    Go To 100
  End If
End If

!      Display results
Write (nout,99999) 'Multivariate Normal probability =', prob

100  Continue

99999 Format (1X,A,F7.4)
End Program g01hbfe

```

## 10.2 Program Data

```
G01HBF Example Program Data
 4 0.00001 'c'
 0.0 0.0 0.0 0.0
 1.0 0.9 0.9 0.9
 0.9 1.0 0.9 0.9
 0.9 0.9 1.0 0.9
 0.9 0.9 0.9 1.0
-2.0 -2.0 -2.0 -2.0
 2.0 2.0 2.0 2.0
```

## 10.3 Program Results

```
G01HBF Example Program Results
```

```
Multivariate Normal probability = 0.9142
```

---