G01HBF – NAG Fortran Library Routine Document

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

```
      real FUNCTION G01HBF(TAIL, N, A, B, XMU, SIG, LDSIG, TOL, WK, LWK,
     1                     IFAIL)
      INTEGER              N, LDSIG, LWK, IFAIL
      real                 A(N), B(N), XMU(N), SIG(LDSIG,N), TOL, WK(LWK)
      CHARACTER*1          TAIL
```

3 Description

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

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

The lower tail probability is defined by:

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

The upper tail probability is defined by:

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

The central probability is defined by:

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

To evaluate the probability for $n \geq 3$, the probability density function of $X_1, X_2, \ldots, X_n$ is considered as the product of the conditional probability of $X_1, X_2, \ldots, X_{n-2}$ 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 $n = 3$ the routine D01AJF is used and for $n > 3$ the routine 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

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

5 Parameters

1: TAIL — CHARACTER*1 *Input*

On entry: indicates which probability is to be returned.

If TAIL = 'L' the lower tail probability is returned.

If TAIL = 'U' the upper tail probability is returned.

If TAIL = 'C' the central probability is returned.

Constraint: TAIL = 'C', 'L' or 'U'.

2: N — INTEGER *Input*

On entry: the number of dimensions, n.

Constraint: $1 \leq \mathrm{N} \leq 10$.

3: A(N) — ***real*** array *Input*

On entry: if TAIL = 'C' or 'U' the lower bounds, a_i, for $i = 1, 2, \ldots, n$.

If TAIL = 'L', A is not referenced.

4: B(N) — ***real*** array *Input*

On entry: if TAIL = 'C' or 'L' the upper bounds, b_i, for $i = 1, 2, \ldots, n$.

If TAIL = 'U' B is not referenced.

Constraint: if TAIL = 'C', $\mathrm{A}(i) < \mathrm{B}(i)$, for $i = 1, 2, \ldots, n$.

5: XMU(N) — ***real*** array *Input*

On entry: the mean vector, μ, of the multivariate Normal distribution.

6: SIG(LDSIG,N) — ***real*** array *Input*

On entry: the variance-covariance matrix, Σ, of the multivariate Normal distribution. Only the lower triangle is referenced.

Constraint: Σ 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: $\mathrm{LDSIG} \geq \mathrm{N}$.

8: TOL — ***real*** *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 $\mathrm{N} > 1$, $\mathrm{TOL} > 0.0$.

9: WK(LWK) — ***real*** 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 $\alpha(\mathrm{LWK}/n - 1)$, 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:

LWK ≥ 1, for N $= 1$ or 2,
LWK $\geq 4 \times$ N, for N ≥ 3.

11: IFAIL — INTEGER *Input/Output*

On entry: IFAIL must be set to 0, -1 or 1. Users who are unfamiliar with this parameter should refer to Chapter P01 for details.

On exit: IFAIL = 0 unless the routine detects an error or gives a warning (see Section 6).

For this routine, because the values of output parameters may be useful even if IFAIL $\neq 0$ on exit, users are recommended to set IFAIL to -1 before entry. **It is then essential to test the value of IFAIL on exit**.

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 specified by the routine:

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

IFAIL = 1

On entry, N < 1,
or N > 10,
or LDSIG $<$ N,
or TAIL $\neq$ 'L', 'U' or 'C',
or N > 1 and TOL ≤ 0.0,
or LWK is too small.

IFAIL = 2

On entry, TAIL = 'C' and A(i) $\geq$ B(i), for some $i = 1, 2, \ldots, n$.

IFAIL = 3

On entry, Σ is not positive-definite, i.e., is not a correct variance-covariance matrix.

IFAIL = 4

The requested accuracy has not been achieved, a larger value of TOL should be tried or the length of the workspace should be increased. The returned value will be an approximation to the required result.

IFAIL = 5

Round-off error prevents the requested accuracy from being achieved; a larger value of TOL should be tried. The returned value will be an approximation to the required result. This error will only occur if $n = 3$.

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 $\Phi^{-1}(\mathrm{TOL}/(10\times n))$, where Φ^{-1} is the inverse univariate Normal distribution function.

8 Further Comments

The time taken is related to both the number of dimensions, the range over which the integration takes place ($b_i - a_i$, for $i = 1, 2, \ldots, n$) and the value of Σ 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.

9 Example

The mean and variance of a four-dimensional multivariate Normal distribution are input and a central probability computed and printed.

9.1 Program Text

Note. The listing of the example program presented below uses bold italicised terms to denote precision-dependent details. Please read the Users' Note for your implementation to check the interpretation of these terms. As explained in the Essential Introduction to this manual, the results produced may not be identical for all implementations.

```
*     G01HBF Example Program Text
*     Mark 15 Release. NAG Copyright 1991.
*     .. Parameters ..
      INTEGER          NIN, NOUT
      PARAMETER        (NIN=5,NOUT=6)
      INTEGER          NMAX, LWK
      PARAMETER        (NMAX=10,LWK=2000)
*     .. Local Scalars ..
      real             PROB, TOL
      INTEGER          I, IFAIL, J, N
      CHARACTER        TAIL
*     .. Local Arrays ..
      real             A(NMAX), B(NMAX), SIG(NMAX,NMAX), WK(LWK),
     +                 XMU(NMAX)
*     .. External Functions ..
      real             G01HBF
      EXTERNAL         G01HBF
*     .. Executable Statements ..
      WRITE (NOUT,*) 'G01HBF Example Program Results'
*     Skip heading in data file
      READ (NIN,*)
      READ (NIN,*) N, TOL, TAIL
      IF (N.LE.NMAX) THEN
         READ (NIN,*) (XMU(J),J=1,N)
         DO 20 I = 1, N
            READ (NIN,*) (SIG(I,J),J=1,N)
   20    CONTINUE
         IF (TAIL.EQ.'C' .OR. TAIL.EQ.'c' .OR. TAIL.EQ.'U' .OR. TAIL.EQ.
     +       'u') READ (NIN,*) (A(J),J=1,N)
         IF (TAIL.EQ.'C' .OR. TAIL.EQ.'c' .OR. TAIL.EQ.'L' .OR. TAIL.EQ.
     +       'l') READ (NIN,*) (B(J),J=1,N)
         IFAIL = -1
*
```

```
         PROB = G01HBF(TAIL,N,A,B,XMU,SIG,NMAX,TOL,WK,LWK,IFAIL)
*
         IF (IFAIL.EQ.0 .OR. IFAIL.GT.3) THEN
            WRITE (NOUT,*)
            WRITE (NOUT,99999) 'Multivariate Normal probability = ',
     +        PROB
         END IF
      END IF
      STOP
*
99999 FORMAT (1X,A,F6.4)
      END
```

9.2 Program Data

```
G01HBF Example Program Data
 4  0.0001 '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
```

9.3 Program Results

```
 G01HBF Example Program Results

 Multivariate Normal probability = 0.9142
```