NAG Library Routine Document

G05LXF

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

G05LXF sets up a reference vector and generates an array of pseudorandom numbers from a multivariate Student's t distribution with ν degrees of freedom, mean vector a and covariance matrix $\frac{\nu}{\nu-2}C$.

2 Specification

```
      SUBROUTINE G05LXF(MODE, DF, M, XMU, C, LDC, N, X, LDX, IGEN, ISEED, R,
     1                  LR, IFAIL)
      INTEGER           MODE, DF, M, LDC, N, LDX, IGEN, ISEED(4), LR, IFAIL
      double precision  XMU(M), C(LDC,M), X(LDX,M), R(LR)
```

3 Description

When the covariance matrix is nonsingular (i.e., strictly positive-definite), the distribution has probability density function

$$f(x) = \frac{\Gamma\left(\frac{(\nu+m)}{2}\right)}{(\pi v)^{m/2}\Gamma(\nu/2)|C|^{\frac{1}{2}}}\left[1+\frac{(x-a)^{\mathrm{T}}C^{-1}(x-a)}{\nu}\right]^{\frac{-(\nu+m)}{2}}$$

where m is the number of dimensions, ν is the degrees of freedom, a is the vector of means, x is the vector of positions and $\frac{\nu}{\nu-2}C$ is the covariance matrix.

The routine returns the value

$$x = a + \sqrt{\frac{v}{s}}z$$

where z is generated by G05LYF from a Normal distribution with mean zero and covariance matrix C and s is generated by G05LCF from a χ^2-distribution with ν degrees of freedom.

One of the initialization routines G05KBF (for a repeatable sequence if computed sequentially) or G05KCF (for a non-repeatable sequence) must be called prior to the first call to G05LXF.

4 References

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

Wilkinson J H (1965) *The Algebraic Eigenvalue Problem* Oxford University Press, Oxford

5 Parameters

1: MODE – INTEGER *Input*

On entry: selects the operation to be performed.

MODE $= 0$
: Initialize and generate random numbers.

MODE $= 1$
: Initialize only (i.e., set up reference vector).

MODE $= 2$
Generate random numbers using previously set up reference vector.

Constraint: MODE $= 0$, 1 or 2.

2: DF – INTEGER *Input*

On entry: ν, the number of degrees of freedom of the distribution.

Constraint: DF ≥ 3.

3: M – INTEGER *Input*

On entry: m, the number of dimensions of the distribution.

Constraint: M > 0.

4: XMU(M) – ***double precision*** array *Input*

On entry: a, the vector of means of the distribution.

5: C(LDC,M) – ***double precision*** array *Input*

On entry: matrix which, along with DF, defines the covariance of the distribution. Only the upper triangle need be set.

Constraint: C must be positive semi-definite to ***machine precision***

6: LDC – INTEGER *Input*

On entry: the first dimension of the array C as declared in the (sub)program from which G05LXF is called.

Constraint: LDC $\geq$ M.

7: N – INTEGER *Input*

On entry: n, the number of random variates required.

Constraint: N ≥ 1.

8: X(LDX,M) – ***double precision*** array *Output*

On exit: the array of pseudorandom multivariate Student's t vectors generated by the routine, with X(i,j) holding the jth dimension for the ith variate.

9: LDX – INTEGER *Input*

On entry: the first dimension of the array X as declared in the (sub)program from which G05LXF is called.

Constraint: LDX $\geq$ N.

10: IGEN – INTEGER *Input*

On entry: must contain the identification number for the generator to be used to return a pseudorandom number and should remain unchanged following initialization by a prior call to G05KBF or G05KCF.

11: ISEED(4) – INTEGER array *Communication Array*

On entry: contains values which define the current state of the selected generator.

On exit: contains updated values defining the new state of the selected generator.

12: R(LR) – ***double precision*** array *Input/Output*

On entry: if MODE $= 2$, the reference vector as set up by G05LXF in a previous call with MODE $= 0$ or 1.

On exit: if MODE $= 0$ or 1, the reference vector that can be used in subsequent calls to G05LXF with MODE $= 2$.

13: LR – INTEGER *Input*

On entry: the dimension of the array R as declared in the (sub)program from which G05LXF is called. If MODE $= 2$, it must be the same as the value of LR specified in the prior call to G05LXF with MODE $= 0$ or 1.

Constraint: LR $> \mathrm{M} \times (\mathrm{M} + 1) + 1$.

14: 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.

On exit: IFAIL $= 0$ unless the routine detects an error (see Section 6).

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.**

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, MODE $\neq 0$, 1 or 2.

IFAIL $= 2$

On entry, M < 1.

IFAIL $= 3$

On entry, DF ≤ 2.

IFAIL $= 5$

The covariance matrix C is not positive semi-definite to ***machine precision***.

IFAIL $= 6$

On entry, LDC $<$ M.

IFAIL $= 7$

On entry, N < 1.

IFAIL $= 9$

On entry, LDX $<$ N.

IFAIL = 10

On entry, invalid value for IGEN. IGEN must be the same as the value as specified in the prior call to G05LXF with MODE = 0 or 1.

IFAIL = 12

The reference vector R has been corrupted or M has changed since R was set up in a previous call with MODE = 0 or 1.

IFAIL = 13

On entry, LR < M × (M + 1) + 1.

7 Accuracy

The maximum absolute error in LL^{T}, and hence in the covariance matrix of the resulting vectors, is less than $(m\epsilon + (m+3)\epsilon/2)$ times the maximum element of C, where ϵ is the ***machine precision***. Under normal circumstances, the above will be small compared to sampling error.

8 Further Comments

The time taken by G05LXF is of order nm^3.

It is recommended that the diagonal elements of C should not differ too widely in order of magnitude. This may be achieved by scaling the variables if necessary. The actual matrix decomposed is $C + E = LL^{\mathrm{T}}$, where E is a diagonal matrix with small positive diagonal elements. This ensures that, even when C is singular, or nearly singular, the Cholesky factor L corresponds to a positive-definite covariance matrix that agrees with C within ***machine precision***.

9 Example

This example prints ten pseudorandom observations from a multivariate Student's t-distribution ten degrees of freedom, means vector

$$\begin{bmatrix} 1.0 \\ 2.0 \\ -3.0 \\ 0.0 \end{bmatrix}$$

and C matrix

$$\begin{bmatrix} 1.69 & 0.39 & -1.86 & 0.07 \\ 0.39 & 98.01 & -7.07 & -0.71 \\ -1.86 & -7.07 & 11.56 & 0.03 \\ 0.07 & -0.71 & 0.03 & 0.01 \end{bmatrix},$$

generated by G05LXF. All ten observations are generated by a single call to G05LXF with MODE = 0. The random number generator is initialized by G05KBF.

9.1 Program Text

```
*     G05LXF Example Program Text
*     Mark 21 Release. NAG Copyright 2004.
*     .. Parameters ..
      INTEGER          NOUT, LDC, LDX, LR
      PARAMETER        (NOUT=6,LDC=5,LDX=100,LR=LDC*LDC+LDC+2)
*     .. Local Scalars ..
      INTEGER          DF, I, IFAIL, IGEN, J, M, MODE, N
*     .. Local Arrays ..
      DOUBLE PRECISION C(LDC,LDC), R(LR), X(LDX,LDC), XMU(LDC)
      INTEGER          ISEED(4)
*     .. External Subroutines ..
```

```
      EXTERNAL         G05KBF, G05LXF
*     .. Executable Statements ..
      CONTINUE
*
      WRITE (NOUT,*) 'G05LXF Example Program Results'
      WRITE (NOUT,*)
*
*     Initialize the seed to a repeatable sequence
      ISEED(1) = 1762543
      ISEED(2) = 9324783
      ISEED(3) = 42344
      ISEED(4) = 742355
*
*     Choose the random generator to use
      IGEN = 1
*
*     Initialize the random generator
      CALL G05KBF(IGEN,ISEED)
*
*     Set the number of variables and variates
      M = 4
      N = 10
*
*     Input the upper triangle portion of the covariance matrix
      C(1,1) = 1.69D0
      C(1,2) = 0.39D0
      C(1,3) = -1.86D0
      C(1,4) = 0.07D0
      C(2,2) = 98.01D0
      C(2,3) = -7.07D0
      C(2,4) = -0.71D0
      C(3,3) = 11.56D0
      C(3,4) = 0.03D0
      C(4,4) = 0.01D0
*
*     Input the means
      XMU(1) = 1.0D0
      XMU(2) = 2.0D0
      XMU(3) = -3.0D0
      XMU(4) = 0.0D0
*
*     Set the degrees of freedom
      DF = 10
*
      MODE = 0
      IFAIL = 1
*
*     Set up reference vector and generate N numbers
      CALL G05LXF(MODE,DF,M,XMU,C,LDC,N,X,LDX,IGEN,ISEED,R,LR,IFAIL)
*
      IF (IFAIL.EQ.0) THEN
*        Display the results
         DO 20 I = 1, N
            WRITE (NOUT,99999) (X(I,J),J=1,M)
   20    CONTINUE
      ELSE
         WRITE (NOUT,99998) IFAIL
      END IF
*
99999 FORMAT (1X,10F10.4)
99998 FORMAT (1X,' ** G05LXF returned with IFAIL = ',I5)
      END
```

9.2 Program Data

None.

9.3 Program Results

```
 G05LXF Example Program Results

     3.0999   -5.3200   -6.8459    0.1218
     0.1668    7.0595   -2.7861   -0.1162
     0.9310   11.5035    0.5182   -0.0219
    -0.7092    1.2452   -1.0941   -0.0633
     0.7267  -10.2979    0.0582    0.0576
     1.1014   13.6137   -3.3920   -0.0590
     0.0400   -1.6104    2.3485    0.0032
     2.1146    7.9958  -12.5358   -0.0471
     2.7721  -13.4943   -2.4307    0.1480
     0.0397  -18.3675    0.9874    0.1178
```