NAG Library Routine Document

G05RYF

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

G05RYF 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 G05RYF(MODE, N, DF, M, XMU, C, LDC, R, LR, STATE, X, LDX,
     1                  IFAIL)
      INTEGER           MODE, N, DF, M, LDC, LR, STATE(lstate), LDX, IFAIL
      double precision  XMU(M), C(LDC,M), R(LR), X(LDX,M)
```

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

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

where z is generated by G05SKF from a Normal distribution with mean zero and covariance matrix C and s is generated by G05SDF from a χ^2-distribution with ν degrees of freedom.

One of the initialization routines G05KFF (for a repeatable sequence if computed sequentially) or G05KGF (for a non-repeatable sequence) must be called prior to the first call to G05RYF.

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: a code for selecting the operation to be performed by the routine.

MODE $= 0$
Set up reference vector only.

MODE $= 1$
Generate variates using reference vector set up in a prior call to G05RYF.

MODE $= 2$
Set up reference vector and generate two-way table.

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

2: N – INTEGER *Input*

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

Constraint: N ≥ 0.

3: DF – INTEGER *Input*

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

Constraint: DF ≥ 3.

4: M – INTEGER *Input*

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

Constraint: M > 0.

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

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

6: 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***

7: LDC – INTEGER *Input*

On entry: the first dimension of the array C as declared in the (sub)program from which G05RYF is called.

Constraint: LDC $\geq$ M.

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

On entry: if MODE $= 1$, the reference vector as set up by G05RYF in a previous call with MODE $= 0$ or 2.

On exit: if MODE $= 0$ or 2, the reference vector that can be used in subsequent calls to G05RYF with MODE $= 1$.

9: LR – INTEGER *Input*

On entry: the dimension of the array R as declared in the (sub)program from which G05RYF is called. If MODE $= 1$, it must be the same as the value of LR specified in the prior call to G05RYF with MODE $= 0$ or 2.

Constraint: LR $\geq$ M $\times$ (M $+ 1) + 2$.

10: STATE($lstate$) – INTEGER array *Communication Array*

Note: $lstate =$ LSTATE, the dimension of STATE as supplied to the initialization routine G05KFF or G05KGF.

On entry: contains information on the selected base generator and its current state.

On exit: contains updated information on the state of the generator.

11: 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.

12: LDX – INTEGER *Input*

On entry: the first dimension of the array X as declared in the (sub)program from which G05RYF is called.

Constraint: LDX $\geq$ N.

13: 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, N < 1.

IFAIL $= 3$

On entry, DF ≤ 2.

IFAIL $= 4$

On entry, M < 1.

IFAIL $= 6$

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

IFAIL $= 7$

On entry, LDC $<$ M.

IFAIL $= 8$

The reference vector R has been corrupted or M has changed since R was set up in a previous call to G05RYF with MODE $= 0$ or 2.

IFAIL $= 9$

On entry, LR $\leq$ M $\times$ (M $+ 1) + 1$.

IFAIL = 10

On entry, STATE vector was not initialized or has been corrupted.

IFAIL = 12

On entry, LDX < N.

7 Accuracy

Not applicable.

8 Further Comments

The time taken by G05RYF 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 with 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 G05RYF. All ten observations are generated by a single call to G05RYF with MODE = 2. The random number generator is initialized by G05KFF.

9.1 Program Text

```
*     G05RYF Example Program Text
*     Mark 22 Release. NAG Copyright 2006.
*     .. Parameters ..
      INTEGER          NOUT
      PARAMETER        (NOUT=6)
      INTEGER          MSEED, MSTATE, LDC, LDX, LR
      PARAMETER        (MSEED=1,MSTATE=633,LDC=5,LDX=100,
     +                 LR=LDC*LDC+LDC+2)
*     .. Local Scalars ..
      INTEGER          DF, GENID, I, IFAIL, J, LSEED, LSTATE, M, MODE,
     +                 N, SUBID
*     .. Local Arrays ..
      DOUBLE PRECISION C(LDC,LDC), R(LR), X(LDX,LDC), XMU(LDC)
      INTEGER          SEED(MSEED), STATE(MSTATE)
*     .. External Subroutines ..
      EXTERNAL         G05KFF, G05RYF
*     .. Executable Statements ..
      WRITE (NOUT,*) 'G05RYF Example Program Results'
      WRITE (NOUT,*)
*     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
*     Initialize the seed
      SEED(1) = 1762543
*     GENID and SUBID identify the base generator
      GENID = 1
      SUBID = 1
*     Initialize the generator to a repeatable sequence
      LSTATE = MSTATE
      LSEED = MSEED
      IFAIL = 1
      CALL G05KFF(GENID,SUBID,SEED,LSEED,STATE,LSTATE,IFAIL)
      IF (IFAIL.NE.0) THEN
         WRITE (NOUT,99998) IFAIL
         GO TO 40
      END IF
*     Set up reference vector and generate N numbers
*     Choose MODE = 2
      MODE = 2
      IFAIL = 1
      CALL G05RYF(MODE,N,DF,M,XMU,C,LDC,R,LR,STATE,X,LDX,IFAIL)
      IF (IFAIL.NE.0) THEN
         WRITE (NOUT,99997) IFAIL
         GO TO 40
      END IF
*     Display the variates
      DO 20 I = 1, N
         WRITE (NOUT,99999) (X(I,J),J=1,M)
   20 CONTINUE
*
   40 CONTINUE
*
99999 FORMAT (1X,10F10.4)
99998 FORMAT (1X,' ** G05KFF returned with IFAIL = ',I5)
99997 FORMAT (1X,' ** G05RYF returned with IFAIL = ',I5)
      END
```

9.2 Program Data

None.

9.3 Program Results

```
 G05RYF Example Program Results

     1.4957  -15.6226   -3.8101    0.1294
    -1.0827   -6.7473    0.6696   -0.0391
     2.1369    6.3861   -5.7413    0.0140
     2.2481  -16.0417   -1.0982    0.1641
    -0.2550    3.5166   -0.2541   -0.0592
```

```
     0.9731   -4.3553   -4.4181    0.0043
     0.7098   -3.4281    1.1741    0.0586
     1.8827   23.2619    1.5140   -0.0704
     0.9904   22.7479    0.1811   -0.0893
     1.5026    2.7753   -2.2805   -0.0112
```