TRAINING & RESEARCH

GRADUATE INSTITUTE OF STATISTICS

TEXAS A&M UNIVERSITY · COLLEGE STATION

WISHART VARIATE GENERATORS

by

W. B. Smith and R. R. Hocking
Texas A&M University

Institute of Statistics
Texas A&M University

Technical Report # 8

National Aeronautics and Space Administration
Research Grant NGR 44-001-095

July 1971

WISHART VARIATE GENERATORS

by

W. B. Smith and R. R. Hocking
Institute of Statistics
Texas A&M University

Language

ASA Standard FORTRAN

Description and Purpose

By using the Bartlett (1933) decomposition subroutine WSHRT generates the matrix of sample variances and covariances from a random sample of size n distributed as a p-variate normal, $N_p(\mu, \Sigma)$. The joint distribution of these sample variances and covariances is called the Wishart distribution.

By generating p(p-1)/2 independent variates $a_{ij}(1 \leq j < i \leq p)$ distributed identically as $N_1(0, 1)$ and p-variates $a_{ii}(1 \leq i \leq p)$ independently distributed as χ^2_{N-i+1}, form the p x p matrix $A = (a_{ij})$, where $a_{ij} = 0(i < j)$. The variate transformation $B = (b_{ij}) = AA'$, with Jacobian

$$J^{-1} = 2^p \prod_{i=1}^{p} a_{ii}^{p-i+1} , \qquad (1)$$

yields as the joint distribution of the unique b_{ij}(B is symmetric),

$$\frac{|B|^{(N-p-1)/2} \exp(-\frac{1}{2} \operatorname{tr}(B))}{2^{Np/2} \pi^{p(p-1)/4} \prod_{i=1}^{p} \Gamma(\frac{N-i+1}{2})} . \qquad (2)$$

Equation (2) corresponds to the joint density function of a sample variance-covariance matrix constructed from a random sample of size N from $N_p(0, I)$.

By decomposing Σ into the product of a lower triangular matrix D and its transpose (see Faddeva (1959) for a set of recursion formulas), the unique elements of the matrix (SA) = D B D' will have a Wishart distribution with parameters N and Σ. It should be noted that if μ is known then N = n, otherwise N = n - 1.

Wijsman (1957) arrives [illegible] above decomposition by using orthogonal matrices depending on certain random vectors, whereas, Kshirsagar (1959) arrives at the same result by transformations rather than by densities and Jacobians. To generate the standard univariate normal deviates required, the Box and Muller (1958) formulas are used, while χ^2 variates are generated by means of the Wilson and Hilferty (1931) approximation. Reliable use of the Wilson-Hilferty χ^2 deviates requires the degrees of freedom to exceed 20, but the authors have had satisfactory Wishart variates from much smaller values. Each formula, of course, depends upon the availability of a reliable uniform (i.e. rectangular (0, 1)) random number generator. The authors have had unsatisfactory results with the generator made available by the computer company and have resorted to the use of a subroutine (called RANDC) written by a colleague, Dr. C. E. Gates. The WSHRT output subroutine is highly dependent upon the input variate and any substitution of a local uniform pseudo random number generator should be preceded by careful checking of the reliability of that generator.

Structure

The subroutine definition statement is:

SUBROUTINE WSHRT (D, N, NP, SA)

Formal parameters

D	Real Array	input:	corresponds to D in theory, i.e. $DD' = \Sigma$.
N	Integer Scalar	input:	$N = \begin{cases} n, & \text{if } \mu \text{ known} \\ n-1, & \text{if } \mu \text{ unknown} \end{cases}$.
NP	Integer Scalar	input:	corresponds to p in theory, the dimension of the normal variate.
SA	Real Array	output:	corresponds to SA in theory.

Note: The input D is used instead of Σ since each cycle would require the recomputation of D.

Additional Comment

The above technique for generating the sample covariance matrix (Wishart distribution) requires only $p(p+1)/2$ standard univariate normal deviates, whereas, straightforward generation of such a random sample would require np standard normal variates. For small sample sizes ($n \leq 10$) and moderate sized p a straightforward calculation of the sample variance-covariance matrix is recommended. However, for the larger sample sizes the above algorithm would be recommended as the Wilson-Hilferty approximation to χ^2 variates would be valid.

Execution time of subroutine WSHRT is approximately .033 seconds on an IBM 360/65 computer using a WATFIV computer (developed by the University of Waterloo and modified for local use by the Data Processing Center, Texas A&M University).

References

Bartlett, M. S. (1933). On the theory of statistical regression. Proc. Roy. Soc. Edinburgh, 53, 260-283.

Box, G. E. P. and Muller, M. E. (1958). A note on the generation of random normal deviates, Ann. Math. Statist., 29, 610-611.

Faddeva, V. N. (1959). Computational Method of Linear Algebra. New York, Dover.

Kshirsagar, A. M. (1959). Bartlett decomposition and Wishart distribution. Ann. Math. Statist., 30, 239-241.

Wijsman, R. A. (1957). Random orthogonal transformations. Ann. Math. Statist., 28, 415-422.

Wilson, F. S. and Hilferty, M. M. (1931). The distribution of chi-square. Proc. Nat. Acad. Sci., U.S.A., 17, 684-688.

```
      SUBROUTINE WSHRT (D, N, NP, SA)
      IMPLICIT REAL*8(A-H,O-Z)
C
C     WISHART VARIATE GENERATOR
C
      DIMENSION B(NP,NP),C(NP,NP),D(NP,NP),SA(NP,NP),A(NP,NP)
C
C     CHSQF IS A FUNCTION STATEMENT OF THE WILSON-HILFERTY (1931)
C     CHI-SQUARE APPROXIMATION
C
      CHSQF(E,F)=E*(1.-2./(9.*E)+F*DSQRT(2./(9.*E)))**3
C
C     RNORF AND RRORF ARE FUNCTION STATEMENTS OF THE BOX-MULLER (1958)
C     STANDARD NORMAL TRANSFORMATIONS
C
      RNORF(E,F)=(DSQRT(-2.*DLOG(E)))*DCOS(2.*PI*F)
      RRORF(F,F)=(DSQRT(-2.*DLOG(E)))*DSIN(2.*PI*F)
      PI=3.14152927
C
C     LOAD A(I,J),I<J,WITH INDEPENDENT N(0,1) VARIATES
C
      DO 1 I=1,NP
      DO 1 J=1,NP
    1 A(I,J)=0.
      IF(NP*(NP-1)/2) 14,5,2
      K=0
      DO 4 I=2,NP
      II=I-1
      DO 4 J=1,II
      CALL RANDC(KY,U1)
      CALL RANDC(KY,U2)
C
C     RANDC GENERATES A UNIFORM VARIATE (DUE TO DR.C.E.GATES)
C
```

```
      K=K+1
      K1=K-2*(K/2)
      IF(K1.EQ.0) GO TO 3
      A(I,J)=RNORF(U1,U2)
      GO TO 4
    3 A(I,J)=RRORF(U1,U2)
    4 CONTINUE
    5 I=0
    6 I=I+1
C
C     LOAD A(I,I) WITH THE SQUARE ROOT OF INDEPENDENT
C     CHI-SQUARE (N-I+1 D.F.) VARIATES
C
      DF=N-I+1
      CALL RANDC(KY,U1)
      CALL RANDC(KY,U2)
      Y1=RNORF(U1,U2)
      Y2=RRORF(U1,U2)
      A(I,I)=DSQRT(CHSQF(DF,Y1))
      IF(I-NP) 7,8,14
    7 I=I+1
      DF=N-I+1
      A(I,I)=DSQRT(CHSQF(DF,Y2))
      IF(I-NP) 6,8,14
    8 CONTINUE
C
    9 DO 10 M=1,NP
      DO 10 K=1,M
      B(M,K)=0.
      DO 10 J=1,K
   10 B(M,K)=B(M,K)+A(M,J)*A(K,J)
      DO 11 M=2,NP
      M1=M-1
      DO 11 K=1,M1
   11 B(K,M)=B(M,K)
      RN=N
      DO 12 L=1,NP
      DO 12 M=1,NP
      C(L,M)=0.
      DO 12 K=1,L
   12 C(L,M)=C(L,M)+D(L,K)*B(K,M)
      DO 13 I=1,NP
      DO 13 J=1,NP
      SA(I,J)=0.
      DO 13 K=1,J
   13 SA(I,J)=SA(I,J)+C(I,K)*D(J,K)/RN
C
C     SA IS ARRAY WHOSE ELEMENTS HAVE A WISHART (N,SIGMA) DISTRIBUTION
C
   14 RETURN
      END
```