5-1-2002

JMASM3: A Method for Simulating Systems of Correlated Binary Data

Todd C. Headrick

Southern Illinois University at Carbondale, headrick@siu.edu

Follow this and additional works at: http://digitalcommons.wayne.edu/jmasm

Part of the Applied Statistics Commons, Social and Behavioral Sciences Commons, and the Statistical Theory Commons

JMASM3: A Method for Simulating Systems of Correlated Binary Data

Todd C. Headrick
Southern Illinois University at Carbondale

An efficient algorithm is derived for generating systems of correlated binary data. The procedure allows for the specification of all pairwise correlations within each system. Intercorrelations between systems can be specified qualitatively. The procedure requires the simultaneous solution of a system of equations for obtaining the threshold probabilities to generate each system of binary data. A numerical example is provided to demonstrate that the procedure generates correlated binary variables that yield correlations in close agreement with the specified population correlations.

Keywords: Correlated binary variables, Simulation, Uniform random deviates

Introduction

The availability of the desktop computer has made simulation and Monte Carlo techniques widely applicable in statistical research. For example, Monte Carlo methods may be used to compare the small sample properties of a test statistic with its competitors or whether these properties are consistent with the statistic's asymptotic approximation (Headrick & Rotou, 2001). Markov chain Monte Carlo methods (e.g., the Gibbs or slice sampler, Robert & Casella, 1999) are also commonly used to generate posterior distributions to carry out Bayesian analyses. Further, these techniques are now applicable to many areas of research interest. Some examples include: bootstrap tilting (Hesterberg, 2001); conditional logistic regression (Mehta, Patel, & Senchaudhuri, 2000); and likelihood inference with missing data (Gilks, Richardson, & Spiegelhalter, 1998).

There may be occasions when it is desirable to investigate the properties of statistics that involve correlated binary data. Most procedures that generate correlated binary variables are based on an underlying joint distribution that is multivariate normal (e.g., Bahadur, 1961; Emirch & Piedmonte, 1991; Leisch, 2001). This approach presents a problem because correlated binary variables with normal covariance structures may not provide a realistic simulation (Oman & Zucker, 2001). Further, the popular Bahadur (1961) method requires the specification of higher order correlations. These higher order correlations are nuisance parameters that make the Bahadur (1961) procedure arduous (Parzen, Bellamy, Lipsitz, & Fitzmaruice, 2001).

Thus, the purpose of this study is to derive a simple procedure that generates correlated binary variables. The proposed method simulates systems of multivariate binary data using *a priori* specified marginal probabilities and only requires a uniform random number generator.

Todd C. Headrick is Assistant Professor of statistics, Mail Code 4618, 222-J Wham Building, Southern Illinois University–Carbondale, IL, 62901. His areas of research interest are statistical computing, nonparametric statistics, and optimization. Email him at headrick@siu.edu.

Notation and a theoretical derivation of the proposed method are first provided. A numerical example is subsequently worked to demonstrate the procedure. *Mathematica* (Wolfram, Version 4.0, 1999) notebooks are available from the author that simultaneously solves systems of equations that provide the probability threshold points for implementing the procedure. Fortran 77 source code is also provided in the Appendix to demonstrate the algorithm.

Methodology

Let π_i be a pre-specified constant such that $\pi_i \in (0,1)$ and U_i be a $K \times 1$ vector of uniform random deviates on the interval $(0, 1)$ where $i = 1,...,T$. Let Y_1 represent a $K \times 1$ vector of a sequence of independent Bernoulli trials defined by

$$Y_1 = \begin{cases} 1 & U_1 < \pi_1 \\ 0 & U_1 > \pi_1 \end{cases} \qquad (1)$$

It follows that

$$E[Y_1] = \mu_{Y_1} = \pi_1, \text{ and} \qquad (2)$$

$$Var[Y_1] = \mu_{Y_1}(1 - \mu_{Y_1}) = \pi_1(1 - \pi_1). \qquad (3)$$

For subscripts $i = 2,...,T$, let Y_i represent the i-th $K \times 1$ vector of binary variables resulting from a sequence of independent trials defined by

$$Y_i = \begin{cases} Y_1 & U_i < \pi_i \\ Y_1 + 1 & U_i > \pi_i \text{ and } Y_1 = 0 \\ 1 - Y_1 & U_i > \pi_i \text{ and } Y_1 = 1 \end{cases} \qquad (4)$$

Thus, we have

$$E[Y_i] = \mu_{Y_i} = \pi_1\pi_i + (1 - \pi_1)(1 - \pi_i), \qquad (5)$$

$$E[Y_1 Y_i] = \mu_{Y_1}\pi_i = \pi_1\pi_i, \text{ and} \qquad (6)$$

$$Var[Y_i] = \mu_{Y_i}(1 - \mu_{Y_i}), \quad \forall_{i>1}. \qquad (7)$$

Given the definitions of Y_1 and Y_i in (1) and (4), let their measure of correlation, denoted as $\rho_{Y_1 Y_i}$, be defined as

$$\rho_{Y_iY_i} = \frac{Cov\ [Y_i,Y_i]}{\sqrt{Var[Y_i]}\sqrt{Var[Y_i]}}$$

$$= \frac{E[Y_iY_i] - E[Y_i]E[Y_i]}{\{(E[Y_i^2] - (E[Y_i])^2) \times (E[Y_i^2] - (E[Y_i])^2)\}^{1/2}}$$ (8)

$$= \frac{E[Y_iY_i] - E[Y_i]E[Y_i]}{\{(E[Y_i](1 - E[Y_i])) \times (E[Y_i](1 - E[Y_i]))\}^{1/2}},$$

where $E[Y_i^2] = E[Y_i]$ and $E[Y_i^2] = E[Y_i]$ because Y_i and Y_i are binary realizations of either 1 or 0. Substituting equations (2), (3), (5), (6), and (7) into (8) yields:

$$\rho_{Y_iY_i} = \frac{\pi_i\pi_i - \mu_{Y_i}\mu_{Y_i}}{\sqrt{\mu_{Y_i}(1 - \mu_{Y_i})}\sqrt{\mu_{Y_i}(1 - \mu_{Y_i})}}$$ (9)

Given T vectors of Y_i, define X_{ij} as the j-th $K \times 1$ vector of binary variables generated from Y_i as follows:

$$X_{ij} = \begin{cases} Y_i & U_{ij} < \pi_{ij} \\ Y_i + 1 & U_{ij} > \pi_{ij} \text{ and } Y_i = 0 \\ 1 - Y_i & U_{ij} > \pi_{ij} \text{ and } Y_i = 1 \end{cases}$$

with $j = 1, \ldots, k, \ldots, N$ (10)

where the second index (j) runs faster than the first. As such, there are NT vectors of X_{ij}. Note that the NT vectors of X_{ij} are independent of the T vectors of U_i and the constants of π_{ij} are not necessarily equal π_i.

Taking expectations with respect to X_{ij} and the product term of $X_{ij}X_{ik}$ gives

$$E[X_{ij}] = \mu_{X_{ij}} = \pi_i\pi_{ij} + (1 - \pi_i)(1 - \pi_{ij}),$$ (11)

$$E[X_{ik}] = \mu_{X_{ik}} = \pi_i\pi_{ik} + (1 - \pi_i)(1 - \pi_{ik}),$$ (12)

$$E[X_{ij}X_{ik}] = \pi_i\pi_{ij}\pi_{ik} + (1 - \pi_i)(1 - \pi_{ij})(1 - \pi_{ik}),$$ (13)

and hence

$$Var[X_{ij}] = \mu_{X_{ij}}(1 - \mu_{X_{ij}}),$$ (14)

$$Var[X_{ik}] = \mu_{X_{ik}}(1 - \mu_{X_{ik}}).$$ (15)

Taking expectations with respect to X_{ij} and the product term $X_{ij}X_{ik}$ for $i = 2, \ldots, T$ yields

$$E[X_{ij}] = \mu_{X_{ij}} = \pi_{ij}(\pi_1\pi_i + (1 - \pi_1)(1 - \pi_i)) + (1 - \pi_{ij})(1 - (\pi_1\pi_i + (1 - \pi_1)(1 - \pi_i)))$$ (16)

$$E[X_{ik}] = \mu_{X_{ik}} = \pi_{ik}(\pi_1\pi_i + (1 - \pi_1)(1 - \pi_i)) + (1 - \pi_{ik})(1 - (\pi_1\pi_i + (1 - \pi_1)(1 - \pi_i)))$$ (17)

$$E[X_{ij}X_{ik}] = \pi_{ij}\pi_{ik}(\pi_1\pi_i + (1 - \pi_1)(1 - \pi_i)) + (1 - \pi_{ij})(1 - \pi_{ik})(1 - (\pi_1\pi_i + (1 - \pi_1)(1 - \pi_i)))$$ (18)

$$Var[X_{ij}] = \mu_{X_{ij}}(1 - \mu_{X_{ij}}), \text{ and}$$ (19)

$$Var[X_{ik}] = \mu_{X_{ik}}(1 - \mu_{X_{ik}}).$$ (20)

The correlations $\rho_{X_{ij}X_{ik}}$ and $\rho_{X_{ij}X_{ik}}$ can be determined by substituting the expressions in (11) through (15) and (16) through (20) into analogous definitions of (8) expressed in terms of the X_{ij}.

Given specified constants of π_i, an efficient number of correlated X_{ij} vectors for each integer of $i = 1$, the

selection of $N = 3$ yields a system of three equations (for three pairwise correlations $\rho_{X_{11}X_{12}}$, $\rho_{X_{11}X_{13}}$, $\rho_{X_{12}X_{13}}$) solvable in terms of three unknowns (probability threshold points of π_{11}, π_{12}, π_{13}). This system is constructed by first substituting the right-hand sides of (11) and (12) into (14) and (15). Subsequently substituting (11) through (15) into three equations of the form in (8) yields

$$\rho_{X_{11}X_{12}} =$$

$$\frac{[\pi_1\pi_{11}\pi_{12} + (1 - \pi_1)(1 - \pi_{11})(1 - \pi_{12})] - [\pi_1\pi_{11} + (1 - \pi_1)(1 - \pi_{11})][\pi_1\pi_{12} + (1 - \pi_1)(1 - \pi_{12})]}{\sqrt{[\pi_1\pi_{11} + (1 - \pi_1)(1 - \pi_{11})](1 - [\pi_1\pi_{11} + (1 - \pi_1)(1 - \pi_{11})])}\sqrt{[\pi_1\pi_{12} + (1 - \pi_1)(1 - \pi_{12})](1 - [\pi_1\pi_{12} + (1 - \pi_1)(1 - \pi_{12})])}}.$$ (21)

$$\rho_{X_{11}X_{13}} =$$

$$\frac{[\pi_1\pi_{11}\pi_{13} + (1 - \pi_1)(1 - \pi_{11})(1 - \pi_{13})] - [\pi_1\pi_{11} + (1 - \pi_1)(1 - \pi_{11})][\pi_1\pi_{13} + (1 - \pi_1)(1 - \pi_{13})]}{\sqrt{[\pi_1\pi_{11} + (1 - \pi_1)(1 - \pi_{11})](1 - [\pi_1\pi_{11} + (1 - \pi_1)(1 - \pi_{11})])}\sqrt{[\pi_1\pi_{13} + (1 - \pi_1)(1 - \pi_{13})](1 - [\pi_1\pi_{13} + (1 - \pi_1)(1 - \pi_{13})])}}.$$ (22)

$$\rho_{X_{12}X_{13}} =$$

$$\frac{[\pi_1\pi_{12}\pi_{13} + (1 - \pi_1)(1 - \pi_{12})(1 - \pi_{13})] - [\pi_1\pi_{12} + (1 - \pi_1)(1 - \pi_{12})][\pi_1\pi_{13} + (1 - \pi_1)(1 - \pi_{13})]}{\sqrt{[\pi_1\pi_{12} + (1 - \pi_1)(1 - \pi_{12})](1 - [\pi_1\pi_{12} + (1 - \pi_1)(1 - \pi_{12})])}\sqrt{[\pi_1\pi_{13} + (1 - \pi_1)(1 - \pi_{13})](1 - [\pi_1\pi_{13} + (1 - \pi_1)(1 - \pi_{13})])}}.$$ (23)

Specifying π_1 and selecting positive values for $\rho_{X_{11}X_{12}}$, $\rho_{X_{11}X_{13}}$, and $\rho_{X_{12}X_{13}}$ in (21), (22), and (23) and then simultaneously solving these equations gives the probability threshold points of π_{11}, π_{12}, π_{13} to generate X_{11}, X_{12}, and X_{13} with the desired intercorrelations for (10). Systems of equations analogous to (21), (22), and (23) can be created in the same manner for all $i = 2,...,T$ from equations (16) through (20). As such, $3T$ vectors of X_{ij} can be generated with three positive pairwise correlations within each system.

Without loss of generality, feasible solutions (i.e., $0 \le \pi_{ij} \le 1$) are obtainable provided that the following conditions hold with respect to the left-hand sides of (21), (22), and (23)

$$\frac{(\rho_{X_{11}X_{12}})(\rho_{X_{11}X_{13}})}{(\rho_{X_{12}X_{13}})} \le 1,$$ (24a)

$$\frac{(\rho_{X_{11}X_{12}})(\rho_{X_{12}X_{13}})}{(\rho_{X_{11}X_{13}})} \le 1, \text{ and}$$ (24b)

$$\frac{(\rho_{X_{11}X_{13}})(\rho_{X_{12}X_{13}})}{(\rho_{X_{11}X_{12}})} \le 1.$$ (24c)

These conditions must also hold with respect to the other $T-1$ systems of equations.

Because all Y_i for $i = 2,...,T$ are a function of Y_1 from (4), intercorrelations between X_{ij} belonging to *different* systems may also exist and can be analytically determined as follows:

$$\rho_{X_{ij}X_{jj}} = (\rho_{X_{ij}Y_i})(\rho_{Y_iY_j})(\rho_{Y_jX_{jj}}).$$ (25)

Further, and by inspection of (5) and (9), correlations in equation (25) may be negative, zero, or positive. Specifically, correlations are respectively negative, zero, or positive when $\pi_i < 0.5$, $\pi_i = 0.5$, or $\pi_i > 0.5$ in equation (9) for all $i = 2,...,T$. Note that the correlations $\rho_{X_{ij}Y_i}$ and $\rho_{Y_jX_{jj}}$ in (2.5) are derived analogously as (21),..., (23) from (8).

Numerical Example

Suppose it is desired to generate two systems of correlated binary data with correlations $\rho_{X_{11}X_{12}} = 0.40$, $\rho_{X_{11}X_{13}} = 0.50$, $\rho_{X_{12}X_{13}} = 0.60$; and $\rho_{X_{21}X_{22}} = 0.75$, $\rho_{X_{21}X_{23}} = 0.80$, $\rho_{X_{22}X_{23}} = 0.90$ and where the variables between the systems are inversely related. The following steps are taken:

1. Ensure that three pairwise intercorrelations within the systems satisfy (24a), (24b), and (24c).

2. Specify values for π_1 and π_2. Let $\pi_1 = 0.90$ and let $\pi_2 = 0.10$ to obtain the inverse relationships between the two systems.

3. Simultaneously solve the two systems of the form in

(21), (22) and (23) given the specified correlations and values of π_1 and π_2. The numerical equation solver FindRoot (*Mathematica*, 1999) yields solutions of $\pi_{11} = 0.881247$, $\pi_{12} = 0.924095$, $\pi_{13} = 0.972456$, $\pi_{21} = 0.939341$, $\pi_{22} = 0.974722$, and $\pi_{23} = 0.993961$.

4. Use equation (25) to determine the correlations between the variables of the two systems. The correlations are: $\rho_{X_{11}X_{21}} = -0.294484$, $\rho_{X_{11}X_{22}} = -0.331295$, $\rho_{X_{11}X_{23}} = -0.353381$, $\rho_{X_{12}X_{21}} = -0.353381$, $\rho_{X_{12}X_{22}} = -0.397554$, $\rho_{X_{12}X_{23}} = -0.424057$, $\rho_{X_{13}X_{21}} = -0.441726$, $\rho_{X_{13}X_{22}} = -0.496942$, and $\rho_{X_{13}X_{23}} = -0.530071$.

5. Given the parameters and solutions from steps 2 and 3, generate the correlated binary data. Fortran 77 source code is presented in the Appendix to implement this example. Presented in Table 1 are population correlations and computed correlations for both within and between the two systems described in steps 1 and 4. Single samples of binary data with K = 5000 were generated for each of the six variables as described in step 5. Inspection of Table 1 indicates that the proposed method generated correlated binary data that yielded computed correlations that were in close agreement with the population correlations.

Conclusion

As previously mentioned the systems of equations of the form in (21), (22), and (23) solve for the probability threshold points (π_{ij}) for only positive correlations. However, it may be desirable to generate binary data with negative correlations. This can be accomplished by writing additional *if* – *else* structure statements in the Fortran 77 source code that simultaneously reversing the 1's and 0's within each selected vector. Such reversals will allow for the construction of negative correlations within the system(s).

More specifically, suppose that it was desired that $\rho_{X_{11}X_{12}} = -0.40$, and $\rho_{X_{11}X_{13}} = -0.50$ in the numerical example. This can be accomplished by simultaneously reversing the 1's and 0's in the vectors X_{12}, and X_{13}. It should be noted that these changes would also create positive correlations between the systems of: $\rho_{X_{12}X_{21}} = 0.353381$, $\rho_{X_{12}X_{22}} = 0.397554$, $\rho_{X_{12}X_{23}} = 0.424057$, $\rho_{X_{13}X_{21}} = 0.441726$, $\rho_{X_{13}X_{22}} = 0.496942$, and $\rho_{X_{13}X_{23}} = 0.530071$.

It should also be pointed out that as the sample sizes become smaller the probability increases of obtaining vectors of either all 1's or all 0's (i.e., standard deviations of zero). Thus, one way to circumvent this problem in a larger simulation would be to initialize the correlation computed in the subroutine (e.g., PCOR in the Appendix) to zero when either standard deviation is zero and

Table 1. Two systems of correlated binary data generated using the Fortran source code provided in the Appendix. A single sample of size $K = 5000$ was drawn for each of the six variables.

System 1		System 2	
Variables: X_{11}, X_{12}, X_{13}		Variables: X_{21}, X_{22}, X_{23}	
Pop. Correlations Within System 1	Computed Correlations	Pop. Correlations Within System 2	Computed Correlations
$\rho_{X_{11}X_{12}} = 0.40$	0.400780	$\rho_{X_{21}X_{22}} = 0.75$	0.745030
$\rho_{X_{11}X_{13}} = 0.50$	0.503307	$\rho_{X_{21}X_{23}} = 0.80$	0.790314
$\rho_{X_{12}X_{13}} = 0.60$	0.591606	$\rho_{X_{22}X_{23}} = 0.90$	0.893551

Population Correlations Between System 1 and System 2	Computed Correlations
$\rho_{X_{11}X_{21}} = -0.294484$	-0.290305
$\rho_{X_{11}X_{22}} = -0.331295$	-0.330829
$\rho_{X_{11}X_{23}} = -0.353381$	-0.351456
$\rho_{X_{12}X_{21}} = -0.353381$	-0.347825
$\rho_{X_{12}X_{22}} = -0.397554$	-0.390941
$\rho_{X_{12}X_{23}} = -0.424057$	-0.407341
$\rho_{X_{13}X_{21}} = -0.441726$	-0.449708
$\rho_{X_{13}X_{22}} = -0.496942$	-0.504316
$\rho_{X_{13}X_{23}} = -0.530071$	-0.526768

keep count of the zero correlations throughout the simulation. Thus, the total real number of replications could be computed by subtracting the number of times a zero correlation occurs due to standard deviations of zero (e.g., 50 occurrences) from the initialized total number of replications (e.g. 10,000). That is, the total real number of replications would be 9950.

References

Bahadur, R. R. (1961). A representation of the joint distribution of responses to n dichotomous items. In H. Solomon (Ed.), *Studies in item analysis and prediction*. Stanford, CA: Stanford University Press, 158-168.

Blair, R. C. (1987). *RANGEN*, Boca Raton, FL: IBM.Emrich, L. J., & Piedmonte, M. R. (1991). A method for generating high-dimensional multivariate binary variates. *American Statistician*, 45, 302-304.

Gilks, W. R., Richardson, S., & Spiegelhalter, D. J. (1998). *Markov chain Monte Carlo in practice*. Boca Raton, FL: Chapman & Hall/CRC.

Headrick, T. C. & Rotou, O. (2001). An investigation of the rank transformation in multiple regression. *Computational Statistics & Data Analysis*, 38, 203-215.

Hesterberg, T. (2001). *Bootstrap tilting*. The University of Florida Statistics Symposium on Monte Carlo in the New Millenium. Gainesville, FL.

Mehta, C. R., Patel, N. R., & Senchaudhuri, P. (2000). Efficient Monte Carlo methods for conditional logistic regression. *Journal of the American Statistical Association*, 95, 99-108.

Leisch, F. (2001) *The bindata package: Version 2*. Vienna: Author.

Oman, S. D., & Zucker, D. M. (2001). Modeling and generating correlated binary variables. *Biometrika*, 88, 287-290.

Parzen, M., Bellamy, S., Lipsitz, S. R., & Fitzmaurice, G. (2001). *A random effects model for clustered data*. (Unpublished), Graduate School of Business, University of Chicago.

Robert, C. P., & Casella, G. (1999). *Monte Carlo statistical methods*. New York: Springer.

Wolfram, S. (1999). *The Mathematica book* (4th). Wolfram Media/Cambridge University Press.

APPENDIX

```
C********************************************************************
C THIS EXAMPLE PROGRAM (BINARY) GENERATES TWO SYSTEMS OF BINARY
C DATA. EACH SYSTEM CONTAINS THREE VECTORS WITH REQUESTED INTER -
C CORRELATIONS.
C
C       SUBROUTINES CALLED BY PROGRAM BINARY:
C       1. UNI1 (BLAIR, 1987) BASIC UNIFORM (0,1) PSEUDO-RANDOM NUMBER
C                   GENERATOR. PURPOSE IS TO CREATE VECTORS OF U1, U2,
C                   U11, U12, U13, U21, U22, U23 OF LENGTH K.
C       2. PCOR (BLAIR, 1987) COMPUTES THE PEARSON-PRODUCT MOMENT
C                   CORRELATION COEFFICIENT.
C       ARGUMENTS:
C       1. INPUTS:
C               K: VECTOR LENGTH OF BINARY DEVIATES (1 OR 0) TO BE
C                   RETURNED, K=5000.
C               P1, P2: PARAMETERS ON THE INTERVAL (0, 1),
C                   P1=0.90, P2=0.10.
C               P11, P12, P13: THRESHOLD PROBABILITIES FOR SYSTEM 1,
C                   P11=0.939341, P12=0.974722, P13=0.993961.
C               P21, P22, P23: THRESHOLD PROBABILITIES FOR SYSTEM 2,
C                   P21=0.881247, P22=0.924095, P23=0.972456.
C               DSEED(1-8) – INPUT/ OUTPUT: DOUBLE PRECISION
C                   VARIABLE IN THE EXCLUSIVE RANGE (1.D0,2147483647.D0).
C       2. OUTPUTS:
C               Y1, Y2: REAL VECTORS OF LENGTH K CONTAINING THE
C                   CORRELATED BINARY VARIATES TO THE GENERATE THE
C                   SIX VECTORS OF XIJ.
C
C               X11, X12, X13: REAL VECTORS OF LENGTH K CONTAINING THE
C                   REQUESTED CORRELATED BINARY VARIATES FOR
C                   SYSTEM 1. REQUESTED CORRELATIONS ARE:
C                   CORR(X11,X12)=0.40, CORR(X11,X13)=0.50,CORR(X12,X13)=0.60.
C
C               X21, X22, X23: REAL VECTORS OF LENGTH K CONTAINING THE
C                   REQUESTED CORRELATED BINARY VARIATES FOR
C                   SYSTEM 2. REQUESTED CORRELATIONS ARE:
C                   CORR(X21,X22)=0.75, CORR(X21,X23)=0.80,CORR(X22,X23)=0.90.
C
C               R112, R113, R123: COMPUTED CORRELATIONS BETWEEN VECTORS
C                   X11, X12, X13 OF SYSTEM 1.
C
C               R212, R213, R223: COMPUTED CORRELATIONS BETWEEN VECTORS
C                   X21, X22, X23 OF SYSTEM 2.
C
C               R1211, R1212, R1213, R1221, R1222, R1223, R1231, R1232, R1233:
C                   COMPUTED CORRELATIONS BETWEEN VECTORS
C                   X11, X12, X23, X21, X22, X23 OF SYSTEMS 1 AND 2.
C                   (E.G.: R1211 IS THE CORRELATION BETWEEN X11 AND X21)
C********************************************************************
```

```
      PROGRAM BINARY
      REAL Y1(5000), Y2(5000), X11(5000), X12(5000), X13(5000), X21(5000), X22(5000),
      X23(5000), U1(5000), U2(5000), U11(5000), U12(5000), U13(5000), U21(5000),
      U22(5000), U23(5000), P1, P2, P11, P12, P13, P21, P22, P23, R112, R113, R123, R212,
      R213, R223, R1211, R1212, R1213, R1221, R1222, R1223, R1231, R1232, R1233
      INTEGER I, K
      DOUBLE PRECISION DSEED1, DSEED2, DSEED3, DSEED4, DSEED5, DSEED6,
      DSEED7, DSEED8
      READ (*,*) K, P1, P2, P11, P12, P13, P21, P22, P23
C*******************************************************************************
C INITIALIZE DSEED(1-8) AND CALL UNI1 TO GENERATE ARRAYS OF U1, U2, U11,
C U12, U13, U21, U22, U23
C*******************************************************************************
      CALL UNI1 (DSEED1, K, U1)
C*******************************************************************************
C CREATE CALL UNI1 STATEMENTS FOR U2,...,U23
C*******************************************************************************
      DO 10 I = 1, K
      IF ( U1(I) .LE. P1 ) THEN
      Y1(I) = 1.0
      ELSE IF ( U1(I) .GE. P1) THEN
      Y1(I) = 0.0
      ENDIF
      IF ( U2(I) .LE. P2 ) THEN
      Y2(I) = Y1(I)
      ELSE IF ( U2(I) .GE. P2 .AND. Y1(I) .EQ. 0) THEN
      Y2(I)=Y1(I) + 1
      ELSE IF ( U2(I) .GE. P2 .AND. Y1(I) .EQ. 1) THEN
      Y2(I)=Y2(I) – 1
      ENDIF
      IF( U11(I) .LE. P11 ) THEN
      X11(I)=Y1(I)
      ELSE IF ( U11(I) .GE. P11 .AND. Y1(I) .EQ. 0) THEN
      X11(I)=Y1(I) + 1
      ELSE IF ( U11(I) .GE. P11 .AND. Y1(I) .EQ. 1) THEN
      X11(I)=Y1(I) – 1
      ENDIF
C*******************************************************************************
C CREATE IF–THEN STRUCTURES TO GENERATE VECTORS X12 AND X13 FROM Y1.
C*******************************************************************************
      IF( U21(I) .LE. P21) THEN
      X21(I)=Y2(I)
      ELSE IF ( U21(I) .GE. P21 .AND. Y2(I) .EQ. 0) THEN
      X21(I)=Y2(I) + 1
      ELSE IF ( U21(I) .GE. P21 .AND. Y2(I) .EQ. 1) THEN
      X21(I)=Y2(I) – 1
      ENDIF
C*******************************************************************************
C CREATE IF–THEN STRUCTURES TO GENERATE VECTORS X22 AND X23 FROM Y2.
C*******************************************************************************
   10 CONTINUE
```

```
C******************************************************************************
C CALL PCOR TO COMPUTE THE CORRELATIONS
C******************************************************************************
          CALL PCOR (X11, X12, K, R112)
          CALL PCOR (X11, X13, K, R113)
          CALL PCOR (X12, X13, K, R123)
          CALL PCOR (X21, X22, K, R212)
          CALL PCOR (X21, X23, K, R213)
          CALL PCOR (X22, X23, K, R223)
          CALL PCOR (X11, X21, K, R1211)
          CALL PCOR (X11, X22, K, R1212)
          CALL PCOR (X11, X23, K, R1213)
          CALL PCOR (X12, X21, K, R1221)
          CALL PCOR (X12, X22, K, R1222)
          CALL PCOR (X12, X23, K, R1223)
          CALL PCOR (X13, X21, K, R1231)
          CALL PCOR (X13, X22, K, R1232)
          CALL PCOR (X13, X23, K, R1233)
          STOP
          END
```