sium on Modeling and Simulation Methodology, B. Zeigler, Ed. Amsterdam, North-Holland, 1979.
[2] R. W. Blanning, "The sources and uses of sensitivity information," *Interfaces*, vol. 4, no. 4, pp. 32–38, Aug. 1974 (see also vol. 5, no. 3, pp. 24–25, May 1975).
[3] J. J. M. Braat, "The I.P.S.O. control system," *Int. J. Production Research*, vol. 11, no. 4, pp. 417–436, 1973.
[4] J. E. Ertel and E. B. Fowlkes, "Methods for fitting linear spline and piecewise multiple linear regression," in *Proc. Computer Science and Statistics: 8th Ann. Sym. on the Interface*, J. W. Frane, Ed., Health Sciences Computing Facility, Univ. of California, Los Angeles, 1975.
[5] A. Geoffrion, "The purpose of mathematical programming is insight, not numbers," *Interfaces*, vol. 7, no. 1, pp. 81–92, Nov. 1976.
[6] G. B. Goldstein and T. Dushane, *Repro-modeling Applied to the Simplification of Taradcom Computer Models*, Rep. 12243, U.S. Army Tank-Automotive R&D Command, Dec. 1976.
[7] A. Horowitz and W. S. Meisel, *The Application of Repro-modeling to the Analysis of a Photochemical Air Pollution Model*, Rep. EPA-650/4-74-001, Environmental Protection Agency, Dec. 1973.
[8] E. J. Ignall, P. Kolesar, and W. E. Walker, "Using simulation to develop and validate analytic models: some case studies," *Operations Research*, vol. 26, no. 2, pp. 237–253, Mar.–Apr. 1978.
[9] R. L. Keeney and H. Raiffa, *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. New York: Wiley, 1976.
[10] J. P. C. Kleijnen, *Statistical Techniques in Simulation*. New York: Marcel Dekker, 1974/1975.
[11] ——, *Generalizing Simulation Results Through Metamodels*, Working paper, Dep. of Business and Economics, Katholieke Hogeschool, Tilburg, The Netherlands, June 1978.
[12] J. P. C. Kleijnen and P. J. Rens, "IMPACT revisited: A critical analysis of IBM's inventory package 'IMPACT'," *Production and Inventory Management*, vol. 19, no. 1, pp. 71–90, 1978.
[13] J. P. C. Kleijnen, A. J. Van Den Burg, and R. T. Van Der Ham, "Generalization of simulation results: practicality of statistical methods," *European J. Operational Research* (in press).
[14] G. F. Koons and B. Perlic, *A Study of Rolling-mill Productivity Utilizing a Statistically Designed Simulation Experiment*, Research Laboratory, United States Steel Corp., Monroeville, PA, 1977.
[15] R. W. Lawless, L. H. Williams, and C. G. Richie, "A sensitivity analysis tool for simulation with application to disaster planning," *Simulation*, vol. 17, no. 6, pp. 217–223, Dec. 1971.
[16] W. S. Meisel and D. C. Collins, "Repro-modeling: An approach to efficient model utilization and interpretation," *IEEE Trans. Syst., Man, and Cybern.*, vol. SMC-3, no. 4, pp. 349–358, July 1973.
[17] D. C. Montgomery and V. M. Bettencourt, "Multiple response surface methods in computer simulation," *Simulation*, vol. 29, no. 4, pp. 113–121, Oct. 1977.
[18] C. C. Pegels, *Systems Analysis for Production Operations*. New York: Gordon and Breach Science, 1976.
[19] M. R. Rose and R. Harmsen, "Using sensitivity analysis to simplify ecosystem models: A case study," *Simulation*, vol. 31, no. 1, pp. 15–26, July 1978.
[20] W. A. Sherden, "Origin of simultaneity in corporate models," in *Winter Simulation Conf.*, H. J. Highland, T. J. Schriber, and R. G. Sargent, Eds. New York: ACM, 1976.
[21] W. A. H. Thissen, *Investigations into the Club of Rome's World 3 Model: Lessons for Understanding Complicated Models*, Doctoral dissertation, Technical Univ., Eindhoven, The Netherlands, 1978.
[22] J. K. Weeks and J. S. Fryer, "A methodology for assigning minimum cost due-dates," *Management Sci.*, vol. 23, no. 8, pp. 872–881, Apr. 1977.
[23] H. C. Yen and W. Pierskalla, *A Simulation of a Centralized Blood Banking System*, Michael Reese Hospital, Office of Operations Research, Chicago, IL, 1977.

A Simple Technique for the Generation of Correlated Random Number Sequences

S. P. NAWATHE AND B. V. RAO

Abstract—**A simple technique, based on the theory of optimal linear prediction, is presented for the generation of cross-correlated random number sequences. The method is essentially a computer simulation of the process of linear prediction with the error of prediction playing an important role. The algorithm rearranges a random number sequence with reference to another so as to obtain the desired correlation between the two. With no specifications on the distribution of the random number sequence (linear prediction does not assume knowledge of the probability law of the random variables) the method should therefore be applicable to all continuous distributions. The validity of the algorithm has been justified through the derivations of the expressions for the expectation and variance of the correlation coefficient. Satisfactory results have been obtained for three typical distributions, viz., normal, uniform, and exponential. The assumptions made regarding the behavior of the error of prediction have been verified through computations. The method is expected to find applications in various disciplines.**

Manuscript received December 12, 1975; revised May 5, 1976; November 1, 1976; May 17, 1978; and September 25, 1978.

The authors are with the Department of Electrical Engineering, Indian Institute of Technology, Bombay, India.

INTRODUCTION

Monte Carlo simulation provides a convenient means for the qualitative investigation of the behavior of a stochastic system. The scope of this method is determined by the extent to which the statistical characteristics of the random number generator resemble those of the system variables. There are many well-known techniques by which it is possible to generate random numbers having the required distribution characteristics [1], [2].

In some cases, the system variables display statistical interdependence. As an example, consider the problem of reliability analysis of a power system. In this case, failure of one of the power transmission links may cause overloading of another, thereby increasing its failure probability. While simulating such a system, it becomes necessary to generate sequences of random numbers having the prescribed mutual cross correlations among them.

It is well-known that a multivariate normal sample can be generated through appropriate linear combinations of independent normal random variables. These techniques are applicable only to the case of the normal distribution. Moreover, they involve problems like computation of the square root of the covariance matrix or the solution of a system of nonlinear algebraic equations [3].

Li and Hammond [4] suggest a procedure for generating correlated random variables with specified nonnormal probability distribution functions. In this procedure, a multivariate nonnormal sample is obtained through appropriate nonlinear transformation of a multivariate normal sample. The method involves operations like "predistortion" of the desired correlation matrix (in order to compensate for the distortion during the nonlinear transformation), evaluation of the square root of the correlation matrix (for the generation of multivariate normal sample), and transformation from normal to the desired distribution. Although standard procedures are available for all these operations, the overall procedure is quite tedious. If the inverse of the desired distribution is not known in closed form, the transformation is performed by numerical methods. In such a case, the method is prone to quantization errors at two stages. (Evaluation of the error function is equivalent to transformation from normal (Gaussian) to a uniform distribution). Moreover, this method fails if the required correlation matrix does not yield a positive semidefinite predistorted correlation matrix. We propose a simple alternative approach to the same problem.

The problem to be considered in this correspondence is that of the generation of two random number sequences $[X_N]$ and $[Y_N]$ of sample size N, to represent the two jointly distributed (not necessarily normal) random variables X and Y, respectively. The marginal distributions of X and Y are assumed to be known. The marginal distributions of $[X_N]$ and $[Y_N]$, when averaged over the entire sample size, are so adjusted as to match those of X and Y, respectively. (The elements of $[Y_N]$, in general, do not have the

0018-9472/79/0200-0096$00.75

same distribution as Y. However, when averaged over the entire sample size, the (average) distribution of $[Y_N]$ does match that of Y). The interdependence between X and Y is specified by the coefficient of linear correlation between them. The random number sequence $[Y_N]$ is generated by rearranging another random number sequence $[Z_N]$ with reference to $[X_N]$. The process of rearrangement is governed by the optimal linear predictor of Y and the error of prediction. Since the method is based on the theory of optimal linear prediction, knowledge of the joint probability law (which is very difficult to obtain in the case of nonnormal distribution) is not required. The method can easily be extended to the case of multidimensional random variables.

Optimal Prediction

If two random variables are jointly normally distributed, then the conditional expectation of one, for a given value of the other, can be calculated from the knowledge of their means, standard deviations, and the coefficient of correlation [5]. If these moments are not known *a priori*, they may be calculated from the observed data.

It may happen that the joint probability law of the two random variables X and Y is unknown. In the case of nonnormal distributions, even if the joint probability law is known, it is seldom possible to derive an expression for the conditional expectation. In such cases, the prediction problem can be attempted as follows. Let

$E^*(Y \mid X)$ the predictor of Y for given X,
q $= Y - E^*(Y \mid X)$ = the error of prediction,
$L(q)$ the loss function, which, for every true value of the random variables Y and its predictor, assigns a loss or cost.

The choice of the loss function is usually governed more by mathematically manipulative convenience than by any physical significance. The essential characteristics of the loss function are

i) $L(q) = 0$, if $q = 0$;
ii) monotonic, i.e., $L(q_2) \geq L(q_1)$, if $q_2 \geq q_1 \geq 0$;
iii) symmetric, i.e., $L(q) = L(-q)$.

It is desirable, though not essential, that the loss function be convex and continuous in the range of interest [6]. The risk function is defined as

$$R(q) = E[L(q)],$$

which is the expectation of the loss function over the admissible range of errors.

If the conditional distribution function of Y happens to be i) symmetric about the mean $\bar{Y}$ and ii) convex for $\bar{Y} \leq Y$, then it has been shown by Sherman [7] that the random variable that minimizes the risk function is the conditional expectation.

Deutsch [8] has shown that, if the loss function is chosen to be the squared error, then the above theory leads to an optimal predictor, even without the conditional distribution function being required to be either symmetric or convex. This is the principal reason why the loss function is chosen to be the squared error in almost all problems of prediction and estimation.

Best Linear Prediction

Let X and Y be two jointly distributed random variables. The best linear predictor of Y, for given X, is defined as that linear function $(aX + b)$ of the random variable X, that minimizes the mean squared error of prediction

$$E(q^2) = E\{[Y - (aX + b)]^2\}. \tag{1}$$

Expressions for the parameters a and b can be obtained by equating the respective partial derivatives of the above expression to zero. The best linear predictor of Y for $X = x$ is given by [5]

$$E^*(Y \mid X = x) = \bar{Y} - \rho \cdot \frac{\sigma_y}{\sigma_x}[\bar{X} - x] \tag{2}$$

where

$\bar{X}$, $\bar{Y}$ means of X and Y,
σ_x, σ_y standard deviations of X and Y,
ρ coefficient of linear correlation specifying the interdependence between X and Y.

It can be shown that the predictor of (2) is nothing but the projection of the random variable Y onto the random variable X. The existence and uniqueness of the best linear predictor are provided by the well-known projection theorem of abstract Hilbert space theory [9]. The derivation of the linear predictor does not require knowledge of the probability law of the random variables [10]. The only necessary condition for the above predictor to be realizable is that the random variables under consideration be square integrable (i.e., both σ_x^2 and σ_y^2 should be finite and nonnegative).

It is observed that the expression for the best linear predictor coincides with that for the conditional expectation of a jointly normally distributed random variable. Therefore, the best linear predictor could be interpreted alternatively as the conditional expectation, if the random variables under consideration are known to be jointly normally distributed. The error of prediction q is given by

$$q = Y - E^*(Y \mid X) \tag{3}$$

Since there is no reason to believe that the error is more likely to be positive than negative, or vice versa, expectation of the error can be assumed to be equal to zero. This assumption can be verified through actual computations during the simulation:

$$E(q) = 0. \tag{4}$$

Then variance of the error is given by [5]

$$\sigma_q = E(q^2) = \sigma_y^2(1 - \rho^2). \tag{5}$$

Generation of Correlated[1] Random Number Sequences

The problem to be considered is that of the generation of two random number sequences $[X_N]$ and $[Y_N]$ of sample size N, to represent the two jointly distributed random variables X and Y, so that the following conditions are satisfied:

i) The marginal distributions of $[X_N]$ and $[Y_N]$, when averaged over the entire sample size, are the same as those of X and Y, respectively (i.e., according to the specifications described by the practical situation).

ii) The estimate of the coefficient of correlation between $[X_N]$ and $[Y_N]$

$$\rho = \frac{1}{N}\sum_{j=1}^{N} \frac{(x_j - \bar{Y})(y_j - \bar{Y})}{\sigma_x \cdot \sigma_y} \to \rho_d, \qquad \text{as } N \to \infty,$$

where ρ_d is the desired correlation coefficient between X and Y.

The first condition can be satisfied with the help of the techniques described in [1] and [2]. In order to achieve the second, consider (3) rearranged as

$$Y = E^*(Y \mid X = x) + q \tag{6}$$

[1] The term correlated is used to denote pairwise cross correlation between the two random number sequences.

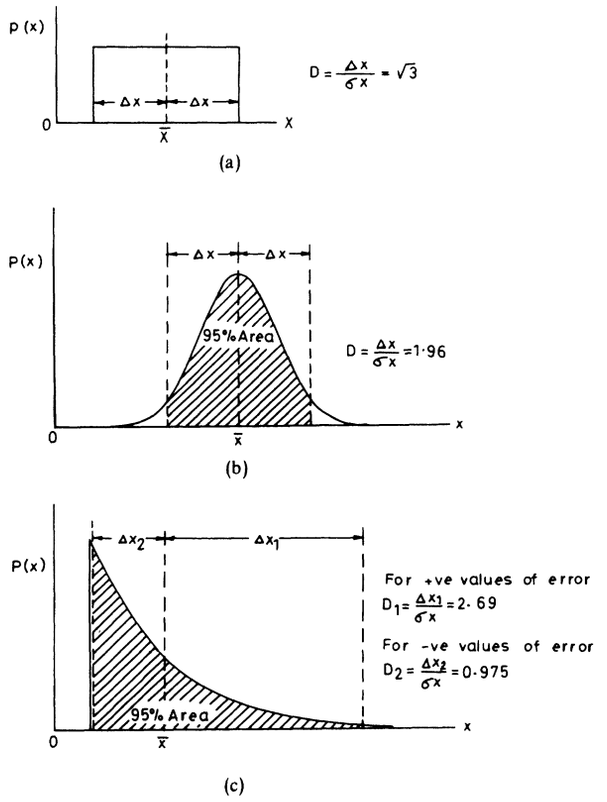

Fig. 1. Fixation of error bounds. (a) Uniform distribution. (b) Normal distribution. (c) Exponential distribution.

From (4)–(6), for every particular value x_j or x, it is possible to specify the range in which the corresponding y_j is likely to be situated with a high probability. Assuming the probability law of the random variable q to be the same as that of Y (this assumption can be verified through computations, see Fig. 5), it is possible to fix upper bounds on the magnitudes of positive and negative errors.

If the range of the random variable Y happens to be finite (as in the case of a uniform distribution) determination of the error bounds is straightforward (see Fig. 1(a)). From (5), it may be observed that the range of the error must be smaller than that of the random variable to be predicted. In the case of those probability laws in which the probability density curve asymptotically approaches the zero value, strictly speaking, the range of the random variable (and hence that of the error as well) extends to infinity. As a result, while fixing the error bounds, one is faced with the absurd task of comparing two infinities. This difficulty can be minimized by a simple approximation as described below.

Although the probability density curve may extend to infinity, in a practical sample, the random numbers are distributed only over a finite range. In the case of the normal distribution, it can be assumed (to a certain approximation) that the practical values of the random variable are restricted to the range bounded by $\pm 1.96\sigma$ (which includes 95 percent area under the probability density curve). Therefore, the error bounds are chosen so as to include 95 percent of the total area under the probability density curve (see Fig. 1(b)).

The same approximation can be used for determining the error bounds in the case of other distributions provided the assumed restricted range of the random variable is consistent with that of the sample. Although it is difficult to give a rigorous explanation for the choice of the cutoff level, the range that includes 95 percent of the area under the probability density curve appears to be a reasonable approximation for the determination of the error bound in the case of the probability density curves which asymptotically approach the zero value.

If the error bounds are tightened (i.e., the area included under the probability density curve is reduced), which is equivalent to assuming a stronger dependence between the two random variables, the magnitudes of the correlation coefficient will increase. However, if the error bounds are slackened (i.e., the area included under the probability density curve is increased), the result will be a decrease in the magnitudes of the correlation coefficient. If the probability law happens to be asymmetric about the mean value, bounds for the magnitudes of positive and negative errors are different (see Fig. 1(c)).

Now the method of obtaining correlated random number sequences can be described as follows. We start with a random number sequence $[X_N]$ of simple size N having the required distribution. For every particular x_j, the best linear predictor of Y can be calculated from (2). (Means and standard deviations of X and Y are specified by their marginal distributions which are assumed to be known. We substitute ρ by ρ_d, the desired correlation coefficient.) The procedure then consists of selecting y_j from an appropriate parent population (one which has the same distribution as that of the random variable Y) of random numbers, in a random manner, such that the error of prediction is within the bounds specified by the probability law of the random variable. In the case of populations of finite size, it may happen that for some x_j, no suitable match (i.e., a random number satisfying the condition of maximum allowable error) is available. Under such circumstances, a random number from the same population, which gives the smallest squared error, can be accepted without seriously affecting the simulation results. It is observed that even if the parent population from which the sequence $[Y_N]$ is generated, is of the same sample size N, on an average, 95 percent of the x_j find suitable matches.

The Algorithm

i) Generate two random number sequences $[X_N]$ and $[Z_N]$ having the required distribution characteristics with the help of the techniques described in [1] and [2]. $[Z_N]$ is to be used as a parent population for generating $[Y_N]$.

ii) For every x_i, calculate the best linear predictor of Y

$$E^*(Y \mid X = x_i) = \bar{Y} - \rho_d \cdot \frac{\sigma_y}{\sigma_x} (\bar{X} - x_i).$$

iii) For every x_i search through $[Z_N]$ until a z_j is found such that

$$q_j^2 = [z_j - E^*(Y \mid X = x_i)]^2 \leq D_1^2 \cdot \sigma_q^2, \qquad \text{for } q_j > 0$$

or

$$q_j^2 \leq D_2^2 \cdot \sigma_q^2, \qquad \text{for } q_j < 0$$

where σ_q^2 is the error variance given by (5) and D_1 and D_2 are the constants to be determined from the probability law of the random variable Y as illustrated in Fig. 1. If the probability law happens to be symmetric about the mean value, then $D_1 = D_2$.

iv) Choose the first z_j, which satisfies the condition of the previous step, to be paired with x_i as the corresponding y_i.

v) If none of the z_j satisfy the condition of step iii), choose that z_j to be paired with x_i, which gives the minimum squared error. Fig. 2 shows the computer flow chart for the above algorithm.

TABLE I
VALUES OF CORRELATION COEFFICIENT

Desired ρ_d	Standard Deviation $\sigma_\rho = \sqrt{1-\rho_d^2}$	Distribution		
		Uniform	Normal	Exponential
-0.8	0.0190	-0.8080	-0.7763	-0.8217
-0.6	0.0252	-0.5892	-0.5827	-0.6634
-0.4	0.0288	-0.3364	-0.3872	-0.5113
-0.2	0.0310	-0.1711	-0.2294	-0.2634
+0.2	0.0310	+0.1937	+0.2445	+0.2470
+0.4	0.0288	+0.4047	+0.3658	+0.4798
+0.6	0.0252	+0.6135	+0.5883	+0.6768
+0.8	0.0190	+0.8015	+0.7853	+0.8339

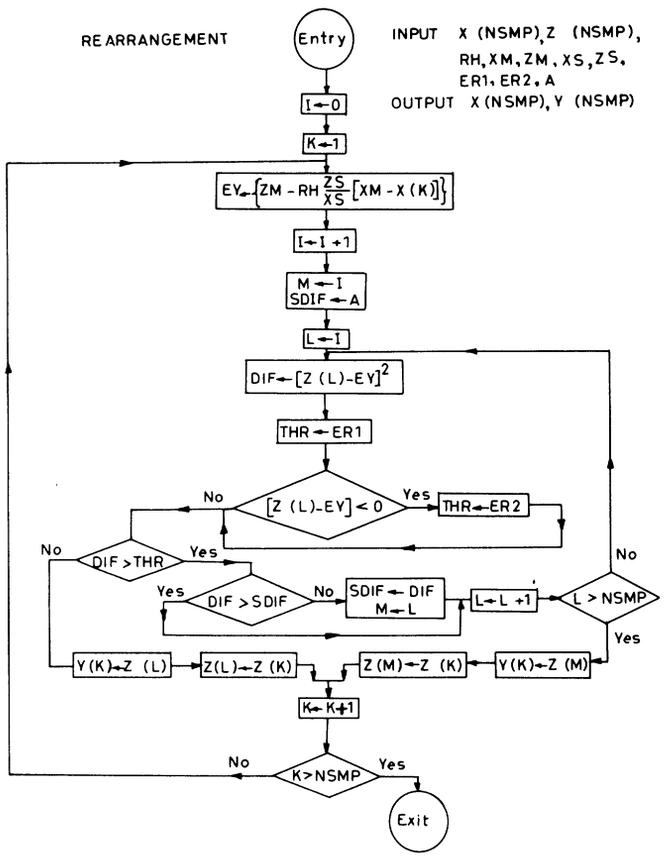

Fig. 2. Computer flowchart for algorithm for generation of correlated random number sequences (see Nomenclature).

RESULTS

The algorithm for the generation of correlated random number sequences was tested for the following three probability laws:

1) *Uniform:*

$$p(x) = \frac{1}{b-a}$$

with $b = 1$ and $a = 0$ so that the mean is 0.5 and the variance is 1/12.

2) *Normal:*

$$p(x) = \frac{1}{\sqrt{2\pi}} \cdot \exp(-x^2/2)$$

where the mean is 0 and the variance is 1.

3) *Exponential:*

$$p(x) = \lambda \cdot \exp(-x \cdot \lambda)$$

with $\lambda = 1$ so that the mean is 1 and the variance is 1.

Reference [1] was used as a source of independent random numbers. The sample size N was equal to 1000. The results are displayed in Table I.

DISCUSSION

i) The conditional expectation $E(Y \mid X)$ is in general a nonlinear function of the random variable X. Therefore, it might be possible to realize a better predictor by including higher order terms in X. However, with reference to the particular problem of the generation of correlated random number sequences, it is felt that the improvements (if there are any) in the final results may not be worth the additional complexity involved in nonlinear prediction.

ii) The generation of correlated random number sequences is essentially a simulation of the linear prediction process. Even if the distribution of the obtained correlation coefficients is difficult to evaluate, one can find out the moments. It can be shown that (see Appendix I for derivations)

$$E(\rho) = \rho_d \qquad \sigma_\rho^2 = \frac{1-\rho_d^2}{N}.$$

Expressions for the moments suggest that the simulation results will tend to improve with increasing magnitude of the correlation coefficient and sample size. This can be verified by generating a number of $[Y_N]$ sequences for the same $[X_N]$ and calculating the moments of the correlation coefficient. The simulation experiment was carried out for the case of the exponential distribution. For each $[X_N]$, one hundred $[Y_N]$ sequences were generated for every particular value of the desired correlation coefficient ρ_d. The results are displayed in Table II. It may be observed that the mean of the correlation coefficient approaches the desired correlation as the sample size increases. Reduction in the variance of the correlation coefficient with the increasing sample size and the magnitude

TABLE II
MOMENTS OF CORRELATION COEFFICIENT

Sample size N	Desired Correlation ρ_d	Mean Correlation $\bar{\rho} = \frac{1}{M}\sum_{j=1}^{M}\rho_j$	Theoretical Variance of ρ $\sigma_\rho^2 = \frac{1-\rho_d^2}{N}$	Sample Variance of ρ $\mathrm{Var}(\rho) = \frac{1}{M}\sum_{j=1}^{M}(\rho_j - \bar{\rho})^2$
	-0.8	-0.8139	0.0036	0.0043
	-0.6	-0.6728	0.0064	0.0069
	-0.4	-0.4899	0.0084	0.0095
100	-0.2	-0.3052	0.0096	0.0104
	+0.2	+0.3230	0.0096	0.0094
	+0.4	+0.5001	0.0084	0.0118
	+0.6	+0.6752	0.0064	0.0070
	+0.8	+0.8111	0.0036	0.0042
	-0.8	-0.7920	0.0018	0.0017
	-0.6	-0.6495	0.0032	0.0045
	-0.4	-0.4449	0.0042	0.0050
	-0.2	-0.2726	0.0048	0.0084
200	+0.2	+0.2433	0.0048	0.0065
	+0.4	+0.4700	0.0042	0.0046
	+0.6	+0.6205	0.0032	0.0041
	+0.8	+0.8061	0.0018	0.0015

X_N and Y_N are exponentially distributed. Number of Y_N sequences generated for each X_N is $M = 100$

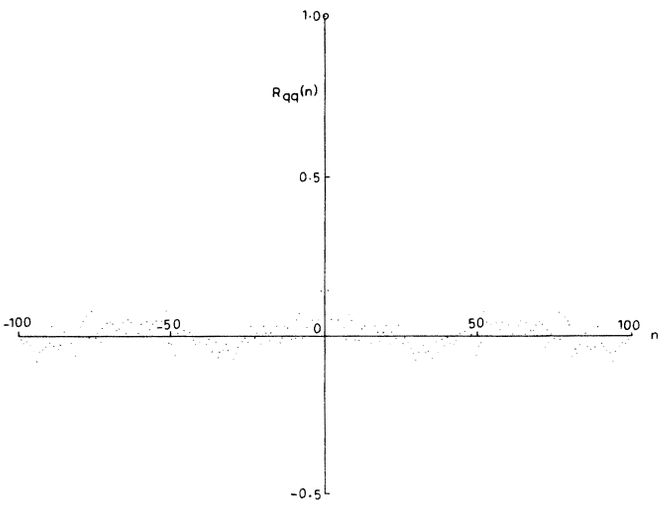

Fig. 3. Normalized autocorrelation function of error of prediction q for exponential distribution with correlation coefficient $\rho_d = 0.8$.

of the desired correlation confirms the above conjecture about the trend of the simulation results.

iii) The algorithm rearranges one random number sequence with reference to another. It is necessary to investigate the effect of this random rearrangement on the statistical characteristics of the random number sequence $[Y_N]$. Figs. 3 and 4 show graphs of the autocorrelation functions of the error of prediction and the random number sequence $[Y_N]$. The impulse-like nature of the autocorrelation functions suggests that the rearrangement has not introduced any significant amount of serial correlation in the random number sequence $[Y_N]$.

iv) Fig. 5 shows the empirical distribution function [11] of the error of prediction for the exponential distribution with correlation coefficient $\rho_d = 0.8$. The nature of the curve confirms the

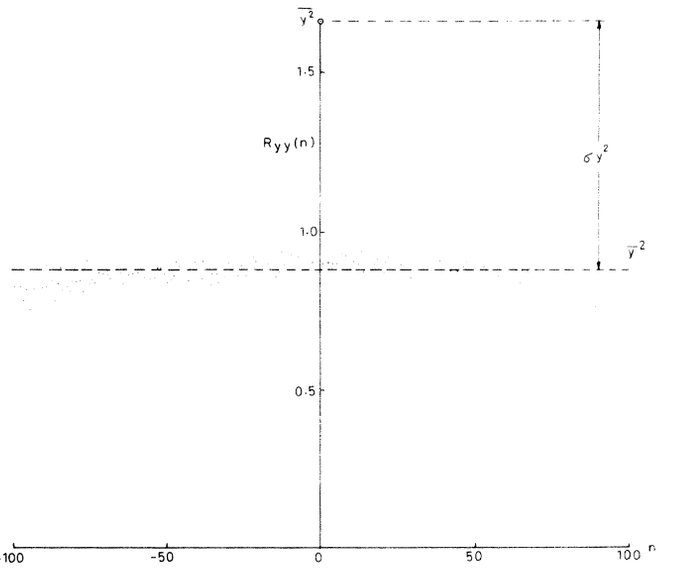

Fig. 4. Autocorrelation function of random number sequence $[Y_N]$ for exponential distribution with correlation coefficient $\rho_d = 0.8$.

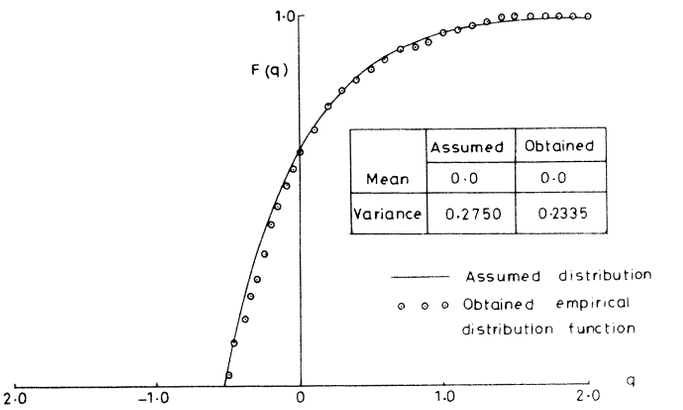

Fig. 5. Empirical distribution function of error of prediction q for exponential distribution with correlation coefficient $\rho_d = 0.8$.

assumption that the error of prediction has the same distribution as that of the random variable Y, with zero mean and variance given by (5).

v) Although the method has been demonstrated only for two random variable, it can easily be extended to multidimensional distributions (see Appendix II), since the theory of linear prediction does not impose any restrictions on the dimensionalities of the random variables.

vi) Since the method is based on the theory of optimal linear prediction (which does not assume an *a priori* knowledge of the probability law of the random variables under consideration), there is no reason which prevents it from being applied to the case of random variables having dissimilar marginal distributions. This particular aspect, in our opinion, deserves deeper investigation elsewhere.

vii) One of the referees has pointed out that the elements y_j of the sequence $[Y_N]$ as generated by the procedure described in the correspondence will, in general, not be identically and independently distributed (i.i.d). The individual distributions of the y_j are bound to differ from each other as the selection of each y_j depends on the value of the corresponding x_j. However, when averaged over the entire sample size, the distribution of $[Y_N]$ does match that of the random variable Y. Therefore, the procedure, although

only approximate in terms of the i.i.d. property, it is believed, would be useful in many practical applications.

Conclusions

The technique described in the correspondence enables generation of random number sequences with prescribed correlation. It has been tested and found satisfactory for three typical probability laws, viz., normal, uniform, and exponential, and hence appears applicable to any continuous distribution. The method is essentially an approximate one especially in the case of nonnormal distributions. Even with this limitation, the results of the Monte Carlo simulation procedure employed here are applicable, e.g., in the type of problems like the one mentioned in the introduction. Moreover, in the absence of a better technique, the present one provides a unified and plausible treatment for problems involving jointly distributed nonnormal random variables.

Acknowledgment

The authors wish to thank the referees for their comments and suggestions.

Appendix I
Derivation of the Moments of the Correlation Coefficient

Since the objective of this correspondence is to generate a prespecified cross correlation between the random number sequences, evaluation of the moments of the correlation coefficient provides an estimate of the errors that are likely to occur during simulation. It is possible to generate an infinite number of $[Y_N]$ sequences which will bear the same correlation with a given $[X_N]$. Therefore, expressions for the moments of correlation coefficient can be derived by taking expectations with respect to Y. The manner in which every y_j is selected can be assumed to be random provided the parent sequence $[Z_N]$ does not exhibit any specific trend. This condition is easily satisfied in almost every random or pseudorandom sequence.

Consider a simulation experiment in which we try to generate a number of $[Y_N]$ sequences bearing the same expected correlation ρ_d with $[X_N]$. Then each y_j will be distributed with the same probability law as that of Y with mean and variance given by

$$E(y_j) = E^*(Y \mid X = x_j) = \bar{Y} - \rho_d \cdot \frac{\sigma_y}{\sigma_x}(\bar{X} - x_j) \tag{7}$$

$$\sigma_{y_j}^2 = \sigma_y^2(1 - \rho_d^2). \tag{8}$$

The correlation coefficient between $[X_N]$ and $[Y_N]$ is given by

$$\rho = \frac{1}{N} \cdot \frac{1}{\sigma_x \cdot \sigma_y} \sum_{j=1}^{N} x_j y_j - \frac{\bar{X} \cdot \bar{Y}}{\sigma_x \cdot \sigma_y}. \tag{9}$$

The expectation of the correlation coefficient is given by

$$E(\rho) = \frac{1}{N} \cdot \frac{1}{\sigma_x \cdot \sigma_y} \sum_{j=1}^{N} x_j \cdot E(y_j) - \frac{\bar{x} \cdot \bar{y}}{\sigma_x \cdot \sigma_y}. \tag{10}$$

The order of expectation and summation can be interchanged since the two are independent of each other. Substituting for $E(y_j)$ from (7) and simplifying it we obtain

$$\begin{aligned} E(\rho) &= \frac{1}{N} \cdot \frac{1}{\sigma_x \cdot \sigma_y} \sum_{j=1}^{N} x_j \left[\bar{Y} - \rho_d \frac{\sigma_y}{\sigma_x}(\bar{X} - x_j)\right] - \frac{\bar{X} \cdot \bar{Y}}{\sigma_x \cdot \sigma_y} \\ &= \frac{\bar{X} \cdot \bar{Y}}{\sigma_x \cdot \sigma_y} - \frac{\rho_d \cdot \bar{X}^2}{\sigma_x^2} + \frac{\rho_d \cdot E(X^2)}{\sigma_x^2} - \frac{\bar{X} \cdot \bar{Y}}{\sigma_x \cdot \sigma_y} \\ &= \frac{\rho_d}{\sigma_x^2}[E(X^2) - (\bar{X}^2)] = \rho_d. \end{aligned} \tag{11}$$

The variance of the correlation coefficient is given by

$$\begin{aligned} \sigma_\rho^2 &= \operatorname{var}\left[\frac{1}{N} \cdot \frac{1}{\sigma_x \cdot \sigma_y} \cdot \sum_{j=1}^{N} (\bar{X} - x_j) \cdot (\bar{Y} - y_j)\right] \\ &= \frac{1}{N^2} \cdot \frac{1}{\sigma_x^2 \cdot \sigma_y^2} \cdot \left[\sum_{j=1}^{N} (\bar{X} - x_j)^2\right] \cdot \operatorname{var}(\bar{Y} - y_j) \\ &= \frac{1}{N^2} \cdot \frac{1}{\sigma_x^2 \cdot \sigma_y^2} \cdot N\sigma_x^2 \cdot \sigma_y^2(1 - \sigma_d^2) \\ &= \frac{1 - \rho_d^2}{N}. \end{aligned} \tag{12}$$

Appendix II
Prediction of Multidimensional Random Variables

Let $[X]$ and $[Y]$ denote the random vectors

$$[X] = [X_1, X_2, \cdots, X_m] \tag{13}$$
$$[Y] = [Y_1, Y_2, \cdots, Y_n].$$

Let $[\bar{X}]$ and $[\bar{Y}]$ denote their respective mean vectors

$$[\bar{X}] = [\bar{X}_1, \bar{X}_2, \cdots, \bar{X}_m] \tag{14}$$
$$[\bar{Y}] = [\bar{Y}_1, \bar{Y}_2, \cdots, \bar{Y}_n].$$

Then the best linear predictor of $[Y]$ for given $[X]$ is

$$[Y^*] = E^*[Y/X] = [\bar{Y}] - \{[\bar{X}] - [X]\} \cdot [\beta] \tag{15}$$

where $[\beta]$ is an $m \times n$ matrix of coefficients so chosen as to minimize the mean squared error $E[\{[Y] - [Y^*]\}^2]$.

The necessary and sufficient condition that $[Y^*]$ be a unique function of random vector $[X]$ (i.e., the projection of $[Y]$ onto $[X]$) satisfying the minimum mean squared error criterion is that the following equality of the corresponding product moments is satisfied:

$$\operatorname{cov}\{[Y^*], [X]\} = \operatorname{cov}\{[Y], [X]\} \tag{16}$$

$$\therefore [\Sigma_X] \cdot [\beta] = [\Sigma_{YX}]$$

$$[\beta] = [\Sigma_X]^{-1}[\Sigma_{YX}] \tag{17}$$

where $[\Sigma_X]$ is the variance-covariance matrix of the random vector $[X]$ and

$$[\Sigma_{YX}] = \begin{bmatrix} \operatorname{cov}(Y_1, X_1) & \operatorname{cov}(Y_2, X_1) \cdots \operatorname{cov}(Y_n, X_1) \\ \operatorname{cov}(Y_1, X_2) & \cdots \\ \vdots & \\ \operatorname{cov}(Y_1, X_m) & \operatorname{cov}(Y_2, X_m) \cdots \operatorname{cov}(Y_n, X_m) \end{bmatrix}.$$

(In case $[\Sigma_X]$ happens to be singular, the difficulty involved in the evaluation of the inverse can be circumvented by the technique given in [12].)

In the actual procedure, random vector $[X]$ (and hence its variance-covariance matrix $[\Sigma_X]$) is known, and the random vector $[Y]$ is so generated as to bear the covariances specified by $[\Sigma_{YX}]$ with the random vector $[X]$. Each component of $[Y]$ can be generated with the help of its corresponding component in the vector

$$[Y^*] = E^*[Y_{Y|X}] = [\bar{Y}] - \{[\bar{X}] - [X]\} \cdot [\Sigma_X]^{-1}[\Sigma_{YX}] \tag{18}$$

The error variances for different components of $[Y]$ can be obtained by substituting for $[Y^*]$ in the expression for the mean squared error.

NOMENCLATURE FOR FLOWCHART OF FIG. 2

I, J, K, L, M, N	Running indices.
$NSMP$	Sample size.
$X(\cdot)$	Random number sequence representing random variable X.
$Y(\cdot)$	Random number sequence representing random variable Y.
$Z(\cdot)$	Random number sequence used as a parent source to generate $Y(\cdot)$.
RH	Desired value of correlation coefficient, (ρ_d).
XM	Mean value of $X(\cdot)$, $(\bar{X})$.
ZM	Mean value of $Z(\cdot)$ is mean value of $Y(\cdot)$, $(\bar{Y})$.
XS	Standard deviation of $X(\cdot)$, (σ_x).
ZS	Standard deviation of $Z(\cdot)$ = standard deviation of $Y(\cdot)$, (σ_y).
EY	Best linear predictor of Y for given X, $[E^*(Y \mid X)]$.
DIF	Squared error, (q^2).
A	Some constant, large enough so that inequality $DIF > SDIF$ (i.e., $DIF > A$) is not satisfied at the first encounter.
$ER1$	Error bound for positive error is $D_1^2 \cdot \sigma_y^2(1 - \rho_d^2)$.
$ER2$	Error bound for negative error is $D_2^2 \cdot \sigma_y^2(1 - \rho_d^2)$.

If the distribution is symmetric about the mean then $ER1 = ER2$.

REFERENCES

[1] RAND Corp., *A Million Random Digits with 100,000 Normal Deviates*. New York: New York Free, 1955.

[2] J. M. Hammersley and D. C. Handscomb, *Monte Carlo Methods*. London: Methuen, 1965.

[3] M. Nakamura, "Obtaining a normal sequence with prescribed covariances from an independent uniform distribution of random variables," *IEEE Trans. Syst. Sci. Cybern.*, vol. SSC-4, no. 3, p. 191, July 1968.

[4] S. T. Li and J. L. Hammond, "Generation of pseudorandom numbers with specified univariate distributions and correlation coefficients," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-5, no. 5, pp. 557-561, Sept. 1975.

[5] Emanuel Parzen, *Modern Probability Theory and Its Applications*. New York: Wiley, 1960.

[6] M. H. De Groot and M. M. Rao, "Bayes estimation with convex loss," *Ann. Math. Statist.*, vol. 34, pp. 839-846, Sept. 1963.

[7] S. Sherman, "Nonmean-square error criteria," *IRE Trans. PGIT*, IT-4, pp. 125-126, Sept. 1958.

[8] Ralph Deutsch, *Estimation Theory*, Englewood Cliffs, NJ: Prentice-Hall, pp. 9-18, 1965.

[9] Emanuel Parzen, "A new approach to the synthesis of optimal smoothing and prediction systems," *Mathematical Optimization. Techniques*, Univ. of California, Berkeley, pp. 75-108, 1963.

[10] ——, "Statistical inference on time series by Hilbert space methods 1," Tech. Rep. no. 23, O.N.R. Contract 225(21), Statistics Dept., Stanford Univ., Jan. 2, 1959.

[11] Marek Fise, *Theory of Probability and Mathematical Statistics*. New York: Wiley, 1963.

[12] C. R. Rao, *Linear Statistical Inference and Its Applications*. New York: Wiley, 1965.

Book Reviews

Optimum Systems Control—A. P. Sage and C. C. White, III (Englewood Cliffs, NJ: Prentice-Hall, 1977, 2nd ed., 413 pp.). *Reviewed by George M. Siouris, Aerospace Guidance and Metrology Center (AFLC), Newark Air Force Station, Newark, OH 43055.*

This authoritative book is the second edition of Professor Sage's "Optimum Systems Control," providing a concise and timely analysis of optimal control theory. Since the appearance of the first edition almost ten years ago, optimum systems control theory has developed into a full-fledged area of study. As everyone who has worked in this area knows, its applications span a wide spectrum of scientific disciplines. Seen in this context, it is indeed a pleasure to welcome the second edition, which is undoubtedly another contribution to the optimal control theory. The present volume has been reduced from the first edition's 562 pp. to 413 pp. It has been completely revised, and the treatment is more geared for the engineer. The revisions have improved and modernized the overall presentation by the addition of new material.

The concepts presented in the second edition are well motivated and only introduced after sufficient justification has been given and provide a well-balanced blend of theory and practical applications. The text takes the student step-by-step through every aspect of modern control theory. At the end of each chapter there are many exercises of varying degree of difficulty for the student to work out and test his mastery of the material presented. Moreover, the text is amply illustrated with completely worked-out examples, which gives it an added dimension as a text.

The present edition is divided into ten chapters. The first chapter defines the problems of deterministic optimum control, state estimation, stochastic control, parameter estimation, and adaptive control, which arise in optimum systems control. A short summary of the subsequent chapters is also given. Chapter 2 provides an introduction to the calculus of extrema of two and more variables. Entitled "Calculus of Extrema and Single-Stage Decision Processes," the chapter discusses such familiar topics as unconstrained extrema, extrema of functions with equality constraints, and nonlinear programming. In Chapter 3, the authors treat the classical calculus of variations. Topics discussed include dynamic optimization without constraints, transversality conditions, sufficient conditions for weak extrema, unspecified terminal time problems, the Euler-Lagrange equations and transversality conditions, dynamic optimization with equality constraints and Lagrange multipliers, and dynamic optimization with inequality constraints. In the classical sense, the problem is to find the particular functions $y(x)$ and $z(x)$ which minimize (or maximize) the integral $I = \int_{x_1}^{x_2} f(x, y, y', z, z')\, dx$ subject to the constraint $\phi(x, y, y', z, z') = 0$. The general form of the transversality condition then takes the form $[(F - \sum_{j=1}^{n} y_j' F_{y_j'})dx + \sum_{j=1}^{n} F_{y_j'}\, dy_j]_a^b = 0$ where $F = f + \sum_{i=1}^{m} \lambda_i \phi_i$ and $\phi_i = 0$ $(i = 1, \cdots, m)$. With regard to modern optimal control theory, one is interested in minimizing the cost function $J(x) = \int_{t_0}^{t_f} \phi[x(t), \dot{x}(t), t]\, dt$ on the interval $[t_0, t_f]$. Chapter 4 is a natural extension of Chapter 3 where more general solutions are obtained. In particular, this chapter covers the Pontryagin maximum principle, the Weierstrass-Erdmann conditions, the Bolza problem with and without inequality constraints, and the Hamilton-Jacobi equations and continuous dynamic programming. All the topics addressed are handled very well, and the authors are on solid ground. Chapter 5 deals with optimum systems control examples. In this chapter an attempt has been made to illustrate those optimal control problems for which closed-loop solutions exist. The linear regulator, the linear servomechanism, the bang-bang control and minimum time prob-