205
4-19-79

DR. 2456

LA-7700-MS

Informal Report

MASTER

Bivariate Distributions with Given Marginals and Fixed Measures of Dependence

University of California

LOS ALAMOS SCIENTIFIC LABORATORY
Post Office Box 1663 Los Alamos, New Mexico 87545

An Affirmative Action/Equal Opportunity Employer

UNITED STATES
DEPARTMENT OF ENERGY
CONTRACT W-7405-ENG. 36

LA-7700-MS
Informal Report
UC-32
Issued: March 1979

Bivariate Distributions with Given Marginals and Fixed Measures of Dependence

Mark E. Johnson
Aaron Tenenbein*

*Visiting Scientist. Graduate School of Business Administration, New York University, New York, NY 10006.

BIVARIATE DISTRIBUTIONS WITH GIVEN MARGINALS AND FIXED MEASURES OF DEPENDENCE

by

Mark E. Johnson and Aaron Tenenbein

ABSTRACT

Two systematic approaches are given for constructing continuous bivariate distributions with specified marginals and fixed dependence measures. Both approaches are based on linear combinations of independent random variables and result in bivariate distributions which can attain the Frechet bounds. The dependence measures considered are the grade correlation coefficient and Kendall's τ. The joint distributions obtained are compared to those of M. Frechet (1951), R. L. Plackett (1965), D. Morgenstern (1956), and G. Kimmeldorf and A. Sampson (1975). Applications to testing for sensitivity in simulation models are discussed.

I. INTRODUCTION

In many simulation applications, it is required to generate dependent pairs of continuous random variables for which there is limited information on the joint distribution. For example, in a portfolio analysis simulation application, a joint distribution for stock and bond returns may have to be specified. Because of lack of data, it may be difficult to specify completely the joint distribution of stock and bond returns. However, it may be realistic to specify the marginal distributions and some measure of dependence between the random variables.

The problem of construction continuous bivariate distributions with specified marginals has been discussed in the literature, which is reviewed in section 2. A general method of constructing a bivariate distribution, whose marginal distribution functions are $F_1(x)$ and $F_2(y)$, is proposed by Nataf (1962) and can be represented as follows.

General Method

i. Consider any two continuous random variables U and V with probability density function h(u, v).

ii. Let $X' = H_1(U)$ and $Y' = H_2(V)$, where $H_1(u)$ and $H_2(v)$ are the cumulative distribution functions of U and V, respectively.

iii. Define:

$$X = F_1^{-1}(X') = F_1^{-1}(H_1(U)) \tag{1.1}$$

and

$$Y = F_2^{-1}(Y') = F_2^{-1}(H_2(V)) \tag{1.2a}$$

or

$$Y = F_2^{-1}(1 - Y') = F_2^{-1}(1 - H_2(V)) . \tag{1.2b}$$

Since X', Y', and 1 - Y' are uniformly distributed over the interval (0, 1), X defined by (1.1) and Y defined either by (1.2a) or (1.2b) will have a joint cumulative distribution whose marginal distribution functions are $F_1(x)$ and $F_2(y)$.

A procedure for constructing bivariate distributions should have the following characteristics.

a. The resulting joint distribution function F(x, y) should attain the upper or lower Frechet bounds given by Mardia (1970) as

$$\max\,[0,\ F_1(x) + F_2(y) - 1] \le F(x, y) \le [\min\ F_1(x),\ F_2(y)]\ , \qquad (1.3)$$

where $F_1(x)$ and $F_2(y)$ are the marginal distribution functions. If this were not the case, the resulting joint distribution will not allow for highly positive or negative correlated random variables in modeling situations.

b. The resulting joint distribution should readily lend itself for the development of random variate generation schemes for use in simulation models.

In this paper we develop two procedures for constructing bivariate distributions whose marginal distributions and measure of dependence, as given by Kendall's τ or the grade correlation coefficient ρ_s, are specified. Both of these procedures are based on this General Method, where the random variables U and V are obtained as linear combinations of U', V', and Z', where these latter three random variables are independent and identically distributed with probability density function g(t). The first procedure, called the WLC (weighted linear combination), defines

$$U = U' \qquad (1.4)$$

$$V = cU' + (1 - c)V' \qquad (1.5)$$

for $0 \le c \le 1$. The second procedure, called the TVR (trivariate reduction) is discussed by Mardia (1970). In this case U and V are defined as

$$U = U' + \beta Z' \qquad (1.6)$$

$$V = V' + \beta Z' \qquad (1.7)$$

for $0 \le \beta < \infty$.

Both of these procedures are easily adaptable to simulation models and contain members which attain the Frechet Upper and Lower Bounds. The WLC procedure attains the upper and lower bounds for c = 1 and Y defined by (1.2a) and 1.2b), respectively. Similarly, as β tends to ∞, the bounds are obtained in the TVR procedure. As a result both procedures have distributions for which ρ_s and τ take on any given value in the interval [-1, 1]. This latter fact is true by a general result proved by Tchen (1976).

In both systems there are two degrees of freedom in constructing (U, V); the weighting factor, β or c, and the underlying density function g(t). The weighting factor affects the measure of dependence. By considering various choices for $g(t)$, we can construct joint distributions with a fixed measure of dependence and fixed marginals. These joint distributions are extremely useful in assessing the sensitivity of a simulation model to this specification of $F(x, y)$. An example of this situation is the portfolio analysis simulation in which we could perform the simulation experiments with various $F(x, y)$ to assess its effect on the analysis.

Section 3 of this paper defines correlation measures ρ_s and τ and describes their properties. Section 4 discusses specific results of the WLC procedure for arbitrary choices for $g(t)$ and considers in detail the case of the normal, uniform, exponential, and double-exponential distributions. Formulae for the cumulative distributions of U and V are derived for use in the General Method equations (1) and (1.2a) or (1.2b). Expressions for τ and ρ_s as a function of the weighting factor c are obtained so that c can be determined for given values of τ or ρ_s. Section 5 discusses the properties of the various joint distributions obtained and discusses the application of these techniques to simulation models.

II. REVIEW OF BIVARIATE DISTRIBUTIONS WITH GIVEN MARGINALS

Several researchers have proposed continuous bivariate distributions which have specified marginals or, equivalently, uniform marginals on the interval $[0, 1]$. This section reviews these distributions with regard to their basic properties for use in simulation applications.

Morgenstern's (1956) bivariate distribution has uniform marginals on $[0, 1]$ and a correlation restricted by $-1/3 \leq \rho \leq 1/3$, so that the Frechet bounds are unobtainable. For many simulation applications, this range of dependence, as measured by ρ is too limited. Methods for generating random variates from the Morgenstern distribution are given by Johnson (1976). Farlie (1960) has generalized the Morgenstern distribution to include additional bivariate distributions having uniform marginals on $[0, 1]$. This generalized family has a limited range of dependence as measured by ρ and cannot attain the Frechet bounds.

Nataf (1962) employed his general method with $h(u, v)$ as the bivariate normal density to obtain a family which can have uniform marginals on $[0, 1]$. Unlike the Morgenstern distribution, Nataf's family attains the Frechet bounds. The correlation between U and V is $(6/\pi)\arcsin(\rho/2)$, so that the full range of ρ is possible. This family can be obtained as a special case of both the WLC and TVR approaches as discussed in sections 4 and 5.

Frechet (1951) proposed a bivariate distribution which is a convex combination of his boundary distributions (see equation 1.3). Because of the method of construction, this bivariate distribution attains its Frechet Bounds. However, the support of this distribution is restricted to the curves $F(X) = F(Y)$, and $F(X) + F(Y) = 1$ and does not include the case of independence.

Kimmeldorf and Sampson (1975a and 1975b) proposed a bivariate distribution having uniform marginals on [0, 1]. The support of this distribution consists of nonintersecting squares centered along the diagonal of the unit square. It is easy to show that the correlation is equal to (using their notation)

$$\rho = (n^3 - n + 3n\beta(\beta - n))/\beta^3,$$

where $n = [\beta]$, and this distribution attains the Frechet Bounds.

Plackett (1965) and Mardia (1967) proposed a bivariate distribution with normal marginals and a parameter ψ, which is called the coefficient of contingency. Mardia (1970) and Kowalski (1973) survey this distribution and Mardia (1970) considers the corresponding bivariate distribution with uniform marginals. For the latter case Mardia shows that the correlation coefficient is (equation (8.3.8) on page 61):

$$\rho = (\psi + 1)/(\psi - 1) - 2\psi \ln\psi \; / \; (\psi - 1)^2 \text{ for } 0 < \psi < \infty \quad (2.2)$$

This family of distributions can achieve its Frechet Bounds and the distribution has full support over the unit square. Mardia (1970) discusses a method of generating random variates from this distribution.

Additional distributions have been proposed by Tenenbein and Gargano (1978) and by Johnson and Ramberg (1977). These distributions are special cases of a general approach to constructing bivariate distributions with specified marginals. Discussion of these distributions is deferred to sections 4 and 5.

III. MEASURES OF DEPENDENCE

The dependence in a bivariate distribution can be described by various measures of association. The usual parametric measure is the Pearson product moment correlation coefficient ρ. Nonparametric measures of association include the grade correlation coefficient ρ_s and Kendall's τ. These latter measures are discussed by Kendall (1962) and Kruskal (1958) and can be defined as follows. Let X and Y be continuous random variables having some joint probability density function. Let (X_1,Y_1), (X_2,Y_2), and (X_3,Y_3) be three independent pairs of observations having the same joint density function. Then

$$\tau = \tau(X, Y) - 2\Pr[(X_1 - X_2)(Y_1 - Y_2) > 0] - 1 \quad (3.1)$$

$$\rho_s = \rho_s(X, Y) = 6\Pr[(X_1 - X_2)(Y_1 - Y_3) > 0] - 3 \quad (3.2)$$

It is easy to show that $|\rho_s|$ and $|\tau|$ are invariant under strictly monotone transformations of X and Y. If X and Y are defined by equations (1.1) and (1.2a), respectively, it follows that for the General Method

$$\rho_s = \rho_S(X, Y) = \rho_S(U, V) \quad (3.3)$$

$$\tau = \tau(X, Y) = \tau(U, V) . \quad (3.4)$$

If X and Y are defined by equations (1.1) and 1.2b), respectively, then

$$\rho_s = \rho_S(X, Y) = -\rho_S(U, V) \quad (3.5)$$

$$\tau = \tau(X, Y) = -\tau(U, V) . \quad (3.6)$$

Throughout this paper we are using τ and ρ_s as measures of dependence rather than ρ. There are two reasons for this approach. Firstly, ρ is not defined if either of the random variables, X or Y, have infinite variances; however, ρ_s and τ are always defined. If ρ is used as a measure of association, the choices of marginal distributions for X and Y are limited. Secondly, the invariance property of $|\rho_s|$ and $|\tau|$ makes them more convenient as measures of dependence for use in the General Method. From equations (3.3),(3.4), (3.5), and (3.6) it is evident that we need only specify the values of ρ_s or τ. This immediately specifies the values of ρ_s (U, V) or τ(U, V) regardless of the form of the marginal distributions of X and Y. On the other hand, the value of ρ must be computed for every choice of marginal distributions. This makes it difficult to use ρ as a measure of dependence in the General Method.

Other expressions for τ(U, V) and ρ_s (U, V) are useful evaluation purposes. Kruskal (1958) shows that

$$\rho_s (U, V) = \rho[H_1(U), H_2(V)] , \qquad (3.7)$$

where $H_1(u)$ and $H_2(v)$ are the marginal distribution functions of U and V, respectively. Since E $H_2(V)$ = 1/2 and V $H_1(U)$ = V $H_2(V)$ = 1/12, it follows that

$$\rho_s(U, V) = 12E[H_1(U)H_2(V)] - 3. \qquad (3.8)$$

Also, it is easy to show that

$$\tau(U, V) = 4 \Pr[U_1 > U_2, V_1 > V_2] - 1 \qquad (3.9)$$

Equation (3.7) immediately implies that if U and V have uniform marginal distributions over the interval (0, 1), then the grade correlation coefficient equals the Pearson product moment correlation coefficient. Hence, for the Kimmeldorf-Sampson distribution system discussed in section 2, equation (2.1) gives the grade correlation coefficent. Similarly, for both Plackett's and Mardia's C-type contingency distributions equation, (2.2) represents the grade correlation coefficent.

In practive, we may have to estimate ρ_s and τ . The former can be estimated by Spearman's rank correlation coefficient r_s and Kendall's τ can be estimated by Kendall's t, as discussed by Kendall (1962) and Kruskal (1958).

IV. THE WEIGHTED LINEAR COMBINATION

The WLC procedure constructs a joint distribution whose marginal distribution functions are specified to be $F_1(x)$ and $F_2(y)$ and whose measure of dependence can be specified to be ρ_s or τ. This procedure begins with two independent random variables U' and V' having the same density function g(t). The random variables U and V are defined by equations (1.4) and (1.5), respectively; then X and Y are obtained from the General Method using equations (1.1) and (1.2a) if the dependence measure is positive or equations (1.1) and (1.2b) if the dependence measure is negative.

In order to apply the WLC procedure, we must obtain expressions for $H_1(u)$, $H_2(v)$, $\rho_s(U, V)$, and $\tau(U,V)$, in terms of c and g(t). The values of $H_1(u)$ and $H_2(v)$ allow us to apply the General Method for a given choice of c and g(t). The expressions for $\tau(U, V)$ and $\rho_s(U, V)$ allow us to specify c for a given choice of g(t) in terms of the required value of either τ or ρ_s.

From equations (1.4) and (1.5) it is obvious that

$$H_1(u) = \int_{-\infty}^{u} g(t)dt \tag{4.1}$$

$$H_2(v) = \iint_{R_1} g(u')g(v')du'dv', \tag{4.2}$$

where $R_1 = \{(u', v'): cu' + (1 - c)v' \leq v\}$ and the joint density function of U and V is

$$h(u, v) = \frac{1}{1 - c} g(u)g\left(\frac{v - cu}{1 - c}\right) \tag{4.3}$$

over the appropriate region. The following theorem gives general expressions for $\rho_s(U, V)$ and $\tau(U, V)$.

<u>Theorem 1</u>

Let U and V be random variables defined by (1.4) and (1.5) with a joint density function given by (4.3) and with marginal distribution functions given by (4.1) and (4.2). It follows that

$$\rho_s(U, V) = 12 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H_1(u)H_2(v)h(u, v)dudv - 3 \tag{4.4}$$

$$\tau(U, V) = 4 \int_{-\infty}^{0} G_2\left(\frac{1 - c}{c} t\right) g_2(t)dt, \tag{4.5}$$

where

$$g_2(t) = \int_{-\infty}^{\infty} g(w + t)g(w)dw \tag{4.6}$$

and

$$G_2(t) = \int_{-\infty}^{t} g_2(x)dx. \tag{4.7}$$

Proof

Equation (4.4) follows directly from equation (3.8). To prove (4.5) consider equations (3.9), (1.4), and (1.5). This implies that

$$\tau(U, V) = 4\Pr[U_1' - U_2' > 0,\ c(U_1' - U_2') + (1 - c)(V_1' - V_2') > 0] - 1$$

where U_1', U_2', V_1', and V_2' are independent and have the same density function g(t). Define

$$S = U_1' - U_2$$
$$T = V_1' - V_2 .$$

T and S are independent random variables which are symmetric about 0 and have the density function and distribution function given by (4.6) and (4.7), respectively. Thus

$$\tau(U, V) = 4\Pr[S > 0,\ cS + (1 - c)T > 0] - 1$$
$$= 4E\{\Pr[S > \max\left(0, -\frac{(1 - c)t}{c}\right) | T = t]\} - 1$$

Since S and T are independent,

$$\tau(U, V) = 4\int_{-\infty}^{0} (1 - G_2\left(-\frac{(1 - c)t}{c}\right))g_2(t)dt + 4\int_{0}^{\infty} [1 - G_2(0)]g_2(t)dt - 1$$

Since T is symmetric about 0, the above reduces to equation (4.5).

Using equations (4.1), (4.2), (4.3), (4.4), and (4.5), we have evaluated $H_1(u)$, $H_2(v)$, $h(u, v)$, $\tau(U, V)$, and $\rho_s(U, V)$ for all values of c and for the cases where g(t) is uniform, exponential, and double exponential. These integrals were quite tedious to perform; however, the process was facilitated by the use of MACSYMA (1977), an MIT computer program which performs multiple integration and algebraic symbol manipulation. The corresponding results were obtained for the normal distribution by using the following properties of the bivariate normal distribution given by Kruskal (1958):

$$\rho_s = \frac{6 \arcsin(\rho/2)}{\pi} \qquad (4.8)$$

$$\tau = (2 \arcsin \rho)/\pi . \qquad (4.9)$$

The normal distribution is equivalent to Nataf's (1962) distribution.

Table 1 shows the expressions for $H_1(u)$ and $H_2(v)$. Tables 2 and 3 show the formulae for ρ_s and τ as functions of c and provide numerical comparisons for the four distributions discussed. Table 4 shows the joint distribution h(u, v). Note that c = 0 implies independence, in which case h(u, v) = g(u)g(v). As an example, suppose we wish to generate random variables whose ρ_s = -0.4 using the double exponential distribution. Using the appropriate formula in Table 2 or interpolating in Table 3, the value c = 0.299 yields a value of ρ_s = (U, V) = 0.4. We would use this value of c in equations (1.4) and (1.5) to generate U and V. We would then use equations (1.1) and (1.2b) to generate X and Y using the appropriate values of $H_1(u)$ and $H_2(v)$ from Table 1. The resulting random variable pair (X, Y) is a sample from a bivariate distribution with ρ_s = -0.4 and marginal distribution functions $F_1(x)$ and $F_2(y)$.

The double exponential, normal, and uniform distributions are members of the class of exponential power distributions discussed by Box and Tiao (1973; page 156), where

$$g(t) = K \exp[-0.5|t|^q] .$$

The double exponential and normal distributions correspond to q = 1 and 2, respectively, whereas the uniform distribution is the limiting case as $q \to \infty$. The double exponential has the heaviest tails, the normal distribution has moderate tails, and the uniform distribution has diffuse tails. The resulting three bivariate distributions obtained by the WLC method with the corresponding values of c adjusted to give a constant measure of dependence will show different probabilities within the (X', Y') region, where X' = $F_1(X) = H_1(U)$ and Y' = $F_2(Y) = H_2(V)$. As can be deduced from Table 4, the normal and double exponential provide full support for the joint distribution of (X', Y') over the region (0 < X' < 1, 0 < Y' < 1). The uniform distribution results in a bivariate distribution for which the probability in portions of the upper left and lower right regions of the unit square is zero.

The exponential and normal distributions are members of the gamma family, where

$$g(t) = \frac{1}{\Gamma(n)} t^{n-1}e^{-t} . \tag{4.11}$$

The exponential is near the extreme of positive skewness; the normal is the limiting case as $n \to \infty$. The resulting two bivariate distributions obtained by the WLC method with the corresponding value of c adjusted to give a constant measure of dependence will show different probabilities within the (X', Y') range. Table 4 implies that the exponential results in a bivariate distribution which can assign zero probability in a portion of the lower right-hand region but has complete support everywhere else within the unit square.

The implications of these comparisions on applications to simulation methods are discussed in section 6.

V. THE TRIVARIATE REDUCTION PROCEDURE

The TVR procedure constructs a joint continuous distribution with fixed marginals and measure of dependence in a similar fashion to the WLC procedure. The difference is that the random variables U and V are constructed by equations (1.6) and (1.7), respectively. In this section we derive the corresponding results for $H_1(u)$, $H_2(v)$, h(u, v), ρ_s(U, V), and (U, V) as a function of β and g(t) and consider the same choices for g(t) as in section 4.

From equations (1.6) and (1.7) it follows that

$$H_1(u) = H_2(v) = \iint_{R_3} g(u')g(z')du'dz' \tag{5.1}$$

where

$$R_3 = \{(u', z'): \ u' + \beta z' \leq u\} \tag{5.2}$$

$$h(u, v) = \int_{-\infty}^{\infty} g(u - \beta z)g(v - \beta z)g(z)dz \tag{5.3}$$

The following theorem gives general expressions for ρ_s(U, V) and (U, V).

Theorem 2

Let U and V be identically distributed random variables defined by (1.6) and (1.7) with a joint density function given by (5.3) and common marginal distribution function given by (5.1). It follows that $\rho_s(U, V)$ is given by equation (4.4) and $\tau(U, V)$ is given by

$$\tau(U, V) = 4 \int_{-\infty}^{\infty} [G_2(\beta t)]^2 g_2(t)dt - 1, \tag{5.4}$$

where $G_2(t)$ and $g_2(t)$ are given by equations (4.7) and (4.6), respectively.

Proof

Equation (4.4) has been shown in Theorem 1. To show that (5.4) is true, consider equations (3.9), (1.6), and (1.7). It follows that

$$\tau(U, V) = 4\Pr[U_1' - U_2' + \beta(Z_1' - Z_2') > 0,$$
$$V_1' - V_2' + \beta(Z_1' - Z_2') > 0] - 1,$$

where U_i', V_i', and Z_i' $(i = 1, 2)$ are independent and have the same density function g(t). Let

$$T = Z_1' - Z_2'$$
$$S_1 = U_1' - U_2'$$
$$S_2 = V_1' - V_2'.$$

Then T, S_1, S_2 are independent random variables, which are symmetric about zero and have the density function and distribution function given by (4.6) and (4.7), respectively. Thus

$$\tau(U, V) = 4E\{\Pr[S_1 > -\beta t, S_2 > -\beta t | T = t]\} - 1$$
$$= 4 \int_{-\infty}^{\infty} [1 - G_2(-\beta t)]^2 g_2(t)dt - 1 .$$

As a result equation (5.4) follows.

Using equations (5.1), (5.3), (4.4), and (5.4) the same computations were carried out for the TVR procedure as have been done for the WLC procedure of the previous section. Table 5 gives the common marginal distribution function of U and V. Tables 6 and 7 yield the expressions for $\rho_s(U, V)$ and $\tau(U, V)$ as functions of β and provide numerical comparisons. Table 8 gives the joint density function of U and V.

The same trends in the support of the corresponding bivariate distributions with uniform marginals are evident for the TVR procedure as for the WLC procedure. The TVR procedure is symmetric in the sense that the marginal distributions of U and V are the same. The normal distribution case of the TVR procedure is equivalent to using Nataf's (1962) joint distribution.

VI. APPLICATION TO SIMULATION MODELS

Many simulation models require the specification of a joint continuous bivariate distribution as input. Eilen and Fowkes (1973) and Hall (1977) consider the problem in risk simulation. If there is an adequate theory or sufficient data upon which to base a specific bivariate distribution the problem is well defined. Johnson and Kotz (1972) have compiled a large selection of bivariate and multivariate distributions and Fishman (1973) and Johnson (1976) have discussed methods of generating random variates from given bivariate distributions.

In many situations there really is no adequate theory or sufficient data to be able to specify a unique bivariate distribution. However, it may be realistic to specify the marginal distributions of the random variables and a measure of dependence between them. If this is the case, the problem is not well defined because there are many bivariate distributions having these properties. A variety of assumed joint distributions must be used to assess their sensitivity to the results of the simulation model.

In this report we have presented the TVR and WLC procedures, each of which constructs bivariate continuous distributions with fixed marginals and depen- dence measure. By using either the TVR procedure or WLC procedure with g(t) being double exponential, normal, and uniform, we can assess the effect of tail weight in g(t) on the final results of the simulation model. By using either procedure with g(t) being exponential and normal, we would be able to assess the effect of skewness in g(t) on the final results of the simulation model.

For choices of g(t) discussed in this paper we have not been able to consider more extreme heavy-tailed distributions, such as the Cauchy, or more skewed distributions, such as the Pareto. These latter distributions are difficult to apply to construction schemes because closed-form solutions for the cumulative distributions of U and V are difficult to obtain. The intermediate choices for the exponential power family (4.10) were likewise intractable. Intermediate choices for the gamma family (4.11), where n is an integer, can be pursued.

The one-parameter families considered in section 2, which achieved the Frechet bounds, can also be applied to simulation models. The correlation coefficients given by equations (2.1) and (2.2) for Kimmeldorf-Sampson and the Plackett-Mardia distributions, respectively, also represent the grade correlation coefficient because of the results of section 3. Consequently, the Kimmeldorf-Sampson bivariate distribution can be used as h(u, v) in the General Method with β adjusted to yield a specified value of ρ_S. Similarly, Plackett's or Mardia's distribution can be used with ψ adjusted to yield a specific value of ρ_S .

In our methods we have adopted a systematic approach in construction bivariate distributions with fixed marginals and dependence measure. We have not investigated the possibility of using these bivariate distributions to fit real data. Preliminary work by Johnson (1976) indicates that it is feasible to fit these distributions using a modified maximum likelihood scheme.

<u>Table 1</u>

$H_1(u)$ AND $H_2(v)$ FOR THE WEIGHTED LINEAR COMBINATION PROCEDURE

$H_1(u)$ (DISTRIBUTION)	$H_2(v)$
$H_1(u) = u$ for $0 \le u \le 1$ (UNIFORM)	$v^2/(2b(1 - b)$ for $0 \le v \le b$ $(2v - b)/(1 - b)$ for $b \le v \le 1 - b$ $1 - (1 - v)^2 /(2b(1 - b))$ for $1 - b \le v \le 1$ where $b = \min(c, 1 - c)$
$H_1(u) = \Phi(u)$ (STANDARD NORMAL)	$\Phi(v/\sqrt{c^2 + (1 - c)^2})$
$H_1(u) = \frac{1}{2}\exp(u)$ for $u < 0$ $= 1 - \frac{1}{2}\exp(-u)$ for $u > 0$ (DOUBLE EXPONENTIAL)	<u>Case 1: $c \neq \frac{1}{2}$</u> $[(1 - c)^2\exp(v/(1 - c)) - c^2\exp(v/c)]/(2(1 - 2c))$ for $v < 0$ $1 - [(1 - c)^2\exp(-v/(1 - c)) - c^2\exp(-v/c)]/(2(1 - 2c))$ for $v > 0$ <u>Case 2: $c = \frac{1}{2}$</u> $(1 - v)\exp(2v)/2$ for $v < 0$ $1 - (v + 1)\exp(-2v)/2$ for $v > 0$
$H_1(u) = 1 - \exp(-u)$ for $u > 0$ (EXPONENTIAL)	$1 - [(1 - c)\exp(-v/(1 - c) - c\exp(-v/c)]/(1 - 2c)$ for $v > 0$ and $c \neq \frac{1}{2}$ $1 - (2v + 1)\exp(-2v)$ for $v > 0$ and $c = \frac{1}{2}$

Table 2

FORMULAE FOR ρ_s AND τ AS FUNCTIONS OF c IN THE WEIGHTED LINEAR COMBINATION PROCEDURE

DISTRIBUTION	$\rho_s(U, V)$	$\tau(U, V)$
UNIFORM	$\frac{c(10 - 13c)}{10(1 - c)^2}$ for $0 < c < .5$; $\frac{3c^3 + 16c^2 - 11c + 2}{10c^3}$ for $.5 < c < 1$	$\frac{4c - 5c^2}{6(1 - c)^2}$ for $0 < c < .5$; $\frac{11c^2 - 6c + 1}{6c^2}$ for $.5 < c < 1$
NORMAL	$\frac{6}{\pi} \arcsin \frac{c}{2\sqrt{c^2 + (1 - c)^2}}$	$\frac{2}{\pi} \arcsin \frac{c}{\sqrt{c^2 + (1 - c)^2}}$
DOUBLE EXPONENTIAL	$\frac{c(9 - 18c^2 + 14c^3 - 3c^4)}{2(2 - c)^2}$	$\frac{c(3 + 3c - 2c^2)}{4}$
EXPONENTIAL	$\frac{c(3 - 2c)}{2 - c}$	c

Table 3

NUMERICAL VALUES OF ρ_s AND τ IN THE WEIGHTED LINEAR COMBINATION PROCEDURE

	Value of $\rho_s(U, V)$				Value of $\tau(U, V)$			
c	UNIFORM	NORMAL	DOUBLE EXPONENTIAL	EXPONENTIAL	UNIFORM	NORMAL	DOUBLE EXPONENTIAL	EXPONENTIAL
0	0	0	0	0	0	0	0	0
.1	.107	.106	.122	.147	.072	.070	.082	.100
.2	.231	.232	.259	.289	.156	.156	.176	.200
.3	.373	.379	.401	.424	.255	.258	.279	.300
.4	.533	.537	.542	.550	.370	.374	.388	.400
.5	.700	.690	.674	.667	.500*	.500*	.500*	.500*
.6	.837	.819	.789	.771	.630	.626	.612	.600
.7	.924	.912	.883	.862	.745	.742	.721	.700
.8	.972	.967	.950	.933	.844	.844	.824	.800
.9	.994	.993	.988	.982	.928	.930	.918	.900
1.0	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

* It can be shown that $\tau = \frac{1}{2}$ when $c = \frac{1}{2}$ for any continuous distribution (see equation (4.5)).

Table 4

THE JOINT DENSITY FUNCTION: h(u, v)
FOR THE WEIGHTED LINEAR COMBINATION PROCEDURE

UNDERLYING DENSITY	h(u, v)	SUPPORT
UNIFORM $g(t) = 1$ $(0 < t < 1)$	$1/(1 - c)$	$cu < v < cu + 1 - c$ $0 < u < 1$
NORMAL $N(0, 1)$	UNDER BIVARIATE NORMAL $E(U) = E(V) = 0$ $\sigma_u^2 = 1$ $\sigma_v^2 = c^2 + (1 - c)^2$ $\rho = c/\sqrt{c^2 + (1 - c)^2}$	$-\infty < u < \infty$ $-\infty < v < \infty$
DOUBLE EXPONENTIAL $g(t) = \frac{1}{2} \exp(-\|t\|)$ $(-\infty < t < \infty)$	$(\exp[-\|u\| - \|v - cu\|/(1 - c)])/(4(1 - c))$	$-\infty < u < \infty$ $-\infty < v < \infty$
EXPONENTIAL $g(t) = \exp(-t)$ $(0 < t < \infty)$	$(\exp(-v - u + 2cu))/(1 - c)$	$cu < v < \infty$ $0 < u < \infty$

Table 5

$H_1(u) = H_2(u)$ FOR THE TRIVARIATE REDUCTION PROCEDURE

DISTRIBUTION	$H_1(u) = H_2(u)$	
UNIFORM	Case 1: $\beta \le 1$	
	$u^2/2\beta$	for $0 \le u \le \beta$
	$u - \beta/2$	for $\beta \le u \le 1$
	$1 - (u - \beta - 1)^2/2\beta$	for $1 \le u \le 1 + \beta$
	Case 2: $\beta > 1$	
	$u^2/2\beta$	for $0 \le u \le 1$
	$(2u - 1)/2\beta$	for $1 \le u \le \beta$
	$1 - (u - \beta - 1)^2/2\beta$	for $\beta \le u \le 1 + \beta$
STANDARD NORMAL	$\Phi(u/\sqrt{1 + \beta^2})$	
DOUBLE EXPONENTIAL	Case 1: $\beta \ne 1$	
	$(\beta^2 \exp(u/\beta) - \exp(u))/(2\beta^2 - 2)$	for $u < 0$
	$1 - (\beta^2 \exp(-u/\beta) - \exp(-u))/(2\beta^2 - 2)$	for $u > 0$
	Case 2: $\beta = 1$	
	$(2 - u)\exp(u)/4$	for $u < 0$
	$1 - (u + 2)\exp(-u)/4$	for $u > 0$
EXPONENTIAL	$1 - (\exp(-u) - \beta \exp(-u/\beta))/(1 - \beta)$	for $u \ge 0$ and $\beta \ne 1$
	$1 - (u + 1)\exp(-u)$	for $u \ge 0$ and $\beta = 1$

Table 6

FORMULAE FOR ρ_s AND τ AS FUNCTIONS OF β IN THE TRIVARIATE REDUCTION METHOD

DISTRIBUTION		ρ_s	τ
UNIFORM	$\beta \le 1$	$\frac{19\beta^4 - 126\beta^3 + 210\beta^2}{210}$	$\frac{\beta^4 - 6\beta^3 + 10\beta^2}{15}$
	$1 \le \beta \le 2$	$\frac{\beta^7 - 14\beta^6 + 84\beta^5 - 280\beta^4 + 770\beta^3 - 672\beta^2 + 238\beta - 24}{210\beta^3}$	$\frac{15\beta^2 - 14\beta + 4}{15\beta^2}$
	$\beta \ge 2$	$\frac{105\beta^3 - 105\beta + 52}{105\beta^3}$	$\frac{15\beta^2 - 14\beta + 4}{15\beta^2}$
NORMAL		$\frac{6}{\pi} \arcsin\left(\frac{\beta^2}{2(\beta^2 + 1)}\right)$	$\frac{2}{\pi} \arcsin\left(\frac{\beta^2}{\beta^2 + 1}\right)$
DOUBLE EXPONENTIAL		$\frac{16\beta^9 + 152\beta^8 + 588\beta^7 + 1122\beta^4 + 1104\beta^3 + 555\beta^2 + 132\beta + 12}{2(\beta + 1)^4(\beta + 2)^2(2\beta + 1)}$	$\frac{32\beta^7 + 125\beta^6 + 161\beta^5 + 90\beta^4 + 22\beta^3 + 2\beta^2}{2(2\beta + 1)^4(\beta + 1)^3}$
EXPONENTIAL		$\frac{\beta^2(2\beta^2 + 9\beta + 6)}{(\beta + 1)^2(2\beta + 1)(\beta + 2)}$	$\frac{2\beta^2}{(\beta + 1)(2\beta + 1)}$

Table 7

NUMERICAL VALUES OF ρ_s AND τ
IN THE TRIVARIATE REDUCTION PROCEDURE

	Value of ρ_s				Value of τ			
$\frac{\beta}{1 + \beta}$	UNIFORM	NORMAL	DOUBLE EXPONENTIAL	EXPONENTIAL	UNIFORM	NORMAL	DOUBLE EXPONENTIAL	EXPONENTIAL
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.1	0.012	0.012	0.018	0.027	0.008	0.008	0.012	0.018
0.2	0.053	0.056	0.076	0.099	0.036	0.037	0.051	0.067
0.3	0.140	0.148	0.178	0.204	0.093	0.099	0.120	0.138
0.4	0.285	0.295	0.315	0.331	0.191	0.199	0.216	0.229
0.5	0.490	0.483	0.473	0.472	0.333*	0.333*	0.333*	0.333*
0.6	0.702	0.675	0.637	0.617	0.496	0.487	0.466	0.450
0.7	0.855	0.833	0.786	0.756	0.649	0.641	0.606	0.576
0.8	0.945	0.936	0.904	0.877	0.783	0.781	0.747	0.711
0.9	0.988	0.987	0.977	0.965	0.900	0.900	0.881	0.853
1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

* It can be shown that $\tau = 1/3$ when $\beta = 1$ for any continuous distribution (see equation (5.4)).

Table 8a

THE JOINT DENSITY FUNCTION h(u, v)
FOR THE TRIVARIATE REDUCTION PROCEDURE (THE UNIFORM DISTRIBUTION)

Case 1: $\beta \leq 1$

h(u, v)	Support	
1	$\beta < u < 1$	$, \beta < v < 1$
$u - \beta/2$	$\beta < u < 1$	$, 0 < v < \beta$
$v - \beta/2$	$\beta < v < 1$	$, 0 < u < \beta$
$1 - (u - 1)/\beta$	$1 < u < \beta + 1$	$, \beta < v < 1$
$1 - (v - 1)/\beta$	$1 < v < \beta + 1$	$, \beta < u < 1$
$v/\beta - (u - 1)/\beta$	$1 < u < 1 + v$	$, 0 < v < \beta$
$u/\beta - (v - 1)/\beta$	$1 < v < 1 + u$	$, 0 < u < \beta$
v/β	$v < u < \beta$	$, 0 < v < \beta$
u/β	$u < v < \beta$	$, 0 < u < \beta$
$1 - (u - 1)/\beta$	$1 < v < u$	$, 1 < u < \beta + 1$
$1 - (v - 1)/\beta$	$1 < u < v$	$, 1 < v < \beta + 1$

Case 2: $\beta \geq 1$

h(u, v)	Support
$u/\beta - (v - 1)/\beta$	$0 < u < \beta$, $\max(1, u) < v < 1 + u$
$v/\beta - (u - 1)/\beta$	$0 < v < \beta$, $\max(1, v) < u < 1 + v$
$1 - (v - 1)/\beta$	$\beta < u < v$, $\beta < v < 1 + \beta$
$1 - (u - 1)/\beta$	$\beta < v < u$, $\beta < u < 1 + \beta$
u/β	$0 < u < 1$, $u < v < 1$
v/β	$0 < v < 1$, $v < u < 1$

Table 8b

THE JOINT DENSITY FUNCTION $h(u, v)$
FOR THE TRIVARIATE REDUCTION PROCEDURE
(THE DOUBLE EXPONENTIAL DISTRIBUTION)

$h(u, v)$	Support
$h_1(u, v) = \left\{ \begin{array}{l} (2\beta^2 + \beta)\exp(v - v/\beta - u) \\ + (\beta - 2\beta^2)\exp(v - u - u/\beta) \\ - \exp(-v - u) \end{array} \right\} \div (16\beta^2 - 4)$ for $\beta \neq \frac{1}{2}$ $h_1(u, v) = [(4v + 3)\exp(-v - u) - \exp(v - 3u)]/16$ for $\beta = \frac{1}{2}$	$u > v > 0$
$h_2(u, v) = h_1(v, u)$	$v > u > 0$
$h_3(u, v) = \left\{ \begin{array}{l} (2\beta + 1)\exp(v - u) \\ -\beta \exp(v - u + v/\beta) \\ -\beta \exp(v - u - u/\beta) \end{array} \right\} \div (8\beta + 4)$	$u > 0 > v$
$h_4(u, v) = h_3(v, u)$	$v > 0 > u$
$h_5(u, v) = h_1(-u, -v)$	$0 > v > u$
$h_6(u, v) = h_1(-v, -u)$	$0 > u > v$

Table 8c

THE JOINT DENSITY FUNCTION
FOR THE TRIVARIATE REDUCTION PROCEDURE
(THE NORMAL AND EXPONENTIAL DISTRIBUTIONS)

DISTRIBUTION	$h(u, v)$
EXPONENTIAL	$\exp(-u - v)(1 - \exp((2 - 1/\beta)\min(u, v))/(1 - 2\beta)$ for $u, v \geq 0$ and $\beta \neq \frac{1}{2}$ $2 \exp(-u - v)\min(u, v)$ for $u, v \geq 0$ and $\beta = \frac{1}{2}$
STANDARD NORMAL	BIVARIATE NORMAL WITH $E(U) = E(V) = 0$ $\sigma_U^2 = \sigma_V^2 = 1 + \beta^2$ $\rho = \beta^2/(1 + \beta^2)$

REFERENCES

G. Box and G. Tiao, Bayesian Inference in Statistical Analysis (Addison-Wesley Publishing Co., Reading, Massachusetts, 1973).

S. Eilen and T. P. Fowkes, "Sampling Procedures for Risk Simulation," Oper. Res. Q. 24, 241-252 (1973).

D. J. G. Farlie, "The Performance of Some Correlation Coefficients for a General Bivariate Distribution," Biometrika 47, 307-323 (1960).

G. E. Fishman, Concepts and Methods in Discrete Event Digital Simulation (John Wiley & Sons, New York, 1973).

M. Frechet, "Sur les Tableaux de Correlation dont les Marges sont Donnees," Ann. Univ. Lyon, Sect. A, Series 3, 14, 53-77 (1951).

J. C. Hall, "Dealing with Dependence in Risk Simulation," Oper. Res. Q. 29, 202-213 (1977).

M. E. Johnson, Models and Methods for Generating Dependent Random Vectors, unpublished Ph.D. dissertation, University of Iowa, Iowa City, Iowa (1976).

M. E. Johnson and J. S. Ramberg, "A Bivariate Distribution System with Specified Marginals," Los Alamos Scientific Laboratory report LA-6858-MS (1977).

N. L. Johnson and S. Kotz, Distributions in Statistics: Continuous Multivariate Distributions (John Wiley & Sons, New York, 1972).

M. G. Kendall, Rank Correlation Methods (Charles Griffin and Co., Ltd., London, 1962).

G. Kimmeldorf and A. Sampson, "One-Parameter Families of Bivariate Distributions with Fixed Marginals," Commun. Stat. 4, 293-301 (1975).

G. Kimmeldorf, "Uniform Representations of Bivariate Distributions," Commun. Stat. 4, 617-627 (1975).

C. J. Kowalski, "Non-Normal Bivariate Distributions with Normal Marginals," Am. Stat. 27, 103-106 (1973).

W. Kruskal, "Ordinal Measures of Association," J. Am. Stat. Assoc. 53, 814-859 (1958).

K. V. Mardia, "Some Contributions to Contingency-Type Bivariate Distributions," Biometrika 54, 235-249 (1967).

K. V. Mardia, Families of Bivariate Distributions (Hafner Publishing Co., Darien, Connecticut, 1970).

Math Lab Group, MACSYMA Reference Manual (Laboratory for Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1977).

D. Morgenstern, "Einfache Beispiele Zweidimensionaler Verteilungen," Mitteilungsbl. Math. Stat. 8, 234-235 (1956).

A. Nataf, "Détermination des Distributions de Probabilités dont les Marges sont Données," C. R. Acad. Sci., Paris 255, 42-43 (1962).

R. L. Plackett, "A Class of Bivariate Distributions," J. Am. Stat. Assoc. 60, 516-522 (1965).

A. Tchen, "Inequalities for Distributions with Given Marginals," Stanford University Department of Statistics, Technical Report No. 19 (1976).

A. Tenenbein and M. Gargano, "Simulation from Bivariate Distributions with Given Marginal Distribution Functions," New York University Graduate School of Business Administration Working Paper, Series #78-102 (1978).

☆U.S. Government Printing Office: 1979 —677-013/34