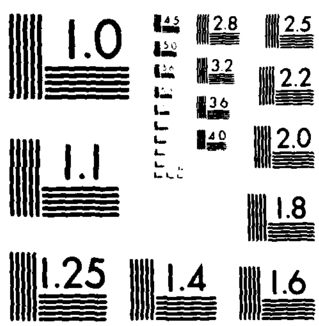
AD-A099 578 NAVAL RESEARCH LAB WASHINGTON DC F/G 12/1
GENERATION OF A MULTIVARIATE DISTRIBUTION FOR SPECIFIED UNIVARI--ETC(U)
MAY 81 I R GOODMAN
UNCLASSIFIED NRL-MR-4425 NL

END
DATE
FILMED
6-81
DTIC



MICROCOPY RESOLUTION TEST CHART

14

NRL-MR-4425

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NRL Memroandum Report 4425	2. GOVT ACCESSION NO. AD-A099578	3. RECIPIENT'S CATALOG NUMBER 9
4. TITLE (and Subtitle) 6 GENERATION OF A MULTIVARIATE DISTRIBUTION FOR SPECIFIED UNIVARIATE MARGINALS AND COVARIANCE STRUCTURE.		5. TYPE OF REPORT & PERIOD COVERED Final report on an NRL problem.
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) 10 I. R. Goodman		8. CONTRACT OR GRANT NUMBER(s) 17
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Research Laboratory Washington, DC 20375 12 28		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61153N, RR014 02 41 79-0719-0-1
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Arlington, VA 22217	11 28	12. REPORT DATE May ~~20, 1~~981
		13. NUMBER OF PAGES 27
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) 16 RR01402		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Statistics; Moments; Covariance; Sampling procedure; Probability transformation; Marginal densities; Joint distributions; Fréchet-Nataf translation method		

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This paper addresses the problem of determining and generating outcomes for a joint multivariate distribution having specified arbitrary marginal distributions and correlation structure.

In addition, a new technique is developed which generates outcomes for essentially any given multivariate random variable from a more conveniently realizable random variable. This generalizes in a natural way the well known univariate transformation of probability sampling technique.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

251950

CONTENTS

GENERATION OF A MULTIVARIATE DISTRIBUTION
FOR SPECIFIED UNIVARIATE MARGINALS AND
COVARIANCE STRUCTURE

I. INTRODUCTION

During a discussion pursuant to a proposed new nonlinear convolution filter of the Wiener type [1], the problem arose of verifying the accuracy of this filter with respect to the standard linear Wiener filter. This would require the generation of a nongaussian stochastic process, since for the gaussian case, for simple additive noise, the linear filter cannot be improved upon. (See, e.g., [5], Chapter 3.) For purposes of sensitivity analysis, it was agreed that choice of the scalar marginal random variables and covariance structure of the stochastic process should be made as wide as possible.

Consequently, the following three-stage statistical problem arose: Given a set of univariate marginal probability distributions:

(1) Determine the class of possible correlations between random variables having these distributions.

(2) Determine, for each allowable set of correlations in (1), a corresponding (joint) stochastic process yielding back the given set of marginals.

(3) Establish a procedure for generating outcomes of each stochastic process specified in (2).

Manuscript submitted October 24, 1980.

This brief memorandum addresses itself to the above problem, and under mild restrictions, answers it in the affirmative, when the set of univariate probability distributions is finite.

In treating this problem, two byproducts resulted: A simple method for the special case of determining a bivariate distribution, given two marginals; and a transform of probability technique from a multivariate uniform to any of a wide class of multivariate distributions, generalizing the well-known univariate procedure. These results are presented in Appendices A and B.

II. ANALYSIS

Following a cursory view of the available literature, Johnson and Kotz' text [2], Chapter 34 (pp. 1-36) and Chapter 42 (pp. 273 et passim) was found to present a brief survey of techniques which relate to part (2) of the problem posed here, without involving explicitly the specification of the covariance structure. Mardia's monograph [3] has a much more extensive treatment of the problem. Some of the results presented here are, in part, extensions of the Fréchet-Nataf 'translation method' for bivariate families (which do not particularly address the specification of the covariance matrices), exhibited in [3].

In the following development, let, for each index i, $i = 1, \ldots, m$, $\underset{\sim}{x}_i$ be any univariate marginal random variable with ordinary probability density function f_i, probability distribution function F_i, mean μ_i, variance $\sigma_{ii} \equiv \sigma_i^2$, median θ_i and mean absolute deviation about the median ν_i, all assumed known.

As usual, $E_F(\underset{\sim}{X})$ and $\mathrm{Cov}_F(\underset{\sim}{X})$ denote expectation and covariance matrix of random variable $\underset{\sim}{X}$ with respect to distribution F, etc.

Ψ denotes the probability distribution function corresponding to the standardized normal scaler distribution $n(0, 1)$, while ψ denotes the probability density function for $n(0, 1)$. $\Omega_m \overset{df}{=} (\rho_{ij})_{1\leq i,j\leq m}$ denotes an arbitrary (positive semidefinite) m by m correlation matrix, where $\rho_{ii} \equiv 1$ and $|\rho_{ij}| \leq 1$, $\rho_{ij} = \mathrm{corr}(\underset{\sim}{x}_i, \underset{\sim}{x}_j) = E(\underset{\sim}{x}_i - \mu_i)(x_j - \mu_j)/(\sigma_i\sigma_j)$.

Denote also $1_m \overset{df}{=} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ (m by 1), $I_m \overset{df}{=} \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$, $0_m \overset{df}{=} \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ (m by 1). $\underset{\sim}{X}_m \overset{df}{=} \begin{pmatrix} \underset{\sim}{x}_1 \\ \vdots \\ \underset{\sim}{x}_m \end{pmatrix}$, $X_m \overset{df}{=} \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$, where $x_i \in R$ is any outcome of marginal random variable $\underset{\sim}{x}_i$ (which can be arbitrary, unless specified) corresponding to f_i; $i = 1, \ldots, m$. All monotonic properties of distribution functions and positiveness of densities are with respect to an assumed interval of support.

For $\Omega_m = I_m$ and $1_m \cdot 1_m^T$, define the functions $F(\cdot;\Omega_m): R^m \to [0,1]$, where for any $X_m \varepsilon R^m$,

$$F(X_m;I_m) \stackrel{df}{=} \prod_{i=1}^{m} F_i(x_i) \quad , \tag{1}$$

$$F(X_m;1_m 1_m^T) \stackrel{df}{=} \min_{i=1,\ldots,m} F_i(x_i) \quad . \tag{2}$$

Define also $F^*_{(m)}: R^m \to [0,1]$, by, for all $X_m \varepsilon R^m$,

$$F^*_{(m)}(X_m) \stackrel{df}{=} \max\left(\sum_{i=1,\ldots,m} F_i(x_i) - (m-1),\ 0\right) . \tag{3}$$

Then the following properties hold (see [3], pp. 30-35, for the bivariate case of $m = 2$):

<u>Lemma 1</u>

(a) $F(\ ;I_m)$, $F(\ ;1_m 1_m^T)$, and $F^*_{(m)}$ are all m-variate probability distributions possessing $F_{1,\ldots,}\ F_m$ as marginal distributions.

(b) For any joint probability distribution $F_{1,\ldots,m}$, possessing marginal distributions $F_{1,\ldots,}\ F_m$ as above, for all $X_m \varepsilon R^m$,

$$F^*_{(m)}(X_m) \le F_{1,\ldots,m}(X_m) \le F(X_m;1_m,1_m^T). \tag{4}$$

(<u>Proof</u>: Omitted as a straightforward generalization of the results in [3].)

□

<u>Remarks</u>

1. Note the case of $F = F(\ ,I_m)$ for (4); $F(\ ,I_m)$ is the probability distribution function corresponding to $\chi_{1,\ldots,}\chi_m$ being statistically

independent.

2. $F(;1_m \cdot 1_m^T)$ corresponds to $\underset{\sim}{x}_1,\ldots,\underset{\sim}{x}_m$ all being monotone increasing functions of each other. This implies the unique relation

$$F_1(\underset{\sim}{x}_1) = F_2(\underset{\sim}{x}_2) = \cdots = F_m(\underset{\sim}{x}_m) \ , \tag{5}$$

i.e.,

$$\underset{\sim}{x}_i = F_i^{-1}(F\ (\underset{\sim}{x}_j)) : \qquad i = 1,\ldots,m \ ,$$

provided each F_i is a monotone increasing function. Thus, if $g:R^m \to R$ is any function,

$$\int_{X\varepsilon R^m} g(X)dF(X;1_m 1_m^T) = \int_{x_1 \varepsilon R} g\left(\begin{pmatrix} x_1 \\ F_2^{-1}F_1(x_1) \\ \vdots \\ F_n^{-1}(F_1(x_1)) \end{pmatrix}\right) dF_1(x_1) \ , \tag{6}$$

among several representations.

3. $F^*_{(m)}$ corresponds to the functional relation

$$\sum_{i=1}^{m} F_i(\underset{\sim}{x}_i) = m - 1 \ . \tag{7}$$

<u>Lemma 2</u>

For any $i \neq j$, $1 \leq i, j \leq m$, and any joint probability distribution $F_{i,j}$ possessing fixed marginal distributions F_i, F_j, and densities f_i, f_j, with $f_j > 0$

$$\alpha_{i,j} \le E_{F_{i,j}}((x_i - x_j)^2) \le \beta_{ij} \tag{8}$$

where

$$\alpha_{ij} \overset{df}{=} E_{H(\ ;1_2 1_2^T)}((\tilde{x}_i - \tilde{x}_j)^2)$$

$$= \int_{x_i \in R} (x_i - F_j^{-1}(F_i(x_i))^2 dF_i(x_i) \tag{9}$$

$$= \int_{x_j \in R} (F_i^{-1}(F_j(x_j)) - x_j)^2 dF_j(x_j) \ ;$$

the latter two relations holding, if F_i and F_j are monotone increasing;

and

$$\beta_{ij} \overset{df}{=} E_{F^*_{(2)}}((\tilde{x}_i - \tilde{x}_j)^2) \tag{10}$$

$$= \int_{x_i \in R} (x_i - F_j^{-1}(1-F_i(x_i)))^2 dF_i(x_i)$$

$$= \int_{x_i \in R} (F_i^{-1}(1-F_j(x_j)) - x_j)^2 dF_j(x_j) \quad .$$

Again, the last two equalities hold, if F_i and F_j are monotone increasing.

Upper equality in (8) holds for $F_{i,j} = F(\ ;1_2 \cdot 1_2^T)$; lower equality holds for $F_{ij} = F^*_{(2)}$.

(Proof:

Using Fubini's Theorem on iterated integration and Lemma 1 (a) for $m = 2$,

$$E_{F_{ij}}(\mathcal{X}_i-\mathcal{X}_j)^2 = \int_{x_j\varepsilon R} \left(\int_{x_i\geq x_j} (x_i-x_j)^2 dF_{(i|j)}(x_i|x_j)\right) dF_j(x_j) \ , \tag{11}$$

$$E_{F(\ ;1_2 1_2^T)}(\mathcal{X}_i-\mathcal{X}_j)^2 = \int_{x_j\varepsilon R} \left(\int_{x_i\geq x_j} (x_i-x_j)^2 dF^{(0)}_{(i|j)}(x_i|x_j)\right) dF_j(x_j) \ .$$

Now the conditional distribution functions $F_{(i|j)}(x_i|x_j) = F_{ij}(x_i,x_j)/f_j(x_j)$ and $F^{(0)}_{(i|j)}(x_i|x_j) = F(x_i,x_j;1_2 1_2^T)/f_j(x_j)$. Thus Lemma 1 (b) (eq. (4)) implies that for all $x_i,x_j\varepsilon R$, $F_{(i|j)}(x_j|x_j) \leq F^{(0)}_{(i|j)}(x_i|x_j)$. Since for x_j fixed, $(x_i-x_j)^2$ is an increasing function in x_i, it follows that (see, e.g., Lehmann [4], page 73, Lemma 1 and, page 112, Example 11) for all $x_i \geq x_j$, for any fixed x_j,

$$\begin{aligned} &\int_{x_i\geq x_j} (x_i-x_j)^2 dF_{(i|j)}(x_i|x_j) \\ &\geq \int_{x_i\geq x_j} (x_i-x_j)^2 dF^{(0)}_{(i|j)}(x_i|x_j) \ . \end{aligned} \tag{12}$$

Substituting (11) into (12) yields the lower inequality in (8).

By a similar argument, reversing inequalities from Lemma 1, yields the upper inequality in (8).)

□

The next result establishes the allowable range of values for $Corr(\mathcal{X}_i,\mathcal{X}_j)$ for given marginal densities f_i and f_j.

<u>Theorem 1</u>

Let $\mathcal{X}_i,\mathcal{X}_j$ correspond to fixed densities f_i,f_j, respectively, which are positive, for $i \neq j$; $1 \leq i,\ j \leq m$.

Then for any joint bivariate distribution F_{ij} possessing marginals f_i,f_j,

$$\gamma_{ij} \leq \mathrm{Corr}_{F_{ij}}(x_i, x_j) \leq T_{ij} \quad , \tag{13}$$

where

$$\gamma_{ij} \stackrel{\mathrm{df}}{=} \frac{\sigma_i{}^2 + \sigma_j{}^2 + (\mu_i - \mu_j)^2 - \beta_{ij}}{2\sigma_i\sigma_j} \quad , \tag{14}$$

$$T_{ij} \stackrel{\mathrm{df}}{=} \frac{\sigma_i{}^2 + \sigma_j{}^2 + (\mu_i - \mu_j)^2 - \alpha_{ij}}{2\sigma_i\sigma_j} \quad , \tag{15}$$

and α_{ij}, β_{ij} are given in Eqs. (9) and (10).

(<u>Proof</u>:

Using Lemma 2, expand out the expectation in Eq. (8), yielding a bound on $E_{F_{ij}}(x_i x_j)$. Subtracting $\mu_i \cdot \mu_j$ and dividing by $\sigma_i\sigma_j$ yields the final results.) $\square$

<u>Remarks</u>

Suppose $i \neq j$; $1 \leq i, j \leq m$.

1. Upper equality to T_{ij} in (3) is achieved for $F_{ij} = F^*_{(2)}$; lower equality to γ_{ij}, for $F_{ij} = F(\cdot, 1_2 \cdot 1_2^T)$. Thus, $-1 \leq \gamma_{ij} \leq T_{ij} \leq 1$.

2. Assuming that F_i and F_j are monotone increasing:

$$T_{ij} = 1 \text{ iff } \alpha_{ij} = (\sigma_i - \sigma_j)^2 + (\mu_i - \mu_j)^2 \quad , \tag{16}$$

$$\gamma_{ij} = -1 \text{ iff } \beta_{ij} = (\sigma_i + \sigma_j)^2 + (\mu_i - \mu_j)^2 \quad . \tag{17}$$

A sufficient condition for $T_{ij} = 1$ to hold is

$$F_i(\cdot) = F_j((\cdot - b'_{ij})/\ a'_{ij}) \ , \tag{18}$$

for any constants b'_{ij} and $a'_{ij} > 0$. (Thus, it holds for $F_i = F_j$.) This value is achieved if $x_i = a'_{ij} \cdot x_j + b'_{ij} = F_i^{-1}(F_j(x_j))$.

A sufficient condition for $\gamma_{ij} = -1$ to hold is

$$F_i(\cdot) = 1 - F_j((b''_{ij} - \cdot)/a''_{ij}) \quad , \tag{19}$$

for any constants b''_{ij} and $a''_{ij} > 0$. This value is achieved if $x_i \equiv -a''_{ij}\, x_j + b''_{ij} \equiv F_i^{-1}(1 - F_j(x_j))$.

3. In particular (corresponding to $b'_{ij} = 0$, $a'_{ij} = 1$, $b''_{ij} = 2\mu_i$, $a''_{ij} = 1$), <u>if</u> $F_i = F_j$ and is a symmetric function about $\mu_i = \mu_j$, i.e. $F_i(x_i) = 1 - F_i(2\mu_i - x_i)$, for all $x_i \varepsilon R$, then it follows that $\alpha_{ij} = 0$ (this does not require the symmetry condition), $\tau_{ij} = 1$, $\beta_{ij} = 4\sigma_i^{\ 2} = 4\sigma_j^{\ 2}$, and $\gamma_{ij} = -1$.

□

Now, denote for any m by m correlation matrix Ω_m, the m-variate zero-mean gaussian distribution $N_m(0_m, \Omega_m)$ with corresponding probability density function $\psi_{\Omega m}$ and probability distribution function $\Psi_{\Omega m}$. Define, for given monotone increasing marginal distribution functions $F_1, \dots,$ F_m, transform $T_m : R^m \rightarrow R^m$, where, for any $X_m \varepsilon R^m$,

$$T_m(X_m) \ \underline{\underline{df}} \ \begin{pmatrix} F_1^{-1}(\Psi(x_1)) \\ \vdots \\ F_m^{-1}(\Psi(X_m)) \end{pmatrix} \quad , \tag{20}$$

with the inverse transform

$$T_m^{-1}(X_m) \overset{df}{=} \begin{pmatrix} \Psi^{-1}(F_1(x_1)) \\ \Psi^{-1}(F_m(X_m)) \end{pmatrix} \quad . \tag{21}$$

Next, define the function $F(\ ;\Omega_m): R^m \to [0,1]$, where for any $X_m \varepsilon R^m$,

$$F(X_m;\Omega_m) = \Psi_{\Omega_m}(T_m^{-1}(X_m)) \quad . \tag{22}$$

(See [3], pp. 30-35, for the case m = 2.)

Finally, let $\underset{\sim}{V}_m = \begin{pmatrix} \underset{\sim}{V}_1 \\ \vdots \\ \underset{\sim}{V}_m \end{pmatrix}$ be any random variable corresponding to $N_m(0_m,\Omega_m)$. (Thus, each $\underset{\sim}{v}_i$ is distributed as $n(0,1)$.)

The major result of this is presented in the following theorem:

<u>Theorem 2</u> (Extension of Fréchet-Nataf's Translation Method [3])

Let $F_{1,\ldots,}F_m$ be m given monotone increasing scalar marginal probability distribution functions with corresponding positive densities $f_{1,\ldots,}f_m$, and random variables $\underset{\sim}{x}_{1,\ldots,}\underset{\sim}{x}_m$.

<u>Then</u>, for any correlation matrix $\Omega'_m = (\rho'_{ij})_{1\le i,j\le m}$ (which must be always at least positive semidefinite) such that

$$\gamma_{ij} \le \rho'_{ij} \le T_{ij} \quad , \tag{23}$$

where γ_{ij} and T_{ij} are given in eqs. (14) and (15), for all $i \ne j$, $1 \le i,j \le m$:

(a) There exists a corresponding correlation matrix $\Omega_m = (\rho_{ij})_{1\le i,j\le m}$ such that $F(\ ;\Omega_m)$ is a joint m-variate probability distribution of random variable $\underset{\sim}{X}_m \overset{df}{=} \begin{pmatrix} \underset{\sim}{X}_1 \\ \vdots \\ \underset{\sim}{X}_m \end{pmatrix}$, which, possesses marginal probability distribution functions, densities and random variables $F_1,\ldots,F_m, f_1,\ldots,f_m$, and $\underset{\sim}{X}_1,\ldots,\underset{\sim}{X}_m$, respectively, such that

$$\underset{F(\cdot;\Omega_m)}{\mathrm{Corr}}\,(X_m) = \Omega'_m \quad . \tag{24}$$

(b) Random variable $\underset{\sim}{X}_m$ can be considered in the following functional relationship with respect to $\underset{\sim}{V}_m$:

$$\underset{\sim}{X}_m = T_m(\underset{\sim}{V}_m) \quad . \tag{25}$$

(c) For $\Omega'_m = (T_{ij})_{1\le i,j\le m}(T_{ii} \equiv 1)$, the upper bound in eq. (23), we can choose $\Omega_m = 1_m \cdot 1_m^T$ and $F(\cdot;1_m 1_m^T)$ is the same as previously defined in Eq. (2).

(d) For $\Omega'_m = I_m$, we can choose $\Omega_m = I_m$ and $F(\cdot;I_m)$ is the same as previously defined in Eq. (1).

(e) For Ω'_m with at least some entires $\rho'_{ij} = \gamma_{ij}$, the lower bounds in (23), we can choose corespondingly in Ω_m, $\rho_{ij} = -1$, which yields the bivariate $F_{ij} = F(\;;\binom{1}{-1}\cdot\binom{1}{-1}^T) = F^*_{(2)}$, as given in Eq. (3).

<u>Proofs</u>:

(b) follows from (22). (c), (d), (e) all follow from basic properties of m-variate gaussian distributions.

For (a): Without loss of generality, consider for any Ω_m, the first marginal component of $F(X_m;\Omega_m)$:

$$(F(X_m;\Omega_m))_1 = F\left(\begin{pmatrix} x_1 \\ +\infty \\ \vdots \\ +\infty \end{pmatrix};\Omega_m\right) = \Psi_{\Omega_m}\left(\begin{pmatrix} \Psi^{-1}(F_1(x_1)) \\ \Psi^{-1}(1) \\ \vdots \\ \Psi^{-1}(1) \end{pmatrix}\right)$$

$$= \Psi_{\Omega_m}\left(\begin{pmatrix} \Psi^{-1}(F_1(x_1)) \\ +\infty \\ \vdots \\ +\infty \end{pmatrix}\right) = \Psi(\Psi^{-1}(F_1(x_1))) = F_1(x_1).$$

Thus, $\{F(\ ;\Omega_m)|\Omega_m$ is arb. pos. semidefinite$\}$ is a family of m-variate distributions of $\underset{\sim}{X}_m$ yielding back the required marginals. Clearly, for any fixed $i \neq j$, $\mathrm{Corr}_{(F(\ ;\Omega_m)}(\underset{\sim}{X}_i, \underset{\sim}{X}_j)$ is indeed a continuous function of Ω_m - in fact, of ρ_{ij} - which achieves its maximum value (see Theorem 1 and ensuing remarks) T_{ij}, for $\Omega_m = 1_m 1_m^T$ ($\rho_{ij} = 1$), and its minimal value γ_{ij}, for Ω_m such that $\rho_{ij} = -1$, i.e., for bivariate marginal $F_{ij} = F^*_{(2)} = F(\ ;\binom{1}{-1}\binom{1}{-1}^T)$. Thus, $\mathrm{Corr}_{F(\ ;\Omega_m)}(\underset{\sim}{X}_i, \underset{\sim}{X}_j)$ can take any value between γ_{ij} and T_{ij} by proper choice of Ω_m.)

<u>Remarks</u>

1. If $F_1 = F_2 = \cdots = F_m$ is a symmetric function about mean μ_1 ($F_1(x_1) \equiv 1 - F_1(2\mu_1 - x_1)$), then $\gamma_{ij} = -1$ and $T_{ij} = 1$, and thus, Ω'_m in Theorem 2 can be arbitrary (and (23) is always satisfied).

2. Rewriting q. (24) in scalar form

$$\rho'_{ij} = \underset{F(\ ;\Omega_m)}{\mathrm{Corr}}(x_i, x_j) \equiv \underset{F_{ij}\left(\ ;\begin{pmatrix} 1 & \rho_{ij} \\ \rho_{ij} & 1 \end{pmatrix}\right)}{\mathrm{Corr}}(x_i, x_j) \quad , \tag{26}$$

for any given $\gamma_{ij} \leq \rho'_{ij} \leq T_{ij}$, where F_{ij} is the corresponding bivariate marginal distribution function of F, and $\Omega_m = (\rho_{ij})_{1 \leq i,j \leq m}$.

Lancaster and Kendall & Stuart (see [3], page 33) have shown that for all $1 \leq i, j \leq m$,

$$|\rho'_{ij}| \leq |\rho_{ij}| \tag{27}$$

with strict inequality holding in (27) for $i \neq j$, unless the F_i's are all gaussian distributions.

For any given ρ'_{ij}, with strict inequalities both holding in (23), solving for ρ_{ij} in (26) is equivalent to solving for ρ_{ij} in:

$$\rho'_{ij} = H_{ij}(\rho_{ij}) \tag{28}$$

$$\stackrel{df}{=} (1/(\sigma_i \cdot \sigma_j)) \cdot \iint_{\binom{v_i}{v_j} \in R^2} (F_i^{-1}(\Psi(v_i)) - \mu_i)(F_j^{-1}(\Psi(v_j)) - \mu_j) \cdot \Psi_{\begin{pmatrix} 1 & \rho_{ij} \\ \rho_{ij} & 1 \end{pmatrix}}(v_i, v_j) dv_i dv_j$$

where

$$\Psi_{\begin{pmatrix} 1 & \rho_{ij} \\ \rho_{ij} & 1 \end{pmatrix}}(v_i, v_j) = (1/(2\pi \cdot \sqrt{1-\rho_{ij}^2})) \cdot \exp-\{ \tfrac{1}{2}(1-\rho_{ij}^2)^{-1} \cdot (v_i^2 + v_j^2 - 2\rho_{ij} v_i v_j) \} \quad . \tag{29}$$

Thus, for purposes of implementation, the integral on the right hand side of (28), $H_{ij}(\rho_{ij})$ should be tabulated as a function of ρ_{ij}, for $-1 < \rho_{ij} < 1$. Then that ρ_{ij} satisfying (28) could be found by a graphical method.

3. For $m=2$, an alternative simple method for constructing the joint bivariate probability distribution for specified marginals F_1 and F_2 is presented in Appendix A.

4. For general m again, Eq. (25) can be used to generate outcomes of $\underset{\sim}{X}_m$. This thus depends on the generation of outcomes of $\underset{\sim}{V}_m$. A standard procedure, is to first in effect decompose $\Omega_m = P_m D_m P_m^T$, where P_m is the orthonormal matrix of eigenvectors and D_m is the diagonal matrix of eigenvalues of Ω_m. Then the random variable $\underset{\sim}{V}_m$ can be written as

$$\underset{\sim}{V}_m = P_m \underset{\sim}{W}_m, \text{ where random variable } \underset{\sim}{W}_m = \begin{pmatrix} \underset{\sim}{w}_1 \\ \vdots \\ \underset{\sim}{w}_m \end{pmatrix} \text{ is distributed } N_m(0, D_m).$$

Hence, the $\underset{\sim}{w}_i$'s are distributed statistically independent with $\underset{\sim}{w}_i$ distributed $n(0, d_i^2)$ where

$$D_m \overset{\text{df}}{=} \begin{pmatrix} d_1^2 & & 0 \\ & \ddots & \\ 0 & & d_m^2 \end{pmatrix}.$$

Finally, we can apply the fundamental transformation of probability from a uniformly distributed random variable $\underset{\sim}{z}_i$ over $[0,1]$: $\underset{\sim}{w}_i = \Psi^{-1}(\underset{\sim}{z}_i)$; $i=1,..,m$. Hence we can first generate statistically independently outcomes $z_1,\ldots,z_m$, obtaining in turn, outcomes $w_1,\ldots,w_m$, forming $\underset{\sim}{W}_m$, and in turn finally, $\underset{\sim}{V}_m$.

5. An alternative procedure for not only generating $\underset{\sim}{V}_m$, but for obtaining essentially any m-variate random variable by a transform of probability from any convenient initial m-variate random variable - such as a uniformly distributed one over the unit m-cube - is given in Appendix B. This technique could be of especial use when non gaussian random variables are sought and/or when m is large, since no matrix, inversion is required. However, the forms of the marginal probability distribution functions are required.

III. SUMMARY

The problem of constructing and realizing outcomes for a multivariate distribution compatible with arbitrary given scalar marginal distributions and having a specified correlation matrix has been considered here. The problem is essentially answered in the affirmative in Theorem 2. In addition, an apparently new multivariate transformation of probability technique which is triangular in form is presented (Appendix B) for generating outcomes of a given random variable from a more convenient distribution.

IV. REFERENCES

1. Personal Communication with Dr. Abraham Schultz (Code 7932) of the Systems Research Branch, Space Systems Division of Naval Research Laboratory, Wash., D. C. An internal memorandum on the proposed filter will be forthcoming (1980).

2. N. L. Johnson and S. Kotz, Distributions in Statistics: Continuous Multivariate Distributions, John Wiley & Sons, Inc., 1972.

3. K. V. Mardia, Families of Bivariate Distributions, Griffin Statistical Monograph No. 27, Hafner Co., 1970.

4. E. L. Lehmann, Testing Statistical Hypotheses, John Wiley & Sons, Inc. 1959.

5. C. R. Rao, Linear Statistical Inference and its Applications, 2nd Ed., John Wiley & Sons, Inc., 1973.

APPENDIX A

Alternative Construction of a Bivariate Distribution for Specified Marginal Distributions and Covariance Matrix

Let $\underset{\sim}{x}_1$, $\underset{\sim}{x}_2$ be two given marginal scalar random variables with probability density functions f_1, f_2 continuous about their medians θ_1, θ_2 possessing (finite) variances σ_1^2, σ_2^2 ; denote their absolute first moments about their medians by μ_1' , μ_2' , respectively.

Let $\sum = (\sigma_{ij})_{1 \leq i,j \leq 2}$ be any given positive semidefinite matrix, such that

$$\sigma_{ii} = \sigma_i^2 \quad ; \quad i=1,\ 2 \tag{A-1}$$

and

$$|\sigma_{12}| < \mathrm{Min}(\mu_1'^2,\ \mu_2'^2) \quad . \tag{A-2}$$

Define functions f and h_i by, for all outcomes

$X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \varepsilon\ R^2$,

$$f(X) \stackrel{df}{=} \prod_{i=1}^{2} f_i(x_i) + \prod_{i=1}^{2} h_i(x_i) \ , \tag{A-3}$$

$$h_i(x_i) \stackrel{df}{=} \lambda_i \begin{cases} f_i(x_i) & , \text{ for all } x_i \geq \theta_i \\ -f_i(x_i) & , \text{ for all } x_i < \theta_i \end{cases} \tag{A-4}$$

and

$$\lambda_1 = \frac{\sqrt{\sigma_{12}}}{\mu_1'} \quad , \quad \lambda_2 = \frac{\sqrt{\sigma_{12}}}{\mu_2'} \quad ; \quad \text{if} \quad \sigma_{12} \geq 0 \quad ,$$

$$\lambda_1 = \frac{-\sqrt{|\sigma_{12}|}}{\mu_1'} \quad , \quad \lambda_2 = \frac{\sqrt{|\sigma_{12}|}}{\mu_2'} \quad ; \quad \text{if} \quad \sigma_{12} < 0 \quad . \tag{A-5}$$

Then f is a joint probability density function of $\underset{\sim}{X} \stackrel{df}{=} \begin{pmatrix} \underset{\sim}{x}_1 \\ \underset{\sim}{x}_2 \end{pmatrix}$,

having marginal p.d.f.'s f_1, f_2 and such that with respect to f,

$$E(\underset{\sim}{X}) = \begin{pmatrix} E(\underset{\sim}{x}_1) \\ E(\underset{\sim}{x}_2) \end{pmatrix} \qquad \text{and} \qquad Cov(\underset{\sim}{X}) = \Sigma \quad . \tag{A-6}$$

Proof:

Since (A-2) and (A-5) imply $|\lambda_i| \leq 1$, then (A-3) and (A-4) imply that $f(X) \geq 0$, for all $X \varepsilon R^2$. Clearly, by construction in (A-4)

$$\int_{x_i=-\infty}^{+\infty} h_i(x_i)dx_i = 0; \qquad i=1,2 \tag{A-7}$$

implying

$$\int_{X \,\varepsilon\, R^m} f(X)dX = \int_{X \varepsilon\, R^2} \prod_{i=1}^{2} f_i(x_i)dx_i = 1$$

and f is thus a p.d.f.

Also, for example, from (A-7),

$$\begin{aligned} \int_{x_2=-\infty}^{+\infty} f(X)dX_2 &= f_1(x_1) + 0 \\ &= f_1(x_1) , \end{aligned}$$

verifying that the marginal p.d.f.'s of f are f_1 and f_2.

Now

$$E(\underset{\sim}{X}) = \int_{X\varepsilon R^2} X f(X)\,dX = \left(\int_{X\varepsilon R^m} x_i f(X)\,dX\right)_{1\le i\le 2} ,$$

where, for example, using (A-7) again

$$\int_{X\varepsilon R^2} x_1 \; f(X)\,dX = \int_{x_1=-\infty}^{+\infty} x_1 f_1(x_1)\,dx_1 +$$

$$\int_{x_1=-\infty}^{+\infty} x_1 h_1(x_1)\,dx_1 \cdot \int_{x_2=-\infty}^{+\infty} h_2(x_2)\,dx_2$$

$$= E(\underset{\sim}{x}_1).$$

Finally,

$$\mathrm{Cov}(\underset{\sim}{X}) = \left(\int_{X\varepsilon R^2} (x_i - E(\underset{\sim}{x}_i))(x_j - E(\underset{\sim}{x}_j)) f(X)\,dX\right)_{1\le i,j\le 2} ,$$

For i=j, using (A-7)

$$\mathrm{var}(\underset{\sim}{x}_i) = \int_{X\varepsilon R^2} (x_i - E(\underset{\sim}{x}_i))^2 \cdot f(X)\,d(X)$$

$$= \int_{x_i=-\infty}^{+\infty} (x_i - E(\underset{\sim}{x}_i))^2 f_i(x_i)\,dx_i) \cdot \prod_{\substack{1\le j\le 2,\\ j\ne i}} \int_{x_j=-\infty}^{+\infty} f_j(x_j)\,dX,$$

$$+ \int_{x_i=-\infty}^{+\infty} (x_i - E(\underset{\sim}{x}_i))^2 \cdot h_i(x_i)\,dx_i) \cdot \prod_{\substack{1\le j\le 2,\\ j\ne 1}} \int_{x_j=-\infty}^{+\infty} h_i \; (x_j)\,dx_j$$

$$= \int_{x_i=-\infty}^{+\infty} (x_i - E(\underset{\sim}{x}_i))^2 f_i(x_i)\,dx_i$$

$$= \sigma_i^{\,2} ,$$

for $i = 1,2$.

For $i \neq j$, i.e., $i = 1$, $j = 2$,

$$\mathrm{Cov}(\underset{\sim}{x}_1,\underset{\sim}{x}_2) = \int_{X\epsilon\, R^m} (x_1 - E(\underset{\sim}{x}_1)(x_2 - E(\underset{\sim}{x}_2))f(X)\,dX$$

$$= \int_{x_1=-\infty}^{+\infty} (x_1 - E(\underset{\sim}{x}_1))f_1(x_1)\,dx_1) \; (\int_{x_2=-\infty}^{+\infty} (x_2 - E(\underset{\sim}{x}_2))f_2(x_2)dx_2)$$

$$+$$

$$\int_{x_1=-\infty}^{+\infty} (x_1 - E(\underset{\sim}{x}_1))h_1(x_1)\,dx_1)\cdot(\int_{x_2=-\infty}^{+\infty} (x_2 - E(\underset{\sim}{x}_2))h_2(x_2)\,dx_2) .$$

$$= 0 \cdot 0 +$$

$$\int_{x_1=-\infty}^{+\infty} (x_1 - \theta_1)h_1(x_1)\,dx_1)(\int_{x_2=-\infty}^{+\infty} (x_2 - \theta_2)h_2(x_2)\,dx_2) \quad ,$$

using (A-7)

$$= \lambda_1 \cdot \mu_1' \cdot \lambda_2 \cdot \mu_2'$$

$$= \sigma_{12} \qquad .$$

APPENDIX B

Multivariate Transformation of Probability Technique for Generating Outcomes of Arbitrary Random Variables

The univariate transform $Y = F(X)$ of a random variable X with continuous cumulative probability distribution F causes Y to be uniformly distributed over $[0,1]$. This widely known result is perhaps the most useful in simulating outcomes of one dimensional random variables. (See, e.g., the basic texts of Fisz, Wilks, Cramér, Feller, etc.) However, the multivariate extension of this transform is another matter and this author, to the best of his knowledge, has not seen such a result.

This appendix displays a natural extension of the fundamental transform of probability theory, which hopefully will be of help in simulating outcomes of arbitrary multivariate random variables.

Let $\underset{\sim}{X} = \begin{pmatrix} \underset{\sim}{x}_1 \\ \vdots \\ \underset{\sim}{x}_m \end{pmatrix}$ be a given random vector over R^m with positive probability density function f. Define transform $T_f\colon R^m \rightarrow R^m$ by, for any outcome $X = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \in R^m$,

$$T_f(X) \overset{\text{df}}{=} \begin{pmatrix} T_{f,1}(X) \\ \vdots \\ T_{f,m}(X) \end{pmatrix} , \tag{B-1}$$

where

$$T_{f,i}(X) \overset{\text{df}}{=} (1/f_{i+1}(X_{(i+1)}))\cdot\int_{s_i=-\infty}^{x_i} f_i\left(\left(\frac{s_i}{X_{(i+1)}}\right)\right)ds_i , \tag{B-2}$$

for $i = 1, \ldots, m$, and, for $j = 1, \ldots, m$,

$$X_{(j)} \stackrel{df}{=} \begin{pmatrix} x_j \\ x_{j+1} \\ \vdots \\ x_m \end{pmatrix} , \tag{B-3}$$

and f_j is the probability density function of marginal random vector $\underset{\sim}{X}_{(j)}$ where also

$$X_{(m+1)} \stackrel{df}{=} \phi \text{ and } f_{m+1}(X_{(m+1)}) \stackrel{df}{=} 1,$$ and note the special case $X_{(1)} = X$, $f_1 = f$.

Then, the random vector $\underset{\sim}{Y} = T_f(\underset{\sim}{X})$ is uniformly distributed over the m-cube $[0,1]^m$.

Proof:

First note that

$$T_{f,i}(X) < (1/f_{i+1}(X_{(i+1)}))\cdot \int_{s_i=-\infty}^{+\infty} f_i\left(\begin{matrix} s_i \\ X_{(i+1)} \end{matrix}\right) ds_i$$

$$= (1/f_{i+1}(X_{(i+1)}))\cdot f_{i+1}(X_{(i+1)})$$

$$= 1 \quad ,$$

for $i = 1, \ldots, m$, implying, for all $X \varepsilon R^m$,

$$T_f(X) \;\varepsilon\; [0,1]^m \quad . \tag{B-4}$$

Next, it easily follows that

$$\frac{dT_f(X)}{dX} = \left(\frac{\partial T_{f,i}(X)}{\partial x_j}\right)_{1 \le i, j \le m} \tag{B-5}$$

is an upper triangular matrix, i.e.,

$$\frac{\partial T_{f,i}(X)}{\partial x_j} = 0 \qquad \text{(B-6)}$$

if $j<i$; $1\le i,j\le m$, with (2) implying

$$\frac{\partial T_{f,i}(X)}{\partial x_i} = f_i(X_{(i)}) \Big/ f_{i+1}(X_{(i+1)}) , \qquad \text{(B-7)}$$

for $i=1,..,m$.

Hence, (B-5) - (B-7) implies

$$\left|\det\left(\frac{dT_f(X)}{dX}\right)\right| = \left|\prod_{i=1}^{m} \frac{\partial T_{f,i}(X)}{\partial x_i}\right| \qquad \text{(B-8)}$$

$$= f(X) \quad .$$

The positivity of f implies that from (B-2), for $X_{(i+1)}$ fixed, $T_{f,i}(X)$ as a function of x_i only, is monotone increasing for $i=1,..,m$, implying that T is a one-to-one function. Indeed, by defining, for any fixed $X_{(i+1)}$, the function $F_{i;X_{(i+1)}}$, where for any $x \in R$,

$$F_{i;X_{(i+1)}}(x) \ \underline{\underline{df}} \ \int_{s=-\infty}^{x} f_i\left(\left(\frac{s}{X_{(i+1)}}\right)\right) ds \quad , \qquad \text{(B-9)}$$

for $i=1,..,m$, with the special case

$$F_{m;X_{(m+1)}}(x) = F_m(x) = \int_{s=-\infty}^{x} f_m(s)\, ds,$$

and denoting the functional inverse by $F^{-1}_{i;X_{(i+1)}}$, then it follows that the inverse transform $T_f^{-1}: R^m \to R^m$ is obtained, for any outcome $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$, by successive substitution:

$$X \equiv T^{-1}(Y) \stackrel{df}{=} \begin{pmatrix} T_{f,1}^{-1}(Y) \\ \vdots \\ T_{f,m}^{-1}(Y) \end{pmatrix} , \qquad \text{(B-10)}$$

where using the $X_{(i)}$ notation

$$X_m \equiv T_{f,m}^{-1}(Y) \stackrel{df}{=} F_m^{-1}(y_m) ,$$

$$X_{m-1} \equiv T_{f,m-1}^{-1}(Y) \stackrel{df}{=} F^{-1}{}_{m-1;X_m}(y_{m-1}) ,$$

$$X_{m-2} \equiv T_{f,m-2}^{-1}(Y) = F^{-1}{}_{m-2;X_{(m-1)}}(y_{m-2}) ,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots ,$$

$$X_2 \equiv T_{f,2}^{-1}(Y) = F_{2;X_{(3)}}^{-1}(y_2) ,$$

$$X_1 \equiv T_{f,1}^{-1}(Y) = F_{1;X_{(2)}}^{-1}(y_1) . \qquad \text{(B-11)}$$

Combining the one-to-one property of T_f, with (B-4) (B-8) and the matrix relation, for all $Y \in [0,1]^m$,

$$\frac{dT_f^{-1}(Y)}{dY} = \left(\frac{dT_f(X)}{dX}\right)^{-1}_{X = T_f^{-1}(Y)} \qquad \text{(B-12)}$$

yields, for the p.d.f. g of $\underset{\sim}{Y}$ at outcome Y, by the standard transformation of probability,

$$g(Y) = f(T_f^{-1}(Y)) \cdot \left|\det\left(\frac{dT_f^{-1}(Y)}{dY}\right)\right|$$

$$= f(T_f^{-1}(Y))/f(T_f^{-1}(Y))$$

$$= 1 \qquad \bullet$$

Remarks

1. It immediately follows that $\underset{\sim}{Y}$ is uniformly distributed over $[0,1]^m$, <u>iff</u> each scalar marginal random variable component of $\underset{\sim}{Y}$ is statistically independent, identically distributed uniformly over $[0,1]$.
2. Let $\underset{\sim}{Y}$ be a random vector uniformly distributed over $[0,1]^m$. Let f be a given poisitive p.d.f. over R^m. Then the random vector $\underset{\sim}{X} \overset{df}{=} T_f^{-1}(\underset{\sim}{Y})$ has p.d.f. f over R^m.
3. Let $\underset{\sim}{X}$ be a given random vector over R^m with positive p.d.f. f, and let h be any given positive p.d.f. over R^m. Then the random vector $\underset{\sim}{Z} \overset{df}{=} T_h^{-1}(T_f(\underset{\sim}{X}))$ has p.d.f. h over R^m.

$1 \leq i, j \leq m,$

$$|\rho'_{ij}| \leq |\rho_{ij}| \tag{27}$$