BELLCOMM, INC.
955 L'ENFANT PLAZA NORTH, S.W., WASHINGTON, D.C. 20024

COVER SHEET FOR TECHNICAL MEMORANDUM

TITLE- An Alternative to Pseudo-Random Number Generators

TM- 69-2014-5

DATE- April 16, 1969

FILING CASE NO(S)- 310

AUTHOR(S)- H. J. Bixhorn

FILING SUBJECT(S)
(ASSIGNED BY AUTHOR(S))-

ABSTRACT

Two methods are given for generating sets of vectors with properties which approximate those of a multivariate normal distribution.

The first method, which applies to any number of dimensions, forces the univariate marginal distributions of a set of points to approximate N(0,1) with low correlations between the components.

The second method applies to 2-dimensional normal random vectors with identity covariance matrix: it attempts to distribute the points in circles of constant density.

These methods are being considered as replacements for pseudo-random number generators in certain types of Monte Carlo problems.

SEE REVERSE SIDE FOR DISTRIBUTION LIST

BELLCOMM, INC.
955 L'ENFANT PLAZA NORTH, S.W. WASHINGTON, D.C. 20024

SUBJECT: An Alternative to Pseudo-Random Number Generators - Case 310

DATE: April 16, 1969

FROM: H. J. Bixhorn

TM-69-2014-5

TECHNICAL MEMORANDUM

1.0 INTRODUCTION

Various methods for generating pseudo-random samples from uniform and normal populations by computer techniques are in common use. Numerous studies have been made of these methods, and standard tests such as those for goodness-of-fit, independence, and length-of-run are generally performed on the numbers generated by them. (Chambers[1] and Hull and Dobell[2] give excellent surveys with extensive bibliographies.) The results vary, and as is to be expected, some properties are satisfied much better than others.

Use is made of pseudo-random numbers in Monte Carlo simulations, where they are taken as input variables and propagated through a series of processes in order to study the distribution of the resulting quantities. If it were possible to determine this resulting distribution by operating on the input normal random variables analytically, there would be no need for Monte Carlo. The difficulties in working analytically with a theoretical distribution are bypassed in Monte Carlo simulations by producing a finite number of points which, hopefully, represents the distribution. In many studies which now make use of pseudo-random numbers, it would therefore be more appropriate to use a set of numbers which had been forced to resemble the parent population as much as possible. The notion of replacing pseudo-random numbers by a sequence designed to meet the needs of a specific problem has been discussed by several authors. Hull and Dobell state: "We must expect that the very best sequences for a particular purpose may be so carefully tailored to that purpose, that they are no longer random." Approaches taken by Kahn[3] and Hammersley and Handscomb[4] in producing such sequences are given in section 2.0.

A finite number of points can obviously not satisfy all the properties of a normal population. It is possible, however, to generate points which are forced to meet certain properties and to hope that other properties will be more or less satisfied in the process.

In Section 2.0 a method is given for constructing a set of N k-dimensional vectors with univariate marginals which closely approximate N(0,1) and with low correlations between all pairs of components. It is too often assumed by persons using random number generators that if these properties are satisfied, the N points will behave like the multivariate normal distribution $N(\vec{0}^{kx1}, I^{kxk})$.* This is not necessarily the case. Feller[5] (page 99) gives two examples of bivariate distributions which are not normal, but whose marginals are N(0,1). Using his second example we can construct a bivariate distribution which is not $N(\vec{0}^{2x1}, I^{2x2})$, but which has covariance matrix I^{2x2} and marginals N(0,1). The distribution is given by

$$\frac{1}{2}\left[N\left(\vec{0}^{2x1}, \Sigma_1^{2x2}\right) + N\left(\vec{0}^{2x1}, \Sigma_2^{2x2}\right)\right]$$

where

$$\Sigma_1 = \begin{bmatrix} 1 & c \\ c & 1 \end{bmatrix} \quad , \quad \Sigma_2 = \begin{bmatrix} 1 & -c \\ -c & 1 \end{bmatrix}$$

and c satisfies the conditions $c \neq 0$, $-1 < c < 1$.

The N points generated by the method of section 2.0 are therefore examined to see how well they satisfy certain properties of $N(\vec{0}^{kx1}, I^{kxk})$. In section 3.0, vectors are constructed to satisfy the property that $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ has circles of constant density when $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ is distributed as $N(\vec{0}^{2x1}, I^{2x2})$. The marginals are then tested against N(0,1). This particular property was chosen to make the marginals relatively insensitive to orthogonal transformations.

---

*$N\left(\vec{\mu}^{kx1}, \Sigma^{kxk}\right)$ denotes a k-dimensional distribution with mean vector $\vec{\mu}^{kx1}$ and covariance matrix Σ^{kxk}. A random vector $\vec{X}^{kx1} = [X_1,\ldots,X_k]^T$ is said to be distributed as $N(\vec{\mu},\Sigma)$ if $X_1,\ldots,X_k$ have the joint probability density

$$(2\pi)^{-k/2}|\Sigma|^{-1/2} \exp\left[-\tfrac{1}{2}(\vec{x} - \vec{\mu})^T \Sigma^{-1}(\vec{x} - \vec{\mu})\right] \quad .$$

Sets of numbers which are forced to obey certain properties of a given distribution are referred to by Hammersley and Handscomb as "quasi-random". Kahn uses the expression "systematic sampling" for the same purpose. In this study such numbers are said to represent or approximate the parent population to emphasize that the points are used as a substitute for the theoretical population, and that we have no interest in random sampling. For the same reason, tests which depend on the order in which the sample is drawn and which are usually performed on pseudo-random numbers are not performed here. They have relevance to random sampling but not to this study.

2.0 METHOD OF FORCED MARGINALS

A method often used for generating pseudo-random numbers from N(0,1) consists of solving for x_i in the equation $F(x_i) = y_i (i=1,\ldots,N)$ where F(x) is the standard normal distribution function and $\{y_1,\ldots y_N\}$ are pseudo-random numbers from a uniform distribution over the interval (0,1). An analogous technique can be used to force a set of N points to have a distribution closely approximating N(0,1). In place of the pseudo-random numbers $\{y_1,\ldots y_N\}$, the points $\frac{i-1/2}{N}$ or $\frac{i}{N+1}$ $(i=1,\ldots,N)$ may be equated to F(x). Because these points have an empirical cumulative distribution function closely resembling the uniform over (0,1), the empirical cumulative distribution of the corresponding points $\{x_1,\ldots x_N\}$ will closely approximate the standard normal.

To produce points representing the k-dimensional normal population $N(\vec{0}^{k\times1}, I^{k\times k})$, the vectors

$$\begin{bmatrix} x_1 \\ x_{i_{11}} \\ x_{i_{21}} \\ \cdot \\ \cdot \\ \cdot \\ x_{i_{k\ 1,1}} \end{bmatrix} \begin{bmatrix} x_2 \\ x_{i_{12}} \\ x_{i_{22}} \\ \cdot \\ \cdot \\ \cdot \\ x_{i_{k-1,2}} \end{bmatrix} \quad --- \quad \begin{bmatrix} x_N \\ x_{i_{1N}} \\ x_{i_{2N}} \\ \cdot \\ \cdot \\ \cdot \\ x_{i_{k-1,N}} \end{bmatrix}$$

may be taken where $\{x_{i_{j1}}, x_{i_{j2}},\ldots,x_{i_{jN}}\}$ $(j=1,2,\ldots k-1)$ is the jth random permutation of the sequence $\{x_1,\ldots,x_N\}$. Here the univariate marginals are forced to approximate N(0,1) and the random permutations give some semblance of independence between the components. Kahn discusses this method, which he refers to as "systematic sampling".

A method discussed by Hammersley and Handscomb makes it possible to arrange for independence between the components rather than to leave this to the chance of a random permutation. This method, due to van der Corput[6] and expanded by Halton[7], produces a set of vectors with marginals which are approximately independent and uniformly distributed, as follows:

Let the positive integers be expressed in a system of base R, i.e.,

$$n = a_m R^m + a_{m-1} R^{m-1} + \dots + a_1 R + a_o$$

($o < a_i < R$). Writing the digits of these numbers in reverse order, preceded by a point gives

$$\phi_R(n) = a_o R^{-1} + a_1 R^{-2} + \dots + a_m R^{-m-1} \quad .$$

For example, if R = 2,

decimal	R = 2	R = 2	decimal
n = 1	1	$\phi_2(n)$ = .1	.5
2	10	.01	.25
3	11	.11	.75
4	100	.001	.125
5	101	.101	.625

R = 3

decimal	R = 3	R = 3	decimal
n = 1	1	$\phi_3(n)$ = .1	.33...
2	2	.2	.66...
3	10	.01	.11...
4	11	.11	.44...
5	12	.21	.77...

It is clear that the function $\phi_R(n)$ distributes points uniformly over the interval (0,1). If R_1, $R_2,\ldots,R_k$ are relatively prime, the sequence of vectors $\left[\phi_{R_1}(n), \phi_{R_2}(n),\ldots, \phi_{R_k}(n)\right]^T$ $(n=1,\ldots,N)$ will have some semblance of independence between pairs of components for sufficiently large N. For small N, $\phi_{R_i}(n)$ and $\phi_{R_j}(n)$ $(n=1,\ldots,N)$ are linearly correlated. There is an exact linear relationship between the two for $N=R_{i-1}$ $(R_i<R_j)$, but this dependence fades as N becomes larger. Figure 1 shows this for the case of $R_i=5$, $R_j=7$. For this case the points $[x_5(n), x_7(n)]$ lie on a straight line for n = 1,2,3,4. For n>4, the points no longer lie on the same straight line. Larger values of R_i and R_j require higher values of N to drive the correlation down.

It should be noted that the distribution of points $\phi_{R_i}(n)$ $(n=1,\ldots N)$ will be biased to the left in (0,1) unless $N=R^p-1$ where p is any positive integer. In figure 2 the case of $R_i=3$ is shown for a few points. The distribution is clearly biased except for N=2 and N=8.

Choosing a value of N equal to R_i^p-1 eliminates the bias of $\phi_{R_i}(n)$ $(n=1,\ldots,N)$ but causes other difficulties. For large values of R_i only large, widely spread values of N would be acceptable, and then $\phi_{R_j}(n)$ $(j\neq i)$ would be biased. If the value of N (the number of simulations to be run) is known in advance, ways can be found to reduce or eliminate the bias. Studies are being continued on this problem. Knowing N in advance would also allow us to make other improvements in the approximation to the multivariate uniform distribution. These topics will be examined in a future study.

In this report it is assumed that N is not known in advance, and therefore the methods and results given are for the most general case.

The vectors $\left[\phi_{R_1}(n),\ldots, \phi_{R_k}(n)\right]^T$ $(n=1,\ldots,N)$ representing a multivariate uniform distribution are easily transformed into vectors representing $N(\vec{0}^{k\times 1}, I^{k\times k})$. If F(x) is the standard normal distribution function, solution of the equation $F(x_i(n)) = \phi_{R_i}(n)$ for $x_i(n)$ and formation of the vectors

$$(1) \quad \begin{bmatrix} x_1(1) \\ x_2(1) \\ \cdot \\ \cdot \\ \cdot \\ x_k(1) \end{bmatrix} \begin{bmatrix} x_1(2) \\ x_2(2) \\ \cdot \\ \cdot \\ \cdot \\ x_k(2) \end{bmatrix} \quad --- \quad \begin{bmatrix} x_1(N) \\ x_2(N) \\ \cdot \\ \cdot \\ \cdot \\ x_k(N) \end{bmatrix}$$

gives the desired set.

Set (1) noted above was constructed for k=6, N=100 and for k=20: N=500; 1000. R_i (i=1,...,k) was taken to be the first k prime numbers.

The empirical marginal distributions, the means of the marginals, and the correlation coefficients of certain pairs of components of the vectors in these sets were tested against $N(\vec{0}^{k\times 1}, I^{k\times k})$. The empirical distributions of the sums of certain components and of the sums of squares of these components were also tested against their hypothetical distributions. It seemed reasonable to demand that the vectors of set (1) give good fits to their hypothetical distributions under these simple transformations if they are to be used in simulations involving more complicated operations. Some of these tests might also be considered as goodness-of-fit tests of the N k-dimensional points against $N(\vec{0}^{k\times 1}, I^{k\times k})$. The test on the sum of squares, for example, may be thought of as comparing the observed to expected number of points in hyperspheres of successively larger radii. A goodness-of-fit test using subintervals of the k-dimensional space would be useless because of the relatively small values of N and large values of k.

A description of the tests and the results are discussed in Sections 4.0 and 4.1.

3.0 <u>METHOD OF FORCED CIRCLES</u>

If $\vec{X}^{k\times 1}$ is distributed as $N(\vec{0}^{k\times 1}, I^{k\times k})$, it is known that for any orthogonal transformation $T\vec{X}=\vec{Y}$, $\vec{Y}$ is also distributed as $N(\vec{0}^{k\times 1}, I^{k\times k})$. Thus the univariate marginals of both $\vec{X}$ and $\vec{Y}$ are distributed as N(0,1). In Section 2.0 the marginals in set (1) were forced to approximate N(0,1) in the hope that any orthogonal transformation T on the set (i.e., rotation of the coordinate system) would yield vectors with the same property. Tables 2, 4, 6 give a measure of the goodness-of-fit of certain marginals for a small number of rotations. The distribution of $X_1 + X_2$, for example, is equivalent to that of Y_1 (except for a change in variance) where

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \cdot \\ \cdot \\ \cdot \\ Y_k \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & \cdots & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & \cdots & 0 \\ & & 1 & & \mathit{0} \\ & & & 1 & \\ & \mathit{0} & & & \ddots \\ & & & & 1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \cdot \\ \cdot \\ \cdot \\ X_k \end{bmatrix}$$

In this section, vectors representing $N(\vec{0}^{2x1}, I^{2x2})$ are generated without giving the marginals for one orientation of the coordinate system priority over all others. This is done by attempting to duplicate the property that $N(\vec{0}^{2x1}, I^{2x2})$ has circles of constant density with radii r such that r^2 is distributed as χ_2^2.

Application of the notions of the first paragraph of Section 2.0 allows us to generate q points r_i^2 $(i=1,\ldots,q)$ representing the χ_2^2 distribution. Solution of the equation $G(x) = \frac{i-1/2}{q}$ $(i=1,\ldots,q)$ where $G(x)$ is the χ_2^2 distribution function gives the value of r_i^2. Let ℓ points be spaced at equal intervals on the circumference of each circle of radius r_i. Then the $N=\ell q$ points to be generated are given by

$$(2) \qquad \vec{x}\,(i,j) = \begin{bmatrix} x_1(i,j) \\ x_2(i,j) \end{bmatrix} = \begin{bmatrix} r_i \cos\left(\theta_i + \frac{2\pi j}{\ell}\right) \\ r_i \sin\left(\theta_i + \frac{2\pi j}{\ell}\right) \end{bmatrix} \qquad \begin{matrix} i=1,\ldots,q \\ j=0,\ldots,\ell-1 \end{matrix}$$

where the θ_i are generated by van der Corput's method with R=2 over the interval $(0,2\pi)$. Thus $\theta_1 = .5(2\pi)$, $\theta_2 = .25(2\pi)$, $\theta_3 = .75(2\pi)$, $\theta_4 = .125(2\pi)$, etc. Distributing the values of θ_i uniformly on $(0,2\pi)$ causes the marginal distributions to be relatively insensitive to rotation of the coordinate system.

Let us consider the case of $\ell=1$. The points may be thought of as being generated as follows: N(=q) points are distributed at equal intervals on the circumference of a circle of any radius by van der Corput's method. The i-th point is then projected radially onto the circumference of a circle of radius r_i. Extension of this method to higher dimensions would require a way of distributing points uniformly on the surface of a hypersphere. There are well-known random methods for doing this, but use of a random method would defeat the purpose of this study. The techniques of this section have not yet been extended to more than 2 dimensions.

The methods of this section were used to generate sets of 100, 500, and 1000 2-dimensional vectors for different values of ℓ. Tests which were performed on the vectors are discussed in Section 4.2.

4.0 DISCUSSION OF TESTS PERFORMED ON FORCED-MARGINALS GENERATOR

The vectors in set (1) were examined for certain properties which are satisfied by a random vector distributed as $N(\vec{0}^{k\times 1}, I^{k\times k})$. The properties examined were:

i) Each of the k-components has the marginal distribution N(0,1).

ii) The correlation coefficient of any pair of components is zero.

iii) The sum of any ν components is distributed as $N(0,\nu)$.

iv) The sum-of-squares of any ν components is distributed as χ^2_ν.

In each of the tables which follow, a measure P is used to indicate how well set (1) satisfies the property being considered. P is the probability that a random sample of size N from $N(\vec{0}^{k\times 1}, I^{k\times k})$ would give a better result than set (1) gave for the property being examined. P might be thought of as the probability that a hypothetical, perfect pseudo-random number generator would give better results than set (1).

The following notation is used in the tables. (All quantities and distributions refer to set (1)).

N - is the number of k-dimensional vectors in the set.

x_i(i=1,...,k) - refers to the empirical distribution of the i-th marginal i.e., the distribution of the points $\{x_i(1), x_i(2), \ldots, x_i(N)\}$.

m_i - is the mean of the i-th marginal.

ρ_{ij} - is the correlation coefficient between the i-th and j-th components.

$\sum x_i$ and $\sum x_i^2$ - refer to the empirical distributions of the sum and sum-of-squares of the components. The tables show which components are included in the summation.

Thus in Tables 1, 3, 5, the mean m_i of the i-th marginal is given. The corresponding P-value is the probability that a random sample of size N from N(0,1) would yield a mean less in absolute value than $|m_i|$ The P-values associated with the ρ_{ij} give the probability that a random sample of size N from $N(\vec{0}^{2x1}, I^{2x2})$ would yield a correlation coefficient less in absolute value than $|\rho_{ij}|$. For ρ_{ij} P is given by the probability that the absolute value of a t-variable with N-2 degrees of freedom is less than $\rho_{ij} \sqrt{(N-2)/1 - \rho_{ij}^2}$. Because N is sufficiently large, the normal approximation to the t-distribution was used.

In determining the P-values of the empirical distributions x_i, $\sum x_i$, and $\sum x_i^2$, the Kolmogorov-Smirnov statistic was used. For example, in Table 4 the P-value associated with $x_{10}+x_{11}+x_{12}$ is .66. This means that the probability is .66 that a random sample of size 500 drawn from N(0,3) would yield a smaller value of the Kolmogorov-Smirnov statistic than that obtained from the empirical distribution of $x_{10}+x_{11}+x_{12}$. Similarly the probability is less than 10^{-3} that a random sample drawn from χ_3^2 would yield a smaller value of the statistic than that obtained from $x_{10}^2+x_{11}^2+x_{12}^2$.

Tables from Owen[18] were used in all of the above-mentioned cases.

4.1 Results of Testing Forced Marginals Generator

From Tables 1-6 a few patterns emerge:

a) $m_i<0$ for all cases.

b) P increases as the indices of the tabled quantities increase.

c) P increases with the number of x_i involved in the transformation.

d) P decreases with increasing N.

Except for a), these are general trends which are not always clear. b) and c) can be seen in Tables 3-6, but are not always evident in Tables 1 and 2. The occurrence of a) was foreseen in Section 2.0 from the bias in the distribution of $\phi_{R_i}(n)$ (n=1,...,N). Statement b) also seems reasonable in view of the linear dependence between $\phi_{R_i}(n)$ and $\phi_{R_j}(n)$, also mentioned in Section 2.0

It should be noted that even in cases where P becomes large, the corresponding values of m_i and ρ_{ij} remain close to zero. For example, in Table 3 the values of m_i (i=10,...,20) might be acceptable for many problems even though the associated P-values are large.

4.2 Results of Testing Forced-Circles Generator

The tests described in Section 4.0 were also performed on vector set (2) of section 3.0. The results are given in Table 7, where the notation is the same as that used in Tables 1-6. The only difference in Table 7 is that ℓ, the number of points placed on the circumference of each circle is given, and the distribution of x_1-x_2 is examined.

The values of m_i and ρ_{12} are not shown for the cases of ℓ=4,10,20. These had to be zero by the manner in which the vectors were constructed. The values of m_i and ρ_{12} and corresponding P-values are given, however, for ℓ=1. It is seen that as ℓ increases, the P-value of the empirical distribution of $x_1^2 + x_2^2$ increases. ℓ=4 seems to give the best results for all the quantities examined. The only possible advantage in taking larger values of ℓ would be to make the vectors less sensitive to orthogonal transformation.

5.0 CONCLUSIONS

The small values of P in the tables indicate that the generators of Sections 2.0 and 3.0 give excellent results for most of the quantities examined. It is clear, however, that as k becomes large, the P-values of $\sum_1^K x_i$ and $\sum_1^k x_i^2$ are too high to make the "forced-marginals" generator safe to use in simulations requiring vectors of large dimension. For the present this generator seems safe to use for the following values of N and k:

i) $1 \leq k \leq 3$, N=100

ii) $1 \leq k \leq 5$, N=500

iii) $1 \leq k \leq 10$, N=1000

It is possible that some pseudo-random number generators could yield good fits for $\sum_1^k x_i$ and $\sum_1^k x_i^2$, but it is doubtful that all of the k-marginals and correlation coefficients would be as good as those produced here. In such a case a hybrid generator could be built using the "forced-marginals" generator for low dimensions and an adequate pseudo-random number generator for higher dimensions. This would necessitate the ordering of $x_1,\ldots,x_k$ by importance so that the most important variables would be generated by the methods presented here and the remaining variables by a standard pseudo-random number generator.

If the negative bias on all of the means m_i is cause for worry, this can be easily remedied by reversing the signs of all numbers in selected marginals. This would give the selected marginals positive biases.

In problems involving only 2-dimensional vectors the "forced-circles" generator of Section 3.0 is recommended. It appears to give better results than the "forced-marginals" generator, especially for small N.

It is often suggested in the literature that a generator be tested on a problem similar to the one in question, but whose analytic solution is known. Although this is usually difficult in practice, it is also suggested for anyone considering the use of the generators presented here.

The generators presented in this paper were developed with the type of problem mentioned in the introduction in mind. Problems which are dependent on the order in which numbers appear cannot be handled by the methods given in this paper. It is for the person in need of a random number generator to decide whether the nature of his problem warrants the use of these special purpose "forced" generators.

6.0 ACKNOWLEDGMENT

I wish to thank Mr. O. R. Pardo for writing the programs to generate and test the vector sets of this study. A memorandum by him describing these programs will soon appear.

2014-HJB-mdr

H. J. Bixhorn

REFERENCES

1. Chambers, R. P., "Random-number Generation on Digital Computers," IEEE Spectrum, Vol. 4, No. 2, 1967, pp. 48-56.

2. Hull, T. E. and A. R. Dobell, "Random-number Generators" SIAM Rev., Vol. 4, 1962, pp. 230-254.

3. Kahn, H., "Use of Different Monte Carlo Sampling Techniques," Symposium on Monte Carlo Methods, John Wiley and Son, Inc., 1956.

4. Hammersley, J. M. and D. C. Handscomb, Monte Carlo Methods, John Wiley and Sons, Inc., 1964.

5. Feller, W., An Introduction to Probability Theory and Its Applications (Volume II), John Wiley and Sons, Inc.

6. Van der Corput, J. C., "Verteilungsfunktionen," Proc. Kon. Adad. Wet. Amsterdam, Vol. 38, 1935, pp. 813-821.

7. Halton, J. H., "On the Efficiency of Certain Quasi-random Sequences of Points in Evaluating Multi-dimensional Integrals," Numerische Math., Vol. 2, 1960, pp. 84-90.

8. Owen, D. B., Handbook of Statistical Tables, Addison-Wesley Publishing Co., Inc., 1962.

i	m_i	P-VALUE OF m_i	P-VALUE OF x_i
1	-.042	.33	*
2	-.047	.36	*
3	-.035	.27	*
4	-.062	.46	*
5	-.060	.45	*
6	-.106	.71	.09

(* DENOTES $< 10^{-3}$)

i \ j	2	3	4	5	6
1	-.053 .40	-.020 .16	-.046 .35	-.036 .28	-.064 .47
2		-.040 .31	-.068 .49	-.005 .04	-.034 .26
3			-.071 .51	-.042 .32	-.020 .15
4				-.009 .07	-.018 .14
5					-.010 .08

UPPER ENTRY - ρ_{ij}

LOWER ENTRY - P-VALUE OF ρ_{ij}

TABLE 1

FORCED MARGINALS

N = 100 K = 6

i \ j	2	3	4	5	6
1	.04	.04	.04	.28	.32
	.02	.02	.04	.007	.12
2		.02	.10	.24	.10
		*	*	*	.22
3			.07	.08	.53
			.01	.001	.24
4				.12	.36
				.002	.12
5					.23
					.02

(* DENOTES $<10^{-3}$)

UPPER ENTRY - P-VALUE OF $x_i + x_j$

LOWER ENTRY - P-VALUE OF $x_i^2 + x_j^2$

x_i, x_2, x_3	$x_1, \ldots, x_4$	$x_1, \ldots x_5$	$x_1, \ldots x_6$
.38	.91	.84	.81
.19	.42	.60	.63

UPPER ENTRY - P-VALUE OF Σx_i

LOWER ENTRY - P-VALUE OF Σx_i^2

TABLE 2

FORCED MARGINALS

N = 100 K = 6

i	m_i	P-VALUE OF m_i	P-VALUE OF x_i
1	-.009	.17	*
2	-.011	.20	*
3	-.009	.17	*
4	-.024	.41	*
5	-.028	.47	*
6	-.023	.40	*
7	-.036	.58	.002
8	-.042	.65	.005
9	-.021	.36	*
10	-.056	.78	.09
11	-.056	.78	.05
12	-.075	.91	.43
13	-.071	.89	.24
14	-.078	.92	.47
15	-.080	.92	.51
16	-.086	.94	.62
17	-.088	.95	.69
18	-.076	.91	.36
19	-.091	.96	.75
20	-.053	.77	.005

(* DENOTES $< 10^{-3}$)

i \ j	2	5	6	10	11	15	16	20
1	-.015 .27	-.011 .19		-.009 .15		-.018 .31		-.015 .27
5			-.011 .19	-.006 .10		-.026 .44		-.009 .15
10					.045 .68	-.010 .17		-.008 .13
15							.058 .81	.106 .98
19								.050 .74

UPPER ENTRY - ρ_{ij}

LOWER ENTRY - P-VALUE OF ρ_{ij}

TABLE 3
FORCED MARGINALS
N = 500 K = 20

i \ j	2	5	6	10	11	15	16	20
1	* *	.10 *		.08 *		.36 *		.04 .001
5			.02 .002	.37 *		.39 *		.04 .004
10					.52 .02	.54 *		.16 *
15							.70 .46	.39 .85
19								.61 .87

(* DENOTES $< 10^{-3}$)

UPPER ENTRY - P-VALUE OF $x_i + x_j$

LOWER ENTRY - P-VALUE OF $x_i^2 + x_j^2$

x_1, x_2, x_3	x_{10}, x_{11}, x_{12}	x_{18}, x_{19}, x_{20}	x_1, x_{10}, x_{20}
.08	.66	.02	.62
.02	*	.58	.18

$x_1, \cdots, x_5$	$x_6, \cdots x_{10}$	$x_{11}, \cdots x_{15}$	$x_{16}, \cdots x_{20}$	$x_1, x_5, x_{10}, x_{15}, x_{20}$
.36	.68	.98	.87	.86
.56	.19	.67	.91	.31

$x_1, \cdots x_{10}$	$x_{11}, \cdots x_{20}$	$x_2, x_4, \cdots, x_{18}, x_{20}$
.94	.999	.99
.63	.47	.46

$x_1, \cdots x_{20}$
.999
.95

UPPER ENTRY - P-VALUE OF Σx_i

LOWER ENTRY - P-VALUE OF Σx_i^2

TABLE 4

FORCED MARGINALS

N = 500 K = 20

i	m_i	P-VALUE OF m_i	P-VALUE OF x_i
1	-.005	.14	*
2	-.008	.20	*
3	-.008	.20	*
4	-.009	.24	*
5	-.011	.28	*
6	-.011	.28	*
7	-.021	.50	*
8	-.021	.50	*
9	-.019	.45	*
10	-.026	.49	.002
11	-.016	.39	*
12	-.024	.56	*
13	-.044	.83	.16
14	-.044	.83	.13
15	-.049	.88	.22
16	-.045	.84	.08
17	-.039	.79	.01
18	-.059	.94	.49
19	-.042	.81	.03
20	-.046	.86	.06

(* DENOTES $< 10^{-3}$)

	2	5	6	10	11	15	16	20
1	-.009 .22	-.006 .16		-.004 .08		-.006 .16		-.005 .13
5			-.005 .13	* .01		-.012 .30		-.008 .21
10					.032 .68	.003 .07		-.001 .03
15							.008 .21	.053 .90
19								-.083 .99

UPPER ENTRY - ρ_{ij}

LOWER ENTRY - P-VALUE OF ρ_{ij}

TABLE 5
FORCED MARGINALS
N = 1000 K = 20

	2	5	6	10	11	15	16	20
1	* *	.01 *		.002 *		.23 *		.04 .001
5			.004 *	.004 *		.10 *		.02 *
10					.16 .03	.31 *		.18 *
15							.27 .04	.39 .63
19								.48 .75

(* DENOTES $<10^{-3}$)

UPPER ENTRY - P-VALUE OF $x_i + x_j$

LOWER ENTRY - P-VALUE OF $x_i^2 + x_j^2$

x_1, x_2, x_3	x_{10}, x_{11}, x_{12}	x_{18}, x_{19}, x_{20}	x_1, x_{10}, x_{20}
.08	.08	.54	.60
.006	.17	.57	.24

$x_1, \ldots, x_5$	$x_6, \ldots, x_{10}$	$x_{11}, \ldots, x_{15}$	$x_{16}, \ldots, x_{20}$	$x_1, x_5, x_{10}, x_{15}, x_{20}$
.20	.37	.90	.98	.89
.26	.22	.23	.99	.55

$x_1, \ldots, x_{10}$	$x_{11}, \ldots x_{20}$	$x_2, x_4, \ldots, x_{18}, x_{20}$
.85	.995	.96
.72	.79	.43

$x_1, \ldots, x_{20}$
.999
.98

UPPER ENTRY - P-VALUE OF Σx_i

LOWER ENTRY - P-VALUE OF Σx_i^2

TABLE 6
FORCED MARGINALS
N = 1000 K = 20

TABLE 7

FORCED CIRCLES

$\ell = 1$	N = 100:	m_1 = -.001	P = .01
		m_2 = -.011	P = .09
		ρ_{12} = -.027	P = .21
	N = 500:	m_1 = *	P = *
		m_2 = -.003	P = .05
		ρ_{12} = -.009	P = .17
	N = 1000:	m_1 = *	P = .004
		m_2 = -.001	P = .03
		ρ_{12} = -.006	P = .14

P-values of x_1, x_2, $x_1 + x_2$, $x_1 - x_2$, $x_1^2 + x_2^2$ were found for

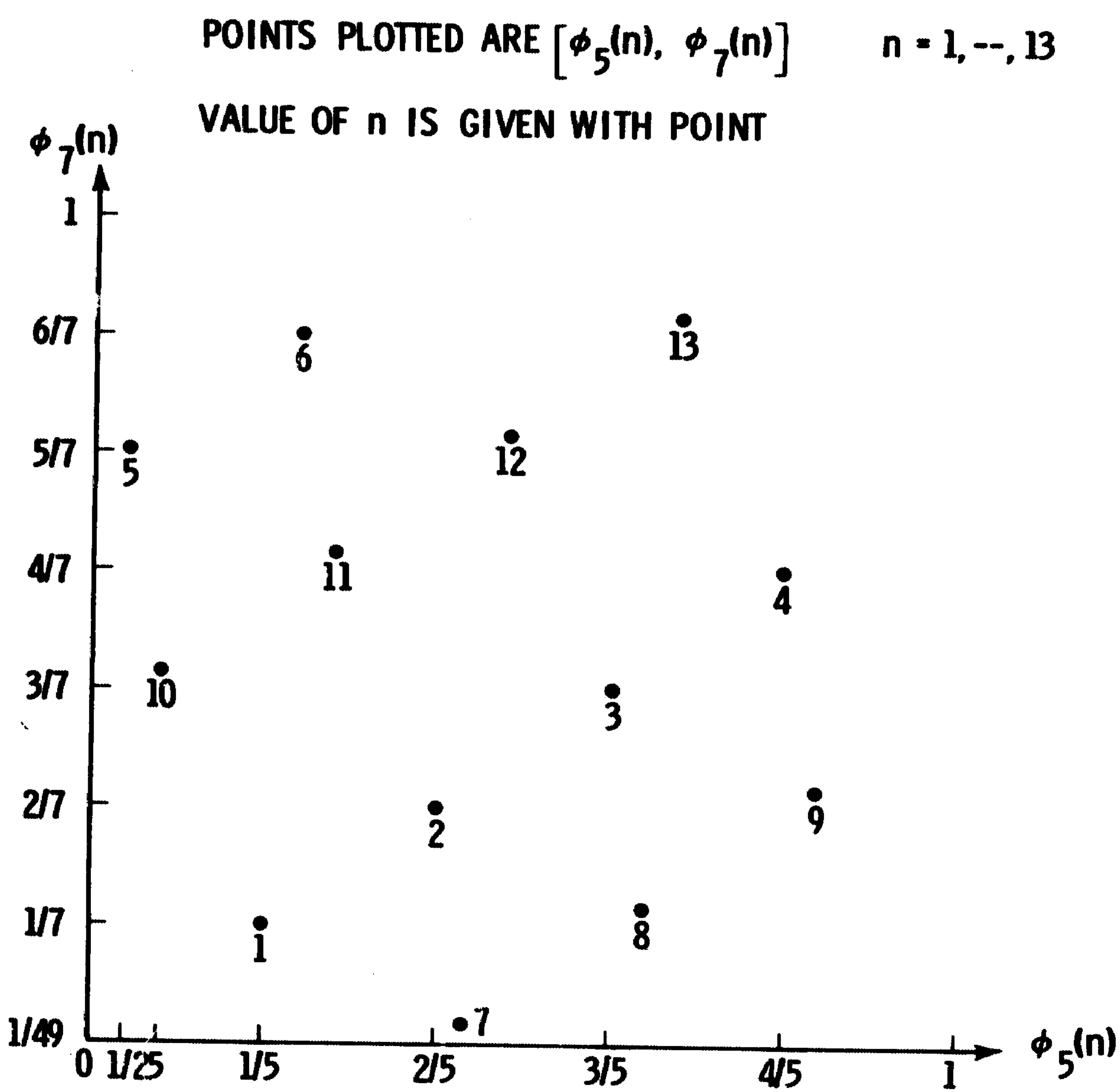
ℓ = 1, 4, 10, 20

N = 100, 500, 1000

In all cases except $x_1^2 + x_2^2$ for values of ℓ and N given below, $P < 10^{-3}$.

ℓ = 10,	N = 100	P = .04
ℓ = 20,	N = 100	P = .73
ℓ = 20,	N = 500	P = .01

*(*) Denotes $<10^{-3}$



POINTS PLOTTED ARE [$\phi_5(n)$, $\phi_7(n)$] n = 1, --, 13
VALUE OF n IS GIVEN WITH POINT
$\phi_7(n)$
1
6/7
5/7
4/7
3/7
2/7
1/7
1/49
0
1/25
1/5
2/5
3/5
4/5
1
$\phi_5(n)$
1
2
3
4
5
6
7
8
9
10
11
12
13

FIGURE 1

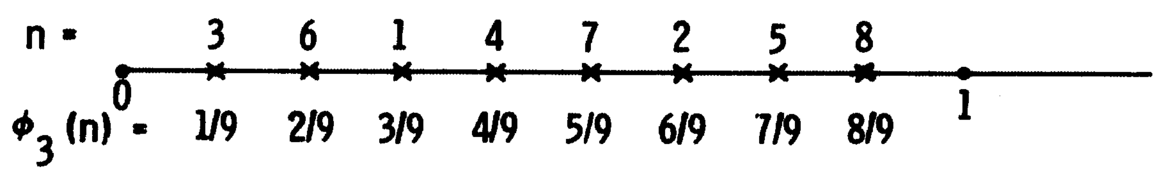


FIGURE 2