Thus,

$$P(|\bar{Z}_n| \geqq t) = 2[1 - P(\bar{Z}_n \leqq t)]$$
$$= 2\left\{1 - \frac{1}{n!} \sum_{i \leqq (n/2)(t+1)} (-1)^i \binom{n}{i} \left[\frac{n}{2}(t+1) - i\right]^n\right\},$$

and in view of the identity

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} (u-k)^n = n!$$

this becomes

$$P(|\bar{Z}_n| \geqq t) = \frac{2}{n!} \sum_{(n/2)(t+1) < k \leqq n} (-1)^k \binom{n}{k} \left[\frac{n}{2}(t+1) - k\right]^n = \Psi_n(t)$$

for $0 \leqq t \leqq 1$. The random variable $\frac{Y}{a}$ is obviously more peaked about zero than Z. Since $\frac{Y}{a}$ and Z fulfil the assumptions of Theorem 1, it follows that $\frac{\bar{Y}_n}{a}$ is more peaked about zero than $\bar{Z}_n$, that is

$$P\left(\left|\frac{\bar{Y}_n}{a}\right| \geqq t\right) \leqq P(|\bar{Z}_n| \geqq t) = \Psi_n(t) \quad \text{for} \quad t \geqq 0.$$

Setting $at = y$, one obtains (4.1).

For $n \to \infty$ the function $\Psi_n(t)$ approaches asymptotically the probability $P(|X| \geqq t\sqrt{3n})$ for the normalized normal random variable X.[4] For $n = 8$ one obtains the following values which indicate a good approximation:

t	.3998	.5254	.6711
$P(\|X\| \geqq t\sqrt{24})$	.05	.01	.001
$\Psi_8(t)$	.049	.0092	.0005.

For smaller values of n, $\Psi_n(t)$ can be easily computed.

A METHOD FOR OBTAINING RANDOM NUMBERS

By H. Burke Horton

Interstate Commerce Commission

The need for large quantities of random numbers to be used in sample design, subsampling, and other statistical problems is well known. Tippett's [1] numbers have been widely used for these purposes, despite criticism directed at their lack of randomness. The following procedure may be of interest to those

[4] Cramér, op. cit., p. 245.

who wish to develop their own random series. The method described below will ultimately be used to record extensive tables of random numbers for general use.

Current methods of producing random numbers usually depend upon single operations of mechanical or electronic devices. These may be described as "single-stage" random number processes. The numerical results are biased to the same extent as the devices from which they are taken.

At this point it is desirable to describe a process which may be called "compound" randomization. Assume two roulette wheels arranged in series so that the first controls the arrangement of symbols on the second wheel, while a turn of the second wheel determines which of its positions is to be observed. If the decimal system is used, the first wheel would have 10! "equally likely" positions, and the second would have 10 "equally likely" positions. If three such wheels were to be chained, the first would require (10!)! positions, the second 10! positions, and the third 10 positions. In general, if n wheels were to be chained, the first would require $10(!)^{n-1}$ "equally likely" positions. It is not practical to design such a machine.[1]

One method of surmounting these difficulties is to shift to the binary system in order to take advantage of the fact that $2! = 2$; or, in general, $2(!)^n = 2$. This property makes feasible the chaining of any number of machines in series; and, furthermore, the machines can be of the same design. If desired, the results taken from a single machine may be chained. Another important feature is the ease of handling binary chains by electronic systems.

The words "equally likely" have been placed in quotation marks thus far to indicate that the probabilities are as nearly equal as manufacturing precision permits. Any simple single-stage device will have some bias, and it is this very lack of true equality that the chaining process is designed to meet. For convenience we may take as our binary symbols $+1$ and -1 rather than the customary 1 and 0. We adhere to the usual rules regarding the sign of a product.

Let p_i be the probability of obtaining $+1$ in the i^{th} trial (or in the i^{th} machine of a chain of machines). $0 \leq p_i \leq 1$. $q_i = 1 - p_i$ represents the probability of obtaining -1 in the i^{th} trial.

Let P_i be the probability of obtaining $+1$ as the *product* of i trials. $Q_i = 1 - P_i$ is the probability of obtaining -1 as the *product* of i trials. The following relationships can be set down immediately:

$$\begin{aligned} P_1 &= p_1 & Q_1 &= q_1 \\ P_2 &= P_1 \cdot p_2 + Q_1 \cdot q_2 & Q_2 &= P_1 \cdot q_2 + Q_1 \cdot p_2 \\ P_3 &= P_2 \cdot p_3 + Q_2 \cdot q_3 & Q_3 &= P_2 \cdot q_3 + Q_2 \cdot p_3 \\ &\cdot & &\cdot \\ &\cdot & &\cdot \\ &\cdot & &\cdot \\ P_i &= P_{i-1} \cdot p_i + Q_{i-1} \cdot q_i & Q_i &= P_{i-1} \cdot q_i + Q_{i-1} \cdot p_i \end{aligned}$$

[1] It has been pointed out by Dr. George W. Brown that a practical solution is possible using any number base, n, by addition of random digits $(0, 1, 2, \cdots n - 1)$ modulo n.

We may calculate the bias, $P_k - \frac{1}{2}$, for a chain of k trials:

$$P_k - \tfrac{1}{2} = \tfrac{1}{2}(P_k - Q_k)$$
$$= \tfrac{1}{2}(P_{k-1} \cdot p_k + Q_{k-1} \cdot q_k - P_{k-1} \cdot q_k - Q_{k-1} \cdot p_k)$$

Factoring, we have

$$P_k - \tfrac{1}{2} = \tfrac{1}{2}(P_{k-1} - Q_{k-1})(p_k - q_k)$$

Substituting for $P_{k-1} - Q_{k-1}$ and factoring again,

$$P_k - \tfrac{1}{2} = \tfrac{1}{2}(P_{k-2} - Q_{k-2})(p_{k-1} - q_{k-1})(p_k - q_k)$$

Continuing the process of substituting and factoring, we obtain

$$P_k - \tfrac{1}{2} = \tfrac{1}{2}(p_1 - q_1)(p_2 - q_2) \cdots (p_k - q_k)$$

(1)
$$P_k - \tfrac{1}{2} = \frac{1}{2}\prod_{i=1}^{k}(p_i - q_i) = \frac{1}{2}\prod_{i=1}^{k}(2p_i - 1).$$

We may write the general formula for P_k:

(2)
$$P_k = \frac{1}{2}\left[1 + \prod_{i=1}^{k}(2p_i - 1)\right].$$

In the special case where all the p_i are equal to a constant, p,

(3)
$$P_k = \tfrac{1}{2}[1 + (2p - 1)^k].$$

This can also be derived directly by expansion of $(p - q)^k$.

If any machine, r, in the chain has no bias ($p_r = \frac{1}{2}$, exactly), the chain itself has no bias, since $2p_r - 1 = 0$. Note also that if for all i, $0 < p_i < 1$, the bias of the complete chain is less than the bias of any component (single or multiple) taken from the chain, because $|(2p_i - 1)| < 1$. Or stated another way, the results taken from any machine, no matter how nearly perfect, can be improved by chaining with another machine, no matter how biased the latter. Even in the limiting case, $p = 1$ (or 0), the magnitude of the bias remains unchanged; in all other cases it is reduced. The bias of final results can be made as small as desired by increasing the length of the chain. Compound randomization can be regarded as an attrition process which may be used to reduce final bias below any preassigned quantity. If the observations taken from two machines in the chain should be perfectly correlated, the only effect is to shorten the chain by two.

In shifting from the binary system to the decimal system, symbol bias will be introduced. In general, symbol bias will be introduced in passing from a given positional system to any other positional system, unless one of the number bases is a rational power of the other.

To illustrate, let us assume that we have a random binary series and wish to obtain a random one-digit decimal series. It will be necessary to tabulate the binary series in blocks of four symbols. The quantities will range from 0000 (binary) to 1111 (binary), or from 00 (decimal) to 15 (decimal), with equal

probabilities. There would be no predominance of either ones or zeros in the overall binary tabulation, as illustrated in the table below.

	Binary System	Decimal System
	0000	0
	0001	1
	0010	2
	0011	3
	0100	4
	0101	5
	0110	6
	0111	7
	1000	8
	1001	9
Tabulation to this point	25 zeros 15 ones	One of each symbol
	1010	10
	1011	11
	1100	12
	1101	13
	1110	14
	1111	15
Overall tabulation	32 zeros 32 ones	(Right digit only) 0–5, 2 each 6–9, 1 each

However, if we look at the right digit of the decimal tabulation, it is clear that the symbols 0 to 5, inclusive, will occur twice as often as the symbols 6 to 9, inclusive. The easiest way of correcting for this bias is simply to reject all two-digit decimal numbers which occur, thereby giving equal probabilities to the ten decimal symbols. The rejection could be accomplished most easily by electronic devices operating on the binary numbers. All numbers greater than 1001 (binary) would be excluded through the operation of a simple four-stage electronic counter.

This simple illustration also demonstrates the inefficiency of converting random four-digit binary numbers to random one-digit decimal numbers. 37.5% of the data are lost in the process of removing bias. A more efficient procedure would be to tabulate the random binary series in blocks of ten digits. The largest number that could occur would be 1 111 111 111 (binary), or 1,023 (deci-

mal). The numbers would have equal probabilities insofar as this is attainable by chaining. To obtain a random three-digit decimal series it would be necessary to reject the numbers above 999 (decimal). This would amount to only 2.34% of the available data. As before, rejection could be accomplished easily in the binary series by use of a ten-stage electronic counter.

Several promising devices are being considered for tabulating random numbers in accordance with the principles discussed herein. Electronic or electrical systems actuated by cosmic rays seem to be the most desirable. Tabulating equipment may be wired to turn out random numbers, possibly as a by-product of other card runs.

If only a few random numbers are needed, they can be obtained by much simpler methods. For example, a coin may be tossed, letting heads and tails represent $+1$ and -1, respectively The product of k successive tosses would be tabulated as the random binary variable. Products equal to $+1$ and -1 would be coded as 1 and 0, respectively. Blocks of binary symbols would then be converted to the decimal system as described above.

REFERENCE

[1] TIPPETT, L. H. C., *Random Sampling Numbers*, Tracts for Computers, No. 15, Cambridge University Press, 1927.

NOTE ON THE ERROR IN INTERPOLATION OF A FUNCTION OF TWO INDEPENDENT VARIABLES

BY W. M. KINCAID

University of Michigan

Suppose that g is a functon of one real variable x and h is an interpolation function such that $g(x) = h(x)$ for $x = x_1, x_2, \cdots, x_n$. Let $f(x) = g(x) - h(x)$ and suppose that $\frac{d^n}{dx^n} f(x)$ exists in an interval containing the points $x_0, x_1, \cdots, x_n$. Then the error in interpolation may be estimated from the well-known relation

$$f(x_0) = \frac{f^{(n)}(\xi)}{n!} (x_0 - x_1)(x_0 - x_2) \cdots (x_0 - x_n), \tag{1}$$

where ξ is some point in the smallest interval containing $x_0, x_1, \cdots, x_n$.

In the most usual case, where $h(x)$ is a polynomial of degree less than n, we have $f^{(n)}(\xi) = g^{(n)}(\xi)$.

It is natural to consider the corresponding situation for functions of two independent real variables x and y. Let g and h be two functions such that $g(x, y) = h(x, y)$ for n points $x = x_i, y = y_i (i = 1, 2, \cdots, n)$. Setting $f(x, y) = g(x, y) - h(x, y)$ as before, we have $f(x_i, y_i) = 0$ for $i = 1, 2, \cdots, n$. Then if (x_0, y_0)