INTRODUCTION

The use of mathematical models in research is a technique familiar to every actuary. The model office and the asset share are just a few examples of models commonly used by actuaries. However, actuaries have really only scratched the surface of the large volume of available techniques involving mathematical models. The realization of means for high speed computation and the concurrent development of so-called “operations research” methods have opened new doors to everyone engaged in the business of substituting “facts for appearances and demonstrations for impressions.” The practical development of techniques such as linear programming, dynamic programming, Monte Carlo experiments, and many others would have been impossible without the advent of the high-speed computer. The actuarial profession needs to keep pace with these developments and give consideration to how these techniques might enable the actuary to better fulfill his calling. The field of potential application is broad, including the provision of better information to assist management in making decisions, cost control, actuarial analysis of such things as mortality experience, retention limits, contingency reserves, etc., to mention only a few.

Mr. Boermeester (TSA, VIII, 1) has illustrated the use of the Monte Carlo method in making mortality studies for a closed group of lives. Basically, the technique consists of simulating the mortality experience of the group by the construction and solution of a model which has the same probabilistic properties. The exposure of a given life to the risk of death, an event with probability $q_x$, is simulated by the “exposure” of a random number, selected from the unit interval on the real line, to the “risk” of being less than or equal to $q_x$, an event with the same probability $q_x$. Solution of the model results in a frequency distribution of claim costs for the group being studied. The only assumption inherent in the method is that the appropriate $q_x$ is the actual probability of death for each life.

The purpose of this paper is to describe the application of the Monte Carlo technique to a practical situation in my Company, and to discuss some of the problems encountered and the solution of these problems.
THE PROBLEM

The specific problem was one of rate-making. A very real problem in the field of group term insurance is the transfer of coverage from one carrier to another by a policyholder who finds himself in a large deficit position with the original carrier. This situation can be avoided if the policyholder is willing to pay an additional charge for a guarantee of an upper limit on the amount of deficit carried forward from one year to the following years. In order to determine such a charge, it is necessary to know the probability of, the expected value of, and the variation of claims in excess of a given amount.

The basic problem to be solved, of course, is that of determining the frequency distribution of the annual claim cost of a given group of lives for a given year. It was desired that the following properties of the group be allowed to vary over rather wide ranges: (1) the size of the group, (2) the age distribution of the group, (3) the sex distribution of the group, (4) the total amount of insurance, and (5) the distribution of the insurance on individual lives.

The analytical solution of this problem would be extremely complex, and indeed any such solution which would be practical from a cost standpoint would necessitate making simplifying assumptions which would raise considerable doubts as to the validity of the conclusions.

Therefore, it was decided to use the Monte Carlo technique, which is admirably suited to a problem of this nature.

THE MONTE CARLO TECHNIQUE

The Datatron 205 was programmed to conduct the Monte Carlo experiment. The basic procedure was very similar to that described by Mr. Boermeester, the core of the program consisting of the comparison of a random number with the probability of death. The comparison routine was repeated for the entire group many times, simulating several trials of the mortality experience for the year.

Input to the program consisted of a deck of cards, one card for each life in the group being tested. Each card contained the age, the sex, and the amount of insurance in force for that life.

Output was in printed form, showing the amount of claims for each trial, the average claims for all trials, and a frequency distribution of claims.

Example

One group which we studied consisted of 306 lives, all male, with ages varying from 23 to 75 and amounts of insurance on individual lives
varying from $2,000 to $10,000. Expected claims for the group, based on  
the 1950-1954 intercompany group mortality experience, were $17,200  
for the year studied. One hundred trials were made. The average amount  
of claims per trial was $17,625, which was very close to the expected  
claims. The frequency distribution of claims is given below:
look-up and number generation methods. A standard table of 1,000 ten-
digit random numbers is stored in the memory of the computer and used
as a starting point. When these 1,000 numbers have been exhausted during
an experiment, a new set of 1,000 numbers is generated from the old as
follows:
1. A number is selected from the old table which I shall designate \( N_1 \).
2. The first number of the new table, \( N_2 \), is obtained by adding the first
   number of the old table, \( N_1 \), to \( N_1 \) and retaining only the ten least
   significant digits of the sum.
3. The second number of the new table, \( N_3 \), is obtained by adding the
   second number of the old table, \( N_1 \), to \( N_2 \) obtained in step 2, again
   retaining only the ten least significant digits of the sum.
4. In general, the \( k \)th number of the new table, \( N_k \), is obtained by adding
   the \( k \)th number of the old table, \( N_1 \), to \( N_{k-1} \) to the \((k - 1)\)th number of the
   new table, \( N_{k-1} \), retaining only the ten least significant digits of the sum.

When any table is exhausted, it is used to construct another table in
exactly the same way.

This routine is considerably faster than the more commonly encoun-
tered ones entailing multiplication, and the time factor is very important
to us since our computer is primarily designed for data processing and
does not operate at the high speeds of the larger computers designed for
scientific applications.

**TESTS OF THE RANDOM NUMBER SUPPLY**

As mentioned, the validity of the results of a Monte Carlo experiment
rests on the randomness of the number supply. The statistical properties
desired for the numbers are exactly those which would result if the num-
bers were obtained by an idealized chance device which selected numbers
from the unit interval independently and with all numbers equally likely.
The numbers produced by a computer subroutine are not random in this
sense, of course, and therefore such numbers should be tested, both the-
oretically and empirically, for various specific properties of uniformly dis-
tributed variables.

We are indebted to Mr. Gordon D. Shellard for a proof of the theorem
that the decimal part of the sum of \( n \) uniformly distributed variables on
the unit interval is itself uniformly distributed on the unit interval. This
theorem provides a theoretical basis for our method of generating random
numbers. With his kind permission, his proof is exhibited in the Appendix.

I have made several \( \chi^2 \) tests of the distribution and independence of
the set of the first 26,000 random numbers generated as I have described, and of various subsets of that set:

1. Test of goodness of fit to the uniform distribution. The unit interval was divided into ten subintervals and the frequency of numbers falling into each subinterval was observed for the entire set of 26,000 numbers and for each table of 1,000 numbers. Also, applications to groups of \( k \) lives will likely assign every \( k \)th random number to the \( k \)th life. Therefore, this test was also applied to these subsets for \( k = 20 \) and \( k = 50 \).

2. Independence test. Again, the unit interval was divided into ten subintervals and a \( 10 \times 10 \) matrix was constructed as follows: tally 1 in row \( i \), column \( j \) when a number in the \( i \)th subinterval is followed by a number in the \( j \)th subinterval. The expected result is equal numbers in all positions of the matrix. Matrices were constructed for the entire set of 26,000 numbers and for each subset of 2,000 consecutive numbers.

3. A study of runs up and down. This test was made for the entire set of 26,000 numbers, and describes the oscillatory nature of the numbers. The number of continuously increasing or decreasing subsequences of length \( l (1 \leq l < 26,000) \) was counted and compared with the theoretical distribution of such runs if the numbers are truly random.

4. A study of runs above and below the mean. This is another test which describes the oscillatory nature of the numbers, and the entire set of 26,000 numbers as well as each table of 1,000 numbers was so tested. The number of subsequences of length \( l \) of numbers all greater or less than \( \frac{1}{2} \) was counted and compared with the theoretical number of such runs if the numbers are truly random.

5. The number of even and odd numbers in the entire set of 26,000 numbers and in each table of 1,000 numbers was counted and compared with the theoretical distribution if the numbers are truly random.

For the most part, these tests gave no significant evidence that the numbers do not have the properties of uniformly distributed variables. Some indication that certain subsets of the 26,000 numbers tested do not have these properties was evident, however. In test 1, three of the fifty subsets of every 50th number fell in the 98\% "tail" and four in the 5\% "tail" of the theoretical \( \chi^2 \)-distribution. In test 2, two of the thirteen subsets of 2,000 consecutive numbers fell in the 1\% tail of the theoretical \( \chi^2 \)-distribution.

All such tests, of course, have a subjective element and no test or
series of tests can establish a sequence of numbers as being random. Also, to sound a philosophical note, if numbers were generated by a truly random process any test which rejected such a sequence would be faulty.

It should be mentioned that any sequence of numbers generated by a computer subroutine with finite input will eventually repeat or "loop." Although our subroutine fits this description, we have not attempted to calculate the period of the sequence generated by this subroutine, since it is most probably so long as to be of no practical significance to our applications.

Since the subroutine is not susceptible to degeneration, no test for this was made.

We feel quite comfortable in using the numbers generated as I have described in this paper.

CONCLUSION

As evidenced by the amount of time and effort which we have devoted to this subject, we feel that the Monte Carlo technique has many potential applications in the fields of actuarial endeavor. Indeed, many mathematical models, although they may be built of very simple variables, may themselves become very complex, particularly if probability concepts are involved. Very often, such situations lend themselves very handily to the Monte Carlo technique. Of course, the practicality of using this technique will depend to a great extent on the speed of the electronic equipment available.

It is my hope that this paper may be of some value to those actuaries who are interested in the application of these newer operations research techniques made possible by high-speed computation to many of the problems with which they are faced. I wish to express my indebtedness and appreciation to Mr. J. S. Hill who provided very helpful guidance and conceived the random number generation process described in this paper, and to Mr. Dale Kain who provided much valuable technical assistance in making the random number tests.

APPENDIX

**Theorem** The decimal part of the sum of \( n \) variables uniformly distributed on the unit interval is itself uniformly distributed on the unit interval.

**Proof** (Due to Mr. Gordon D. Shellard)

An expression for the distribution of the sum, \( y \), of \( n \) uniformly distributed variables, which may be verified by induction, is
\[ f_n(y) = \frac{1}{(n-1)!} \left\{ \begin{array}{ll}
n^{n-1} & \text{if } 0 \leq y \leq 1 \\
^{n-1} - \binom{n}{1}(y-1)^{n-1} & \text{if } 1 < y \leq 2 \\
^{n-1} - \binom{n}{1}(y-1)^{n-1} + \binom{n}{2}(y-2)^{n-1} & \text{if } 2 < y \leq 3 \\
^{n-1} - \binom{n}{1}(y-1)^{n-1} + \ldots + (-1)^{n-1}\binom{n}{n-1}(y-n+1)^{n-1} & \text{if } n-1 < y \leq n \\
 \end{array} \right. \]

If \( x \) is the decimal part of \( y \), then it follows that the distribution of \( x \) on the unit interval is given by

\[ f_n(x) = \frac{1}{(n-1)!} \left\{ \begin{array}{ll}
x^{n-1} & \\
(n+x)^{n-1} - \binom{n}{1}x^{n-1} & \\
(2+x)^{n-1} - \binom{n}{1}(1+x)^{n-1} + \binom{n}{2}x^{n-1} & \\
(3+x)^{n-1} - \binom{n}{1}(2+x)^{n-1} + \binom{n}{2}(1+x)^{n-1} & \\
\ldots & \\
(n-1+x)^{n-1} - \binom{n}{1}(n-2+x)^{n-1} + \binom{n}{n-1}(n-3+x)^{n-1} + \ldots + (-1)^{n-1}\binom{n}{n-1}x^{n-1} & \\
\end{array} \right. \]
The summation in brackets is most easily performed by summing each column separately. The sum of the first column is given by

\[ S_1 = x^{n-1} + (1 + x)^{n-1} + (2 + x)^{n-1} + \ldots + (n - 1 + x)^{n-1} \]

\[ = \left(1 + E + E^2 + \ldots + E^{n-1}\right) x^{n-1} = \frac{E^n - 1}{E - 1} x^{n-1}. \]

In general, the sum of the \( k \)th column (1 ≤ \( k \) ≤ \( n \)) is given by

\[ S_k = (-1)^{k-1} \binom{n}{k-1} \frac{E^{n-k+1} - 1}{E - 1} x^{n-1}. \]

Therefore, the sum of all the terms in brackets is

\[ S = \sum_{k=1}^{n} S_k \]

\[ = \frac{1}{E - 1} \left\{ \left[E^n - \binom{n}{1} E^{n-1} + \binom{n}{2} E^{n-2} + \ldots + (-1)^n \binom{n}{n} \right] \right\} x^{n-1} \]

\[ = \frac{1}{E - 1} \left[ (E - 1)^n - (1 - 1)^n \right] x^{n-1} = \frac{\Delta^n}{\Delta} x^{n-1} = \Delta^{n-1} x^{n-1} \]

\[ = (n - 1)!. \]

Therefore

\[ f_n(x) = \frac{1}{(n - 1)!} (n - 1)! = 1. \]
DONALD A. JONES:

At the Chicago meeting of the Society in June an informal discussion was held to discuss what the Society should do about the study of risk theory. One question raised was "How can the Society aid its members in their study of the theory?" Today two papers are being presented, Mr. Collins' and Dr. Kahn's, which can aid us in our study. The purpose of this discussion is twofold: (1) place Mr. Collins' "basic problem" in risk-theoretic language and (2) give an elementary proof of the theorem given in the Appendix.

First let us contemplate whether Mr. Collins' "basic problem" is better fitted by the collective model of risk theory or by the individual model. In the collective model the claim amounts, given the number of claims, are independently distributed. Moreover, to follow the development in Dr. Kahn's paper they must also be identically distributed. For a relatively small closed group with widely varying ages and amounts, as considered by Mr. Collins, this postulate is not realistic. In the individual model it is convenient to have a closed group of lives for which we know the individual amounts of insurance and probabilities of death. In addition, it is usual to assume independent lives. Our problem sounds like an example of the individual model.

To consider the individual model of risk theory, we need some notations for each of the m lives in the group, assumed to be numbered from 1 to m. For the ith life let $A_i$ be the amount of insurance, $q_i$ (sic) the probability of death within the year, and $Z_i$ an "indicator" random variable, i.e., $Z = 1$ if the ith life dies; $Z = 0$ if the ith life survives. Thus, $Pr\{Z_i = 1\} = q_i$, $Pr\{Z_i = 0\} = 1 - q_i$, and the total claims for the year,

$$C = \sum_{i=1}^{m} A_i Z_i.$$ 

It is in this individual model that Mr. Collins has cast $C$.

Since $C$ is defined as the sum of a fixed number of independent random variables, users of the individual model have approximated its distribution by a normal distribution with mean

$$\sum_{i=1}^{m} A_i q_i$$ 

and variance

$$\sum_{i=1}^{m} A_i^2 q_i(1 - q_i).$$
Both of the authors today warn that this asymptotic normal distribution may be approached slowly. As an alternative to this approximation Mr. Collins has suggested that we observe the experience of a large number of identical groups by simulation on the computer and thus obtain the distribution of \( C \) by statistical inference.

This discussant found it enlightening to look at approximate confidence intervals for the three parameters, \( Pr\{C > 25,000\} \), \( Pr\{C > 30,000\} \), and \( Pr\{C > 35,000\} \) which were constructed as follows. If we let \( p = Pr\{C > 25,000\} \), then the number of times the claims of the 100 Monte Carlo trials exceed 25,000 is a random variable with a binomial distribution (with parameters 100 and \( p \)). The approximate confidence intervals were then obtained by use of the normal approximation to this binomial distribution. These intervals were then compared to some of the approximate solutions obtained by analytical methods following simplifying assumptions. All these solutions fell within the approximate 95% confidence intervals.

It may be worthwhile to remember that Monte Carlo is statistical estimation and thus warrants some statement of the statistical accuracy. Along this line of thought, Mr. Collins' example raises another question. If the mean (or some other characteristic) of the empirical distribution differs from the theoretical mean (known, of course), should we adjust our result?

The theorem in Mr. Collins' paper is the assertion of a proposition for every positive integer, so it is natural to prove it directly by induction. We restate the theorem to include the needed independence of the variables.

**Theorem:** Let \( x_1, x_2, \ldots, \) be a sequence of independent random variables, each uniformly distributed on the unit interval. Let

\[
S_n = \sum_{i=1}^{n} x_i.
\]

Then for every positive integer, \( n \), \( S_n - [S_n] \) is uniformly distributed on the unit interval ([\( u \)] denotes the greatest integer in \( u \)).

**Proof:** It is obvious for \( n = 1 \). We will need the assertion at \( n = 2 \) to prove the inductive step, hence we must do it directly. For every \( t \),

\[
0 < t < 1, \quad Pr\{S_2 - [S_2] \leq t\} = Pr\{x_1 + x_2 \leq t\}
\]

\[
+ Pr\{1 < x_1 + x_2 \leq 1 + t\} = \frac{p^2}{2} + \left[1 - \frac{(1 - t)^2}{2}\right] = t.
\]
To prove the inductive step, we assume $S_{n-1} - [S_{n-1}]$ is uniformly distributed over the unit interval.

$$S_n - [S_n] = S_{n-1} + x_n - [S_{n-1} + x_n]$$

$$= \{S_{n-1} - [S_{n-1}] + x_n\} - \{[S_{n-1} + x_n] - [S_{n-1}]\}$$

$$= \{S_{n-1} - [S_{n-1}]\} + x_n - [S_{n-1} - [S_{n-1}] + x_n].$$

The last equality follows from $[u] - [v] = [u - [v]]$ for all real numbers $u$ and $v$. Now applying the theorem for $n = 2$ to the two variables $S_{n-1} - [S_{n-1}]$ and $x_n$ establishes the theorem.

I thank Mr. Collins for his paper, which I found stimulating.

ROBERT C. TOOKEY:

We are indeed indebted to Mr. Collins for his fine paper which so vividly illustrates how electronic data-processing equipment ushered in a completely new era. The electron, smallest unit of matter, traveling with a speed of light, has become the scientist’s willing and valuable slave. The significance of all this defies description. The newest model electronic machines in particular provide the actuary with his most valuable tool. He can now simulate the experience to obtain the data necessary for his computations. Perhaps the greatest application of these simulated experiences will be in collective risk problems which are so complex that any analytical approach would prove to be extremely cumbersome and costly.

In addition to the calculation of the “extra risk” or “pooling” charge in group life insurance, as the author demonstrates, the Monte Carlo technique has application in determination of nonproportional reinsurance rates. Our operations research affiliate, Peat, Marwick, Caywood, Schiller & Co., using an IBM 1620, computed the claims distribution of a small (around sixty million in force) life insurance company by simulating the experience of 1960 one thousand times. Two distributions were obtained, one assuming a retention limit of $25,000$ on any one life, and the other assuming this limit to be $100,000$. The average policy size was about $4,000$, and, as Table 1 demonstrates, a $25,000$ retention limit produced a distribution of claims well defined by the normal curve. On the other hand, with a retention limit of $100,000$, we end up with a distribution of unusually marked skewness—an unmanageable skewness as far as classical frequency distributions are concerned. Had our client based his computations on the normal distribution, he would have ended up with nonproportional reinsurance premiums that were at least 50% on the low side.
The Datatron used by the author is a fine machine, just as the Model T Ford was a fine car in its day. However, the capacity of the IBM 1620 as well as other modern machines has paved the way for a far broader extension of Monte Carlo techniques than hitherto thought economically feasible. Claims experience in almost any life insurance company can be simulated in a nominal amount of machine time, the cost of which would fall well within the loading of nonproportional premiums. Pseudo-random numbers can be generated as needed and in quantities of 50 million or more using techniques that will now be discussed.

As the author pointed out, tables of random numbers have limited performance in high-speed computing machines because they require extensive storage. For this reason, the standard procedure has been to generate numbers in the machine at the time they are needed.

These mathematical procedures produce sequences of pseudo-random numbers, that is, completely determined sequences of numbers having some of the more easily tested properties of the sequences of similarly distributed truly random numbers. The very determinability of the sequences produced by these procedures provides definite advantages over the physical methods which produce sequences of truly random numbers. A sequence of pseudo-random numbers which satisfactorily passes random number tests can be repeatedly used. On the other hand, if a different sequence is used each time a problem is solved by the Monte Carlo method, the possibility of an unexpected answer's being due to numbers

<table>
<thead>
<tr>
<th>Total Claims in Calendar Year</th>
<th>Number of Trials</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Retention Limit</td>
</tr>
<tr>
<td></td>
<td>Expected Claims</td>
</tr>
<tr>
<td>Less than $100,000</td>
<td>0</td>
</tr>
<tr>
<td>$100,000 to 150,000</td>
<td>5</td>
</tr>
<tr>
<td>$150,000 to 200,000</td>
<td>97</td>
</tr>
<tr>
<td>$200,000 to 252,000</td>
<td>437</td>
</tr>
<tr>
<td>$252,000 to 300,000</td>
<td>340</td>
</tr>
<tr>
<td>$300,000 to 350,000</td>
<td>108</td>
</tr>
<tr>
<td>$350,000 to 400,000</td>
<td>12</td>
</tr>
<tr>
<td>$400,000 to 450,000</td>
<td>1</td>
</tr>
<tr>
<td>Over $450,000</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1,000</td>
</tr>
</tbody>
</table>

* The bracket was $200,000-$262,000 to break at the mean.
† The bracket was $262,000-$300,000 to break at the mean.
used could be checked simply by reproducing the sequence used and mak-
ing statistical tests upon it. Departures from mean behavior are bound
to occur in the long run in truly random sequences, whereas in mathe-
matical procedures these departures can be eliminated.

As pointed out by the author, the middle of the square method used
by Mr. Boermeester (TSA, VIII) had its limitations because on the aver-
age, after several tens of thousands of numbers, it can be expected that
the sequences will degenerate to a sequence of zeros or to cycles of short
period. To prevent this from happening, it was necessary to examine and
test the number produced and if necessary start all over again with a new
random number. For truly large tests, simple methods producing a
greater abundance of pseudo-random numbers are needed.

The object is to produce the digits of one pseudo-random number from
its predecessor by reducing the product of the predecessor and a well-
chosen constant to its least positive residue with respect to a modulus
particularly chosen for the machine. The procedure produces sequences of
numbers closely approximating the sequence of uniformly distributed
random numbers. In the case of the IBM 1620, the number \[1 + 10
\(4r + 1\)] may be used as the multiplier providing it is less than \(10^{10}\) but
is as close thereto as possible. \(r\) must be an integer and \(k\) must be an
integer prime to 10, and \(4r + 1\) must not be divisible by 5. The length
of the period of the sequence of numbers produced by this random gen-
eration process is \(5 \cdot 10^8\). Using 10 digit multipliers, \((\beta = 10)\) the "loop-
ing" period is \(5 \cdot 10^8\). We let \(r = 53\) and \(k = 3\), producing our fixed multi-
plier, the number—9,677,214,091—and then we squared it. From this
result, we removed the 10 right-hand digits to obtain the variable multi-
plier to apply to the original fixed multiplier. From each product so ob-
tained, we removed the middle 12 digits to use as our random sequences.
Three sequences of four digits enabled us to make three trials for each
random number generated since we carried the \(q_x's\) out only four decimal
places. Application of the usual statistical tests on the number sequences
did not reveal any significant departures from mean behavior.

Using a little number theory, we can demonstrate that the generation
process described has a period of \(5 \cdot 10^8\), that is, that the variable multi-
plier does not repeat before 500,000,000 multiplications.

Let \(x\) be the "fixed multiplier" and \(y\) be any particular variable "mul-
tiplier." Then we show that if \(y(x)^{a_1}\) has the same last 10 digits as \(y(x)^{a_2},
\(a_2 \neq a_1\), then \(a_2 - a_1 \geq 5 \cdot 10^8\). In number theory terminology, this rela-
tionship is written:

\[-(1): yx^{a_1} \equiv yx^{a_2} \pmod{10^{10}}\]
APPLICATION OF THE MONTE CARLO TECHNIQUE

(meaning $y(x)^{a_1} \text{ minus } y(x)^{a_2}$ is divisible by $10^{10}$)—and since $(y, 10^{10}) = 1$
(meaning $y$ is prime to $10^{10}$), this is equivalent to (2): $x^{a_1} \equiv x^{a_2} \pmod{10^{10}}$
or (3): $x^a \equiv 1 \pmod{10^{10}}$ where $b = a_2 - a_1$.

Juncosa proved that if $x = [1 + 10 (4r + 1)]^k$ where $4r + 1$ is not divisible by 5 and $(k, 10) = 1$, then $x^{5 \cdot 10^8} \equiv 1 \pmod{10^{10}}$ and $x^k \neq 1 \pmod{10^{10}}$ if $b < 5 \cdot 10^8$.

The proof, which is somewhat too lengthy to be included here, falls in three parts:

(a) any fixed multiplier has a period which is a divisor of $5 \cdot 10^8$;
(b) there exists a fixed multiplier with the full period (a "primitive root");
(c) any multiplier with the form $[1 + 10 (4r + 1)]^k$, where $4r + 1$ is not divisible by 5 and where $k$ is prime to 10 is a primitive root.

Thus, a number fitting the foregoing form can be used in 500 million multiplications, producing, incidentally, 6 billion random digits before looping.

We are just witnessing the very beginning of the application of powerful machine techniques that will simulate the experience, in just a few hours of operation, that would normally transpire in thousands of years. While these techniques will be very useful to the actuary, they will also make great demands on his time and ingenuity. However, the actuary can take some comfort in the knowledge that he can never be completely replaced by an electronic brain because the designers have not been able to come out with a machine that can be taught how to worry.

NATHAN F. JONES:

So far as I know, specific presentation in actuarial journals of Monte Carlo technique has been limited to Mr. Boermeester's paper, to which Mr. Collins refers. The Society is indebted to Mr. Collins for its first presentation of an actuarial problem which could hardly be handled other than by "Monte Carlo" means.

Mr. Collins' problem is, essentially, the problem of rating "aggregate excess of loss," or "stop-loss," insurance or reinsurance. Mr. J. S. Hill mentions the applicability of Monte Carlo methods to this problem in his discussion (TSA, XII, 54) of Mr. Feay's paper on nonproportional reinsurance. Actually, were it not for other underwriting and marketing problems of "stop-loss," Monte Carlo methods would probably be better known to the American actuary in all fields of insurance.

Mr. Collins, like Mr. Boermeester, emphasizes the problem of random number generation. I am sure Mr. Collins would not wish readers of his paper to reach the false conclusion that this is the principal problem for the actuary who needs to use Monte Carlo methods in the solution of real
actuarial problems. Sufficient randomness of the pseudo-random numbers employed is very important. However, simple and elaborate tables of pseudo-random numbers are available, in "hard copy," punched card, and, I am sure, tape form. Even specialized tables are available—e.g., of random normal deviates (abhorred by one author, who says he cannot consider deviates normal).

There is also a voluminous literature on the mechanical generation of pseudo-random numbers; a recent article in *SIAM Review* has a 148-item bibliography. The digital computer manufacturers have standard routines—"software"—for this purpose.

Much more important than random number generation in practice are the estimation of accuracy and the devising of methods—algorithms—for obtaining satisfactory accuracy at reasonable cost in time, both human and machine. Mr. J. E. Walsh, an Associate of this Society, the author of two of the above 148 items on random number generation, has written also on estimation of accuracy in the proceedings, published by Wiley, of a 1954 Symposium on *Monte Carlo Methods*; his second paper in that volume is entitled "A Monte Carlo Technique for Obtaining Tests and Confidence Intervals for Insurance Mortality Rates."

Mr. Hill, in the discussion referred to above, emphasizes these problems indirectly when he says, "In such situations, Monte Carlo techniques can be of assistance, but they tend to use large amounts of computer time when relied on solely." Mr. Collins employed 100 trials to obtain his results. But if anyone attempted to employ Mr. Boermeeester's IBM 650 method to obtain 100 trials of Mr. Collins' type of problem, he would rapidly learn the truth of Mr. Hill's statement. I know, because we once tried.

Even then, how do we (or Mr. Collins) know 100 trials is enough? Reproduction of the mean of the entire claim distribution is not always a sufficient criterion for the uses we may wish to make of the distribution.

Of course, Monte Carlo methods are much more easily applied on a larger scale scientific computer, if one is available. But, even then, the ingenuity in devising "algorithms" on which the actuarial profession traditionally (and rightly) prides itself has a fascinating and very worthwhile opportunity. This is particularly true of sequential or "series" problems.

I suggest, as an appropriate challenge in the pension field, a measure of the probability that costs for a sizable pension fund, ten years hence, will be more than x per cent greater by the unit purchase method than by the entry-age normal (or level-premium) methods. Certainly, for this problem, Mr. Collins' method would rapidly wear out his Datatron 205 (although the method is easy and adequate for his present problem).
There is a substantial literature in this aspect of the field also. See, for example, *International Abstracts in Operations Research*. To those who, like me, tend to find their mathematics inadequate for properly scholarly presentations, we should be glad to make available (i) a presentation of a simplified claim distribution problem, which considers estimation of accuracy and "variance-reduction" in an elementary way; and (ii) a comparison of three manual Monte Carlo methods (and their combination) with approximate integration and other conventional methods for evaluating a moderately complex actuarial function—*provided* you consider letting me (and Mr. Collins) know of your own conclusions and recommendations.

**PATRICK C. FISCHER:**

The method of obtaining pseudo-random numbers used by Mr. Collins is quite interesting and appears to be most useful in computer programs where speed of computation is important and storage is plentiful. However, I wish to point out that alternative methods for generating pseudo-random numbers are neither as few in number nor as uniformly subject to degeneration as one might infer from this article. In particular, there are methods for which the period of the generated numbers has been proved to be adequately long (on the order of $10^8$).

Some of these alternative methods are much more economical of machine storage. The time taken to produce each pseudo-random number may be longer, but one should observe that doubling the length of time necessary to execute a random-number-generating subroutine of a program may increase the total computation time of the program by only a very small fraction.

Two good papers on the generation of pseudo-random numbers are:


**JAMES C. HICKMAN:**

I am delighted to see a paper in our *Transactions* reporting results obtained by the Monte Carlo approach to solving the difficult problem of finding the distribution of total losses caused by deaths among insured lives. Mr. Collins' point that modern electronic computing equipment now makes this approach feasible is well illustrated by his example.

The Appendix containing the theorem upon which Mr. Collins' random-
number generator is based is of great interest. The statement that the probability density function (pdf) of the random variable $Y$ may be found by induction is certainly correct. However, I think it is instructive to point out the long and interesting history of the problem. It was initially solved by Laplace [4]; other solutions by Rietz [5], Hall [1], and Irwin [3] appeared about a century later. Uspensky's well-known probability textbook [6] contains a solution using characteristic functions and the inverse Fourier transform. The fact that the pdf of the sum of $n$ statistically independent and uniformly distributed random variables is made up of arcs of polynomials of degree $n - 1$ and that this pdf appears to become increasingly bell-shaped as $n$ increases, suggests a limiting distribution theorem and adds to the interest in the result.

In finding the pdf of random variable $Y_n$, the sum of $n$ independent random variables each distributed uniformly over the unit interval, by induction Mr. Collins may have in mind the use of the convolution formula

$$f_n(y_n) = \int_0^1 f_{n-1}(y_n - z) f_1(z) \, dz,$$

where $f_n(y_n)$ is the pdf associated with $Y_n$. Note that $Y_n = Y_{n-1} + Z$ where $Z$ is the $n$th random variable in the sum. Students of statistics will recognize this as an example of the change of variable method for finding the pdf of a function of continuous random variables. This method is described within the current actuarial syllabus by Hoel [2].

Employing this method, we can easily show that

$$f_2(y_2) = \int_{y_1}^{y_2} dz = y_2, \quad 0 \leq y_2 \leq 1,$$

$$= \int_{y_1}^{1} dz = 1 - y_2, \quad 1 < y_2 \leq 2,$$

a triangular distribution.

If for the purposes of an induction proof we assume the pdf displayed in the paper true for a sum with $n - 1$ terms, we have for the pdf of $Y_n = Y_{n-1} + Z$

$$f_n(y_n) = \int_0^{y_n} f_{n-1}(y_n - z) \, dz, \quad 0 \leq y_n \leq 1,$$

$$= \int_0^1 f_{n-1}(y_n - z) \, dz, \quad 0 < y_n \leq n - 1,$$

$$= \int_{y_n-n+1}^{1} f_{n-1}(y_n - z) \, dz, \quad n - 1 < y_n \leq n.$$
Substituting for \( f_{n-1}(y_n - z) \) from the inductive hypothesis, we have

\[
f_n(y_n) = \int_0^{y_n} \frac{(y_n - z)^{n-2}}{|n-2|} \, dz, \quad 0 \leq y_n \leq 1,
\]

\[
= \int_0^{y_n-k} \frac{1}{|n-2|} \left[ (y_n - z)^{n-2} - \binom{n-1}{1}(y_n - 1 - z)^{n-2} + \ldots + (-1)^k \binom{n-1}{k}(y_n - z - k)^{n-2} \right] \, dz + \int_{y_n-k}^{1} \frac{1}{|n-2|} \left[ (y_n - z)^{n-2} - \binom{n-1}{1}(y_n - 1 - z)^{n-2} + \ldots + (-1)^{k-1} \binom{n-1}{k-1}(y_n - z - k + 1)^{n-2} \right] \, dz,
\]

\[
k \leq y_n \leq k + 1, \quad k = 1, 2, \ldots, n - 2,
\]

\[
= \int_{y_n - n+1}^{1} \left[ (E_{y_n} - 1)^{n-1} + (-1)^n \right] \frac{(y_n - z - n + 1)^{n-2}}{|n-2|} \, dz, \quad n - 1 < y_n \leq n.
\]

Note the equivalent but shorter expression for \( f_{n-1}(y_n - z) \) when \( n - 1 < y_n \leq n \). We now evaluate the integrals.

\[
f_n(y_n) = \frac{y_n^{n-1}}{|n-1|}, \quad 0 \leq y_n \leq 1,
\]

\[
= \frac{1}{|n-1|} \left[ (-1)(y-z)^{n-1} \Big|_0^1 + \binom{n-1}{1}(y_n - 1 - z)^{n-1} \Big|_0^1 \right.
\]

\[
- \binom{n-1}{2}(y_n - 1 - z)^{n-1} \Big|_0^1 + \ldots
\]

\[
+ (-1)^{k-1} \binom{n-1}{k-1}(y_n - z - k + 1)^{n-1} \Big|_0^1
\]

\[
+ (-1)^k \binom{n-1}{k}(y_n - z - k)^{n-1} \Big|_0^1
\]

\[
= \frac{1}{|n-1|} \left[ y_n^{n-1} - \binom{n-1}{1}(y_n - 1)^{n-1} + \ldots
\]

\[
+ (-1)^k \binom{n-1}{k}(y_n - k)^{n-1} \right], \quad k \leq y_n \leq k + 1,
\]

\[
k = 1, 2, \ldots, n - 2,
\]

\[
= (-1)^n+1 \frac{(y_n - n)^{n-1}}{|n-1|}, \quad n - 1 < y_n \leq n.
\]
Rewriting the last line of the conjectured result as
\[
\frac{(E_1 - 1)^n + (-1)^{n+1}}{n-1} \frac{(y_n - n)^{n-1}}{n-1} = \frac{\Delta^n + (-1)^{n+1}}{n-1} \frac{(y_n - n)^{n-1}}{n-1}
\]
and comparing our result with the conjecture completes the induction proof.

If we now let \( Y_n = [Y_n] + X \) where \([Y_n]\) is the greatest integer in \( Y_n \) then

\[
g(x) = \sum_{[y_n]=0}^{n-1} k([y_n], h(x|y_n))
\]

where \( g(x) \) and \( k([y_n]) \) are the pdfs of the indicated variables and \( h(x|y_n) \) is the conditional pdf of \( X \) given \([Y_n]\). Noting that

\[
h(x|y_n) = f_n([y_n] + x) / k([y_n])
\]

we have Mr. Collins' form of the pdf of \( X \) which he now proceeds very ingeniously to show equal to 1.

A direct method of achieving this result is to consider the two possible cases when \( Y_n - 1 \) is known, and we seek to make a probability statement about \( X \).

\[
\text{Prob}[x \leq a | y_n - 1 = b] = \int_0^{a - (b - [b])} dz + \int_{1 - (b - [b])}^1 dz = a,
\]

if \( b - [b] < a \), and \( \int_{1 - (b - [b])}^{1 - (b - [b]) + a} dz = a \),

if \( a < b - [b] \), \( 0 \leq b \leq n-1 \), \( 0 \leq a \leq 1 \).

We see that in both possible situations the distribution of \( X \) is independent of \( b \) and equal to \( a \). Thus \( X \) has a uniform distribution over the unit interval.

REFERENCES

I wish to express my gratitude for both the quantity and the quality of the discussion which has added so much to the value of the paper. It is clear from these discussions that there is a genuine and spreading interest in the application of modern operations research methods to actuarial problems.

It was certainly not my intention to imply that there is a dearth of satisfactory methods for generating random numbers for Monte Carlo experiments. As Dr. Fischer and Messrs. Tookey and Nathan Jones have all made quite clear, this problem has been extensively dealt with, and there are many methods available. However, each user must find a method which is best suited to his particular situation. The choice of methods will depend on many things, among them the type of equipment and amount of machine time available, the particular problem to be solved, etc. The method described in the paper, which as far as we know is original, appeared to be admirably suited to our situation. This does not mean, of course, that it will be the best method for everyone. In any event, the actuary should be sure that the method used has been adequately tested. The paper describes some common tests which may be used if further testing is necessary.

Dr. Donald Jones makes note of the fact that Monte Carlo experiments are essentially statistical estimation and, as such, that statements as to statistical accuracy may be attached to them. Mr. Jones also raises the question of how many trials are needed to provide an adequate sample, which is again a question of statistical accuracy. Dr. Jones points out that the number of times claims will exceed any fixed amount is a random variable with a binomial distribution and, therefore, approximate confidence intervals can be obtained for the probability of this event. The reader may be interested in the range of these confidence intervals. Referring to the example in the paper, we estimated that there is an 18 per cent chance that claims for one year will exceed $25,000. Using the method described by Dr. Jones, we can state with 95 per cent confidence that this probability lies somewhere between 10 1/2 and 25 1/2 per cent.

In the particular problem considered, however, we were attempting to determine a pooling charge. Determination of a stop-loss premium or non-proportional reinsurance charge considered by Mr. Tookey is basically the same problem. In any of these cases, we are concerned not only with the
probability that claims exceed a certain amount but also with the expected value and variation of the claims in this "tail" of the distribution. I expect that it is considerably more difficult to obtain a good confidence interval for the net premium for any of these benefits. This problem would appear to me to involve so-called "distribution-free methods," of which I am quick to admit I have little knowledge.

Dr. Jones also raises another interesting question. Should we adjust our results to reflect the difference between the theoretical mean and the sample mean? Again I find myself, together with Mr. Jones, among those who sometimes must recognize the inadequacy of their mathematics. However, it does seem likely that, in view of our primary interest in the behavior in the "tail" of the distribution, a discrepancy of the order of 2½ per cent between the two means, such as occurred in my experiment, would not have a significant enough effect on the net premium as to make it a practical matter to consider such an adjustment.

We are indebted to Dr. Hickman for his very interesting treatment of the history of the solutions to the problem of determining the probability distribution function of the sum of uniformly distributed random variables. His presentation of an induction proof for this probability distribution function, which was stated without proof in the Appendix, as well as both his and Dr. Jones's proofs of the theorem stated therein are valuable additions to the material contained in the paper.

Both Mr. Jones and Mr. Tookey suggest further uses and refinements of applications of the Monte Carlo technique to actuarial problems. Certainly, the experiment which I have described is an elementary, albeit useful, application of this technique. I would certainly agree with both that we have not even begun to explore the possibilities of this method. However, as Mr. Tookey points out, these modern operations research techniques, while powerful and useful, will make great demands on the actuary's time and abilities. In presenting this paper, I expressed the hope that discussion would suggest other fruitful paths of inquiry, and that hope has been realized. Once again, I wish to thank all those who discussed the paper for their invaluable contributions to its interest.