TRANSACTIONS OF SOCIETY OF ACTUARIES
1956 VOL. 8 NO. 20

VOL. VIII MEETING NO. 20

TRANSACTIONS

APRIL, 1956

FREQUENCY DISTRIBUTION OF MORTALITY COSTS

JOHN M. BOERMEESTER

INTRODUCTION

IN THE conduct of the life insurance and annuity business, practical questions arise from time to time concerning the determination of limits of insurance or income which may be accepted with respect to any one life or in determining the magnitude of a contingency reserve which is to be established to provide for fluctuations in the expected mortality rates. In order to answer questions of this type and also possibly others, an actuary would want to have available for his study a frequency distribution of the expected costs, for the particular benefits guaranteed under the contracts, deemed to be applicable to the class of lives under consideration.

If the guarantees provide for widely varying benefits, he would normally encounter difficulty, first, in developing formulas theoretically appropriate for the specific situation and, secondly, in making the appropriate calculations to obtain a numerical solution as required for his guidance in the matter of making decisions. The first of these difficulties often stems from the subtleties in the reasoning process required to analyze the unique distributions of the important variables of age, plan, death benefits and amounts of income which are deemed to exist for the specific situation. The second of these difficulties arises from matters concerning time, expense and personnel.

The purpose of this paper is twofold. First, it presents an outline of how these difficulties may be overcome, at least in part, by applying the concept of using a series of random numbers with a view of making mortality studies for each life in a group of annuitants. From such mortality studies emerges the means to compute, with reference to the guarantees attributable to each life, the specific values which are required to form a frequency distribution of costs. The second purpose is to present the results of some experiments which were designed to discover whether or not the concept could be practically conducted on one of the IBM No. 650 electronic data processing machines installed in the John Hancock Mutual Life Insurance Company.

BASIC CONCEPT

Let us first consider this question: What conditions determine whether or not a unit annuity payment will be made at the end of one year to a particular annuitant now at ratable age x?

Suppose a random number is selected from a series of numbers composed of 10^g numbers ranging from 0 to $10^g - 1$. Call this random number ${}^gN_x$ and suppose further that it is compared with $10^g\, q_x$, where g is the number of decimal places in the series of mortality rates q_x. In the text which follows, the assumption is made that each particular life is a sample from a parent universe for which the tabular rates of mortality q_x are the true rates of mortality.

If ${}^gN_x \geq 10^g\, q_x$, the assumption may be made that the annuitant lives one year and that, therefore, the annuity payment will be due. On the other hand, if ${}^gN_x < 10^g\, q_x$, the assumption may be made that the annuitant dies within one year and that, therefore, the annuity payment will not be due.

If an electronic data processing machine can generate a series of random numbers in a small fraction of a second, it may be programmed to make a great number of comparisons of selected values of $10^g\, q_z$ for the range $x \leq z < \omega$ with such a series. In effect, as will be outlined in the following text, a hypothetical cost computation for a particular life or a group of lives may be made in a relatively short time. With successive repetitions of the process, a pattern will develop to indicate the range of probable costs within a given confidence limit.

Let us consider how this concept of using random numbers may be used to calculate the present value of annuity payments of 1 per annum to an annuitant at ratable age x under a hypothetical cost computation for this particular life alone. The following steps designed to accomplish the desired computation have been programmed for an IBM No. 650 machine.

1. Generate a random number ${}^gN_x$ and compare it with $10^g\, q_x$. If ${}^gN_x < 10^g\, q_x$ we may assume that the annuitant dies before attainment of age $x + 1$ and the cost for the annuity will, therefore, be $a_{\overline{0}|} = 0$.
2. If ${}^gN_x \geq 10^g\, q_x$, we may assume that the annuitant lives to age $x + 1$ and the cost for the annuity will be at least $a_{\overline{1}|}$. Generate another random number ${}^gN_{x+1}$, and compare it with $10^g\, q_{x+1}$. If, as a result of the second comparison the annuitant is assumed to die, the cost for the annuity will be $a_{\overline{1}|}$. On the other hand, if the annuitant is assumed to live to age $x + 2$, the operation of comparing a new random number ${}^gN_{x+2}$ with $10^g\, q_{x+2}$ will be performed to establish whether the annuity cost will be greater than or equal to $a_{\overline{2}|}$.
3. In general, if at the rth trial the assumption first may be made that the annuitant dies, then the cost of the annuity for this particular life will be $a_{\overline{r-1}|}$.

A repetition of the above operation performed, say, n times, with each repetition using a different series of random numbers, will provide the means of constructing a frequency distribution of costs for a single life. Any set of n repetitions will, of course, have its own statistical characteristics since, under sampling theory, each sample drawn at random from a homogeneous parent universe would have its own sample statistical values. If n were large, one would expect that the characteristics of the frequency distribution of costs therefor would not differ materially from those which would emerge under the classical approach of analysis.

An experiment designed to conduct a limited number of 100 mortality studies for an annuitant with entry age 65 was performed on an IBM No. 650 machine to give the results shown in Table 1.

TABLE 1

100 MORTALITY STUDIES FOR A SINGLE LIFE
ENTRY AGE 65—ANNUITY PAYMENTS OF 1 PER ANNUM

t	(1) Unit Cost $a_{\overline{t}\rceil}$	(2) Number of Studies with Unit Cost $a_{\overline{t}\rceil}$	(3) Accumulation of Entries in Col. (2)	(4) Accumulated Cost
0	.00	1	1	.00
1	.98	2	3	1.96
2	1.93	2	5	5.82
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
13	10.98	4	47	330.12
14	11.69	4	51	376.88
15	12.38	3	54	414.02
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
25	18.42	3	92	1018.32
26	18.95	3	95	1075.17
27	19.46	2	97	1114.09
28	19.96	3	100	1173.97

The machine time required to make the complete experiment of 100 trials was approximately 5 minutes. A rate of interest of $2\frac{1}{2}\%$ was used in connection with the Annuity Table for 1949, Males, without projection. The average annuity unit value of 11.74 as indicated by the final entry in column (4) of Table 1 is reasonably close to the expected tabular value of $a_{65} = 11.50$. The entry in column (1) of the unit annuity cost corresponding to entry 95 in column (3) does not differ greatly from the theoretical value established by the classical theory for the 95% confidence limit.

A second and parallel experiment of conducting 100 mortality studies for a single life was performed for the same stated conditions, except that different random numbers were generated. This second run produced an average annuity of 11.68; it showed a maximum unit annuity cost of 21.40 as compared to the maximum unit annuity cost of 19.96 shown in the first run.

Appendix B discusses the method used to generate the series of random numbers.

THE GENERALIZED CASE FOR COMMON ANNUITIES

Let us now consider how to apply this concept of using random numbers to make mortality studies for a particular class of n annuitants in order to obtain a distribution of present value costs for specified annuity payments and associated death benefits. For a particular numbered life j within this class of n lives, let x_j = age at entry and I_j = annual income payable. Three common plans of annuities may be present: life annuity with no refund; life annuity with B_j years certain at issue; or life annuity with B_j years of decreasing death benefit at issue, the yearly decrease being I_j for each year of curtate duration to date of death. For practical reasons, I_j and B_j are taken as whole numbers.

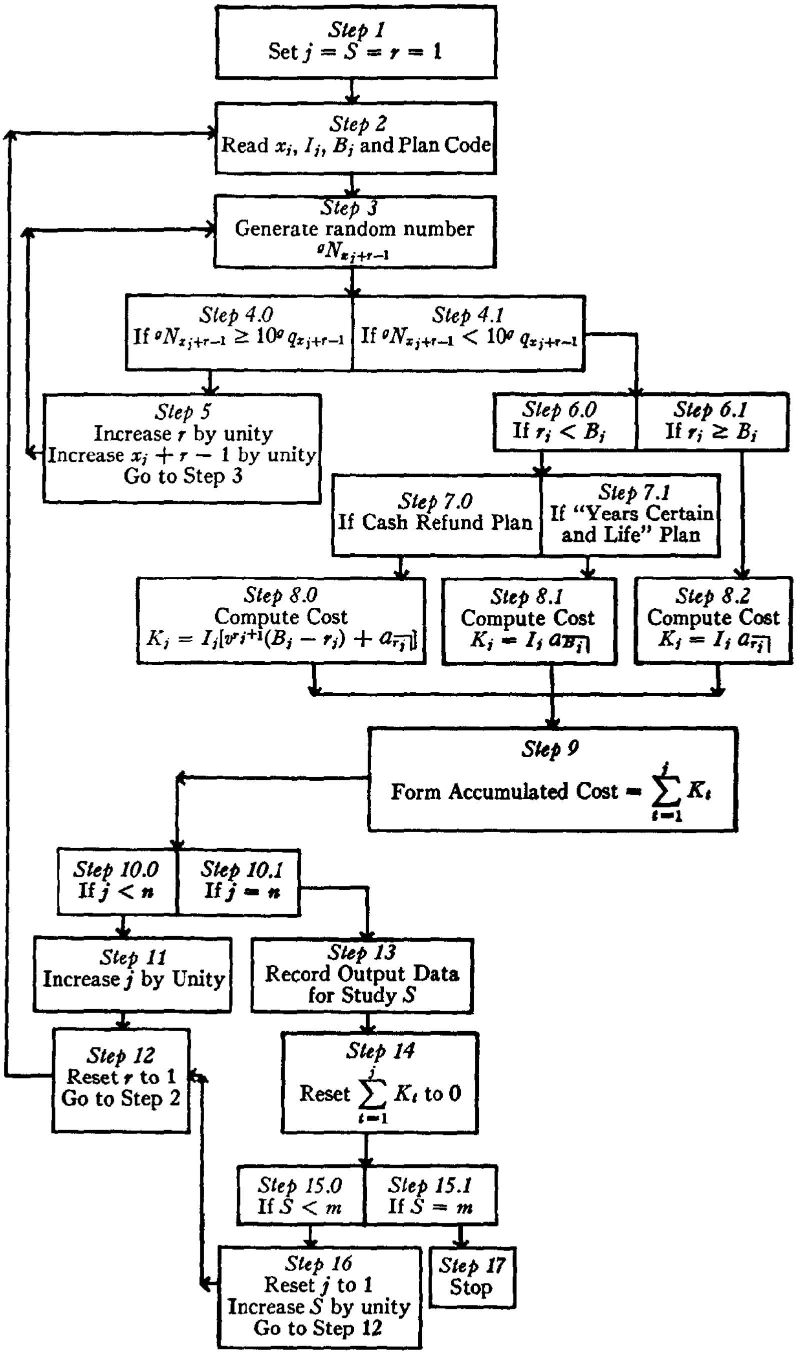
Chart I, which of necessity eliminates some details, will suggest to the interested reader how an electronic data processing machine may be instructed to perform the necessary steps required to produce m different mortality studies and cost computations for the entire group of n lives. Reference should be made to Appendix A, which defines the various symbols used in Chart I.

An experiment designed to test the practicality of the concept with respect to the life annuity with no refund plan for a group of 10 lives was performed on an IBM No. 650 machine. It was assumed that the entry age was 65 for each life, that the income was unity for the first 9 numbered lives within the group, and that the income for the 10th life varied for each of the six trials conducted for the group as a whole. One hundred mortality studies for the group as a whole were performed in each of six trials. Table 2 presents certain data obtained in this experiment with respect to each of the six trials for the group as a whole, each trial requiring 1,000 individual life studies.

The 6,000 individual mortality studies performed for the 6 trials resulted in an average annuity value of 11.54 which compares closely with the expected tabular value of a_{65} = 11.50. It is to be observed that as I_{10} increased in successive trials from 1 to 50, the minimum unit cost encountered gradually diminished while the maximum cost, in general, increased.

CHART I

FLOW CHART FOR COMMON ANNUITY PLANS



Column (6) shows the approximate values for the confidence limit of 95%. In the case of trial No. 1, a contingency fund of 2.57 per unit payment (14.07 − 11.50) would be required if one desired to be able to establish an approximate theoretical 95% chance of meeting all the payments due with respect to the selected group of 10 annuitants. When we observe the entry in column (6) for trial No. 6, we notice that the theoretical contingency fund would have to be materially increased approximately to 6.88 per unit payment. Column (5) similarly gives approximate values for confidence limits of 90%.

TABLE 2

SIX TRIALS OF 100 MORTALITY STUDIES FOR A GROUP OF 10 LIVES

$n = 10$ $B_j = 0$ for $j = 1$ to 10 inclusive
$m = 100$ $I_j = 1$ for $j = 1$ to 9 inclusive
I_{10} is a variable, per column (1)

Trial Number	Annual Income I_{10} Payable to 10th Life (1)	Unit Annuity Cost			Approximate Unit Annuity Cost for Confidence Limit	
		Average (2)	Maximum (3)	Minimum (4)	90% (5)	95% (6)
1.	1	11.61	15.21	6.68	13.61	14.07
2.	2	11.11	15.36	6.65	13.49	14.37
3.	5	11.28	16.21	5.36	14.56	15.01
4.	10	11.52	17.32	4.74	15.61	16.03
5.	25	11.97	20.77	3.23	16.67	17.63
6.	50	11.74	20.06	2.49	17.28	18.38
Average	....	11.54	....	....	....	....

Each of the six trial runs of 100 mortality studies for 10 lives required about 45 minutes on the IBM No. 650 machine. The author believes that the time required for making each trial run could be considerably reduced by applying more refined techniques in programming the operation.

A concluding experiment, which was designed primarily to estimate the time required to conduct a single mortality study for a group of 100 lives, indicated that one such study would require approximately 5 minutes of machine time. As far as the machine is concerned, the time required for a single study for a group of 100 lives is the same as that required for 100 studies for one life.

CONCLUSION

The foregoing experiments, designed to test the practicality of using a series of random numbers to make a hypothetical cost computation, have, I believe, indicated that new techniques for many types of problems are made available to actuaries by the power of new electronic machines. These tests immediately suggest that frequency distributions of costs can be obtained not only for the analysis of closed groups of annuities but also for various classes of insurances. Limitations on the number of lives to be studied and the number of variables to be considered will depend on the type of available equipment, the ingenuity of programmers, and the expense of performing the tests in terms of time and money.

The writer wishes to caution the reader that the technique described herein will not produce sound results unless random and not pseudo-random numbers are used in the comparisons. He also wishes to state an assumption made throughout that no correlation in the rates of death exists among the members of any group. Modifications in the programming suggested by Chart I may, of course, be made to place controls on input and output data or to account for other conditions possibly present in a stated problem, such as the effect of select mortality rates.

Finally, the writer wishes to express his indebtedness to Mr. Ralph E. Traber, F.S.A., who transmitted the germ of the idea which inspired the experiments described herein, and also to Mr. George E. Wallace, A.S.A., for his technical aid.

APPENDIX A

Definitions of Symbols

n = number of lives in class of annuitants

m = number of mortality studies to be performed for the class of n lives

S = numbered study within the total number of m studies which are to be made

j = numbered life within the class of n lives

g = number of decimal places to which q_x is recorded

ω = limiting age in mortality table; *i.e.*, $q_{\omega-1} = 1$

x_j = age at entry of life j

I_j = annual income payable to life j

B_j = number of years certain at issue in the case of a life annuity plan providing for a certain period; or

= number of years following issue during which a decreasing death benefit is payable in the case of a cash refund annuity plan under which the decrease in death benefit each year is I_j; or

= 0 in the case of a life annuity with no refund

r = duration at which the trial is being made to see if the annuitant is to be assumed to die

r_j = duration from date of entry to the beginning of year during which life j is assumed to die

$x_j + r - 1 = z$ = attained age of annuitant at beginning of contract year r

${}^{g}N_{x_j+r-1}$ = random number composed of g digits generated for comparison with $10^g\, q_{x_j+r-1}$

K_j = present value of annuity payments and death benefits for life j

$= I_j\, a_{\overline{r_j}|}$ in event no death benefit is due at end of $r_j + 1$ years; or

$= I_j\, [v^{r_j+1}(B_j - r_j) + a_{\overline{r_j}|}]$ in event the cash refund period has not expired at date of death; or

$= I_j\, a_{\overline{B_j}|}$ in event an annuity certain is payable after the date of death

$\sum_{t=1}^{j} K_t$ = accumulated cost for the first j lives

APPENDIX B

Generation of Random Numbers

In the described experiments the IBM No. 650 machine was programmed to produce a series of "random" numbers. The quote indication ("random") is used because no proof can be given at this time that the numbers produced by the method do in fact represent a series of random numbers. The following outlines the steps of the method used to generate a series of "random" numbers for the case where q_x is recorded to five decimal places:

1. Select a ten digit number greater than 10^9 from an arbitrary source such as the central ten digits of an entry in a column of interest functions having fourteen digits. Call this number ${}^{10}N_t$ and commence the following cycle.
2. Square ${}^{10}N_t$ to form a twenty digit number. Without changing the sequence of digits, extract the 11th to 20th high order digits (*i.e.*, the first ten digits) of $({}^{10}N_t)^2$ to form a new ten digit number, which we will call ${}^{10}N_{t+1}$, of which the highest order digit (*i.e.*, the first) may be 0.
3. Similarly, extract the 6th to 10th high order digits of $({}^{10}N_t)^2$ to form a five digit number, which we will call ${}^{5}N_t$, to become the first of the desired series of random numbers for direct comparison with $10^5\, q_x$.
4. Examine the magnitude of ${}^{10}N_{t+1}$ to test for one of three conditions: (1) if ${}^{10}N_{t+1} > 10^9$, repeat the cycle, generating in the process the next desired five digit random number; (2) if ${}^{10}N_{t+1} < 10^9$, add $9 \cdot 10^9$ to ${}^{10}N_{t+1}$ to produce a

revised value of ${}^{10}N_{t+1}$ for continuing the cycle; (3) if ${}^{10}N_{t+1} = 10^9$, select a new ten digit number from the source indicated in step 1 as the revised value of ${}^{10}N_{t+1}$ for continuing the cycle.

The reader may possibly speculate as to the degree of bias which step 4(2) above introduces into the operation by the relatively frequent addition of $9 \cdot 10^9$. A possible alternative approach could be attempted by performing a cyclical permutation of digits when ${}^{10}N_{t+1} < 10^9$; *i.e.*, extract the high order digit, place it in the low order position and let the digit in the old 9th order position take the new 10th order position, etc.—repeat until a ten digit number is formed which has a nonzero high order digit.

A more satisfying method of generating random numbers could perhaps be devised by programming the machine to use combinations of a commonly accepted table of random digits such as the one published in Volume 24 of *Tracts for Computers* (University of London, University College). The sequences of the digits there published have been subjected to rigorous tests to determine if their choice has been biased. The results of these tests have generally been accepted as proof that they may be deemed satisfactory for most problems which require the selection of random numbers.