RECORD OF SOCIETY OF ACTUARIES
1986 VOL. 12 NO. 4A

SIMULATION METHODOLOGY FOR ACTUARIES

Moderator: ARNOLD F. SHAPIRO
Panelists: EDWARD W. FREES
FARROKH GUIAHI
AARON TENENBEIN
Recorder: CHARLES S. FUHRER

- The underlying theme of this session is the use of simulation, or Monte Carlo, techniques to analyze complex insurance models. Areas to be addressed include:
  - -- Estimating reserves
  - -- Solvency considerations
  - -- Estimation of mortality
  - -- Distribution of aggregate losses

MR. ARNOLD F. SHAPIRO: We have three speakers who will be dealing with simulation from different perspectives. Mr. Farrokh Guiahi is from Hofstra University, and he will be talking about computer simulation with a specific package. Mr. Edward Frees from Wisconsin will be talking about working backwards from the probability of ruin to the initial reserve, reversing in a way what we generally do. Mr. Aaron Tenenbein from New York University will be talking about multivariate distributions and simulations involving those.

MR. EDWARD (JED) W. FREES: I would like to talk about a little trick I discovered for helping me to capture very complicated simulation models in a much easier way. The formal title of the paper I wrote on this subject is "Approximation of the Initial Reserve for Known Ruin Probabilities." It has appeared in the July 1986 edition (Volume 5, Number 3) of *Insurance Mathematics and Economics*. It is a rather technical paper to read, so I will give an

overview of some of the ideas behind it. I am going to remind you about a problem that is particularly actuarial in nature -- the ruin problem. Everybody has seen it on the Part 5 syllabus. It's one of those few problems that the Society is not willing to make an elective at this point. I would like to use a particular kind of methodology entitled "Stochastic Approximation." Now I am not assuming that anyone is familiar with it, and that would probably include the panelists as well, but I would like to give an overview. It is the kind of techniques that you have seen before. Another example is the compound Poisson, not because that is is a good approximation to reality, but just because it is something we have seen at some points in our lives.

Historically, we're used to thinking of modeling an insurance enterprise based on a generic net liability model. It might start with a particular known value and go out over time t. One of the nice things that has been going on in the last 10 years in the literature is that we used to work with particularly simple models like compound Poisson and then maybe enhance them a little. Beard, Pentikäinen and Pesonen [1] wrote about putting something in that model to handle autocorrelation and so they are getting a little bit closer. In particular, some of the English actuaries now are working very hard on trying to incorporate the asset side into the model. I am going to be assuming that we have a generic net liability process, simply claims minus assets or claims minus premiums minus investment income -- anything you can put on a computer. Remember we have a lot of ideas of what should be going on in the real world, but we could never analyze them. We can only analyze very simple things using probabilistic techniques.

When you try to summarize a really hard, hard problem like modeling an insurance enterprise, there are a lot of different ways to summarize the output. One particular way that people are used to talking about, is this notion of a probability of ruin. It is not that the enterprise actually goes bust or anything, but the probability of ruin is a kind of indicator. That's when we think that this risk bearing enterprise is in trouble, and perhaps a solvency guarantee association would want to take over at that point. The probability of ruin is to find when ruin would begin. It is defined to be the first time or the event that the u of t exceeds a particular value, and I'm going to call that particular value the initial reserve, the initial surplus,

the start of capital, or however you would like to think of it. In particular, a lot of the literature has looked at not a particular 10 year or 20 year period, but at an infinite horizon time. Again, that has been for mathematical convenience and not a good approximation to reality. So, I will be thinking of T as a horizon time. We are just looking for the event that at some point this process exceeds this u which is fixed. It is hard to calculate this probability of ruin. It has not been used much in North America just because it is so difficult to calculate.

Another way of restating the problem that I'd like to do for simulation purposes is to talk about not the whole liability process but just to look at the biggest point of process. If we look at the largest value of the process, the probability that that's exceeding the number u, it is the same thing as looking at the entire process. I'll call that event z. That is the largest value of u(t) or another way of describing it is the supremum up to that particular time point. One particular kind of summary measure that people are used to looking at is the odds of or the probability of that event occurring. Now that is not the only kind of summary one would look at, and indeed you can look at a lot of other ones besides that. It is useful though, because European actuaries have a good interpretation, or good idea of what this notion means. It has, at least sometimes, been used in regulation. Again, one of the reasons it hasn't been used is it is very hard to calculate in general.

I would like to pose perhaps another interesting but very different mathematical problem, which is to do the calculation for a fixed probability of ruin that we have some idea should be, perhaps 1% or 1/10%. Then, what height or how much initial surplus should be required so that we achieve this probability, alpha? I am not going to be a statistician at the moment. I am going to assume that in a sense I am a probabilist in that I know something about the distribution function of z. That is, I'm sitting here on my micro-computer, and I can say whether this is going to be a compound Poisson or any one of several types of processes, so we can set that. Now we only set that up stochastically, so that is known.

Now I'd like to remind you about some techniques in numerical analysis. Recalling from numerical analysis, if you have a function of a certain height,

g(x), you would like to find that value theta so that g(theta) is a certain height. In particular, we usually use alpha as equal to zero, but since we are assuming that this is a known constant, it really doesn't matter. This is the same kind of problem, at least from numerical analysis, as finding a minimum or maximum, because, if g is known, as we're assuming it to be, you can just take a first derivative and then go for minimums and maximums.

Here it is a minimization problem. For most problems that even approximate the real world, g is a very complicated thing. It doesn't lend itself to going back and looking at an inverse type of thing, but we can solve it numerically though iteration -- a Newton-Raphson procedure. You take your current best guess and evaluate it at $g(x_n)$. So if you are a little high, then you go down, and if you're a little low, then you go up. It turns out that the derivative value at that point is the right scaling factor to make the amount of change good. Many times this derivative is very difficult to calculate, so you just put in something -- a numerical equivalent or anything -- just to make this algorithm go. All you need is several positive constants that only need to be decreasing. This idea is important.

I want to apply g to this probability calculation, but one of the hard things, and indeed in the actuarial literature it has been a hard problem for the last 30-35 years, is calculating probabilities of ruin. That is why they are not used very much. We can approach this kind of problem where g is only observed with error. For example, the statistical literature describes a situation in which someone might be interested in saying that so much level of a drug is administered and the response that you see to that drug, for instance blood pressure, will not be exactly a certain height but around that height, plus or minus. So it is the same kind of problem except g can only be observed with error. With another kind of problem you might think about applying certain levels of fertilizer to the land. Instead of a minimum for that problem you want to maximize that problem, that is, you would like to maximize that amount of crop yield.

The g that we will be using, though, in this particular kind of approximation procedure is the following: All I would like to do now is find that particular level of surplus so that this function is equal to zero. That is

the equivalent kind of the problem. Now if you just stopped at this, you could apply a Newton-Raphson type argument, iterating and then using this kind of g function. The function g which represents these probabilities is very hard to calculate. But what's not hard is to put an approximation of the function in there. The notation that I am using is an indicator function (I for indicator). If this event occurs, then it is equal to 1, and it is equal to 0, otherwise. What g_n is meant to be is a proxy for g. It's very easy to do, and in the example we'll be thinking about this z as a computer realized event. Either it will exceed u or it won't exceed u. It is wrong because it is just an approximation, but if we do it long enough, then indeed it will work just like a Newton-Raphson kind of argument. But that's a little foreshadowing. I am calling theta to be that best surplus level, and we would like to find that number theta so that the probability of this maximum height, random variable, exceeding that variability, is equal to the prespecified alpha. That's the equivalent to find g -- g(theta) being equal to 0.

That's really kind of the key idea behind this whole algorithm. That is what I am calling $theta_n$'s. They're meant to be approximations of theta that work just like the x_n's for the Newton-Raphson. Because of the minus signs, it works out that, instead of a positive sequence of decreasing constants, we want to use f times f_n, so it is more like a steepest descents method rather than a Newton-Raphson. $Theta_1$ is the initial guess to the algorithm. The z_i's are computer simulated realizations of the u(t) process. We have g divided by g prime which works out that, when you do some of the calculations, that we want to put in a proxy for g prime. If you take a probability and take a derivative of that, that's a density. This f_n is an estimate of this density of the random variable z at theta.

When you write a technical paper, you have an algorithm and you think it works, but usually you have to go through a few steps of justification and those justification steps come in two layers. First, you have to do an analysis that says at least theoretically, if you do something long enough, it will work. That's what this first result says, that is, if you do this thing long enough, the limit of these approximations will be the right value. You need something like that.

A second level is you would like to have an example in the results in Section 4 of the Appendix to show people that not only does this thing work in the limit, but it also works within an hour or two for most problems. Also what is nice for these kinds of arguments is you give not your best guess theta $_n$ at any particular stage, but you can also give an interval estimate. That is, you can give a range for how well you think you are doing and indeed this is a fairly standard confidence interval. A confidence interval will be my best guess theta$_n$ and down 1.96, there are some central limit theorems coming in here (alpha$_1$ minus alpha). There is the usual square root of n, and the f $_n$ is tending towards a constant. These are justifications of the procedure. (See Appendix.)

I tried this on a couple of different things I was interested in. Gerber [2] introduced the idea of autoregressive moving average models. He was interested in actuaries. They work with problems that often don't correspond to reality, and correlation in data is one problem that a lot of times people say, "Of course claims are correlated. How could they not be between financial periods?" So he introduced the autoregressive model to overcome that. Another nice paper by Waters and his student Paptriandafylou [3] introduced the idea of stochastic discounting where interest rates are random variables. That's a rather important model that we don't see very much on the Society syllabus right now. Typically, we've been letting forces of decrements be random variables, but keeping interest rates deterministic. This is one step in that direction. The example I'm going to talk about now, not because it is interesting or because it is practically useful, but because everybody has seen it at some point in his life, is the compound Poisson model. I'll be taking the simplest possible compound Poisson model, that is one where claims are distributed exponentially. So in thinking about the number of claims n $_t$ being the Poisson process driven by something, the amount of claims are exponential and of a constant premium force. For this example, we've assumed that we are in fairly stable interest periods, and we're not explicitly worrying about interest right now. Here's a typical realization in Appendix Section 5, that is it goes down for a while and then all of a sudden we get a claim and it jumps up by that amount, etc. One reason is that this is one of the few models you can actually do calculations on, and you can actually get infinite time horizon calculations. We know that the probability of this process exceeding 21.91 is about 1%. It's known in this case, one of the few cases where it is

known. The types of things included in the tables from my paper are time horizons, known ruin probabilities, approximation of the best reserve, and density estimates (not that they are useful for this, but it gave me a lot of insight into it and so I thought it might help other people as well.) These are upper and lower bounds of the confidence intervals. As the probability of ruin increases, the amount of reserve required decreases. As the time horizon increases, the amount of reserve increases; there is more chance for ruin. The upper bounds by definition are centered about the best guess theta. It turns out that for this kind of procedure, just like Newton-Raphson, it very much matters where you start the procedure theta $_1$ and I've given some recommendations in the paper. In particular, quantal estimation was used. It's not too hard to estimate densities, if you know the right statistical literature to look at. Basically, we are estimating densities by slicker versions of a histogram, but I've also given some references in the paper.

References

1. Beard, R., Pentikäinen, T., and Pesonen, E. *Risk Theory*, London: Chapman and Hall, 1984.

2. Gerber, H. "Ruin Theory in the Linear Model," *Insurance Mathematics and Economics*, 1 (1982), pages 177-84.

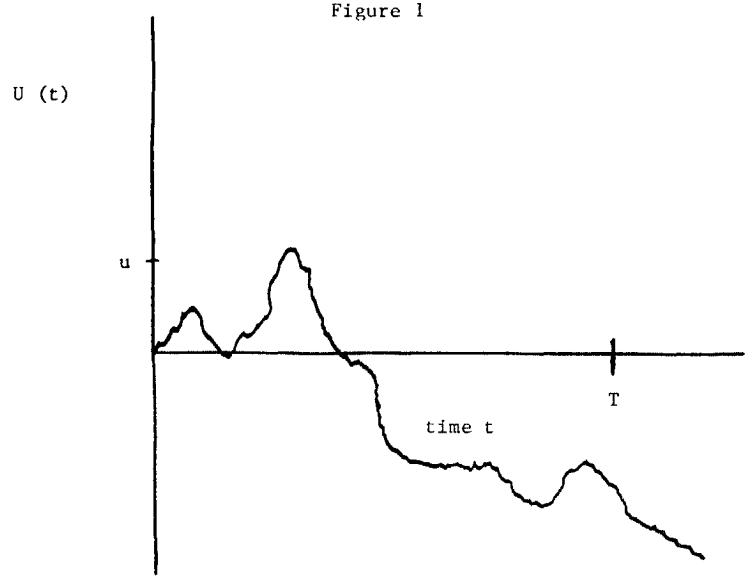
3. Papatriandafylou, A., and Waters, H. "Martingales in Life Insurance," *Scandinavian Actuarial Journal*, (1984), pages 210-30.

APPENDIX

Section 1. Ruin Problem

Define U(t) to be the net liability at time t. We can think of U(t) as the cumulative excess of claims over assets by time t.

Figure 1



We define

P(ruin) = P(U(t) exceeds u at some time $t \leq T$)

where T is the horizon time. A useful auxiliary random variable is

Z = largest value of U(t) over $t \in [0,T]$

$= \sup_{t \leq T} U(t)$,

Here, sup stands for the supremum over a set.

Statement of the Problem

Find the initial surplus 'u' so that P(ruin) = P(Z > u). Here, the probability of ruin, P(ruin) = α, is considered known as is the stochastic structure, symbolized by 'P'.

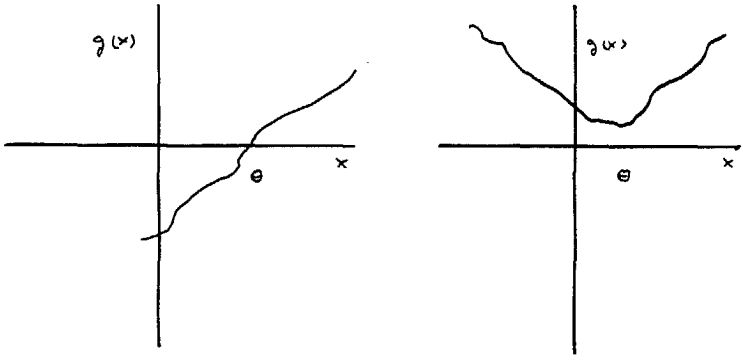
Section 2. Stochastic Approximation

Recall a problem in numerical analysis. Find θ so that

$g(\theta) = \alpha$ = known constant.

Figure 2

Two types of problems in numerical analysis



root-finding function minimization

Two standard methods for solving this problem are:

(1) Newton-Raphson:

$$X_{n+1} = X_n - g(X_n)/g'(X_n)$$

(2) Method of Steepest Descent:

$$X_{n+1} = X_n - a_n\, g(X_n)$$

where $\{a_n\}$ is a decreasing sequence of positive constants. Suppose $g(\cdot)$ can only be observed with "error."

Examples

1) A patient is administered a dosage X of a drug. We would like the response (e.g., blood pressure) to be kept at a certain level.

2) X lbs. of fertilizer is applied to a plot of land. Would like to find the amount θ to maximize the crop yield.

Section 3. The Approximation Procedure

We wish to calculate the function

$$g(x) = \alpha - P(Z > x).$$

This is hard to calculate, but an unbiased estimator

$$g_n = \alpha - I(Z_n > x)$$

is easy to calculate. Here I is the indicator function of a set, that is, $I(Z > x) = 1$ if $Z > x$ and $= 0$, otherwise. Also, $\{Z_i\}$ is an i.i.d. computer simulated sequence of realizations of the largest value of $U(t)$.

Let θ be the desired surplus level, defined by

$$P(Z > \theta) = \alpha$$

which is equivalent to $g(\theta) = 0$. The approximation of θ is a stochastic version of those defined in Section 2. At the n^{th} stage, θ_{n+1} is recursively defined by

Section 4.

$$\theta_{n+1} = \theta_n + n^{-1}\{I(Z_n > \theta_n) - \alpha\}/f_n.$$

Here θ_1 is an arbitrary initial value and f_n is an estimate of the density of Z at θ.

Results of the Procedure:

1. $\lim_{n\to\infty} \theta_n = \theta$ with probability one.

2. $\theta_n \pm 1.96\{\alpha(1-\alpha)/(n f_n)\}^{\frac{1}{2}}$ is an approximate 95% confidence interval for θ.

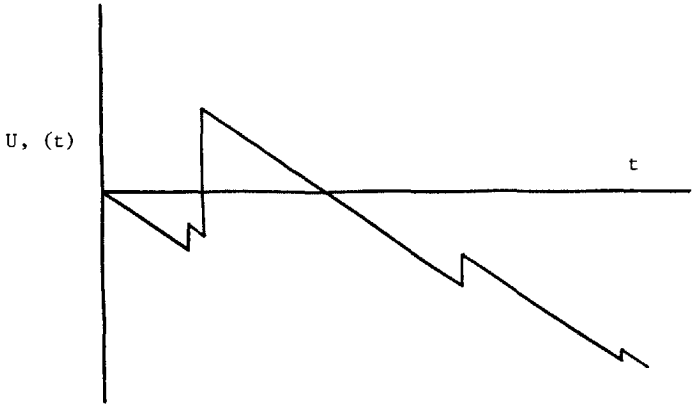
Section 5. Example - Compound Poisson

Let $\{N(t)\}$ be a Poisson process independent of $\{X_k\}$, an i.i.d. sequence of exponential random variables. Define π to be a fixed premium per unit time. Define the associated Compound Poisson process

$$U_1(t) = \sum_{k=1}^{N(t)} X_k - \pi t$$

Figure 3.

Poisson Exponential Liability Process



As a reference point, it is well-known that in this special case that

$$P(\text{ruin}) = P(\sup_{t>0} U_1(t) > 21.91) = .01.$$

MR. FARROKH GUIAHI:

SIMULATING AGGREGATE CLAIMS

Introduction

Determining the distribution of aggregate claims of an insurance portfolio is a subject of immense interest to actuaries. The distribution of aggregate claims may be used to price insurance products, particularly pricing excess-of-loss reinsurance treaties. The fluctuation of aggregate claims over a time horizon is related to the ruin problem.

A large body of literature exists on the subject of aggregate claims. For a recent account, refer to Panjer and Willmot (1984). The focus of that paper (and many other papers cited in its references) is to provide analytic expressions for the distribution of aggregate claims and its moments. The distribution of aggregate claims involves computation of an infinite series where the n-th term in the series is the n-fold convolution of the severity (claim amount) distribution weighted by the probability of having n claims. To evaluate the series, in the special cases where the n-fold distribution of severity can be worked out, one has to resort to numeric methods. Expressions for the moments of the aggregate claims, in terms of the moments of claim count and moments of severity (claim amount), can be relatively easily derived.

This paper emphasizes the simulation approach, which requires generating many aggregate claim values by recourse to simulation. The distribution of aggregate claims is approximated by the empirical (simulated) distribution based on simulated aggregate claim values, provided the size of the simulated sample is sufficiently large. By using a simulation language, one can dispense with programming efforts and concentrate on the modeling and sensitivity aspects of the problem. With current trends toward using software packages as a vehicle for a Decision Support System, GPSS (General Purpose Simulation System), a popular simulation language, was chosen to achieve this end.

In the next section is a brief overview of GPSS. The interested reader should refer to Schriber (1974) or Gordon (1975) for a comprehensive treatment of the subject.

An Overview of GPSS

GPSS is a language used to present models for many systems as well as being a computer program (i.e., a processor). It is particularly suited for simulating queueing systems. Modeling in GPSS means constructing a block diagram for the system to be simulated. The blocks are selected from a set of some 40 predefined block types, which are connected together by directed lines. The inclusion of a given block depends on the logical structure of the problem to be simulated. For each block, one has to provide block information, which consists of Location, Operation, and Operands. Location (optional) is the name given to a block, used when a reference is made to that block from another block. Operation indicates the type of action to be performed by the block. Activation of a block corresponds to the occurrence of events to be simulated by the system. Operands provide the data for the action to be executed by the block.

In GPSS, dynamic entities are called transactions or units of traffic. Transactions are created, moved in the model (i.e., through the blocks), and eventually removed. Initially there are no transactions in the model; at later times transactions are brought into the model and moved through the blocks. As a transaction moves into a block, it causes the occurrence of the event to be simulated. The GPSS processor schedules the future movement of transactions, one at a time, in the model. It performs the actions associated with entries of a transaction into a block, and collects as well as updates statistics of interest to the model builder.

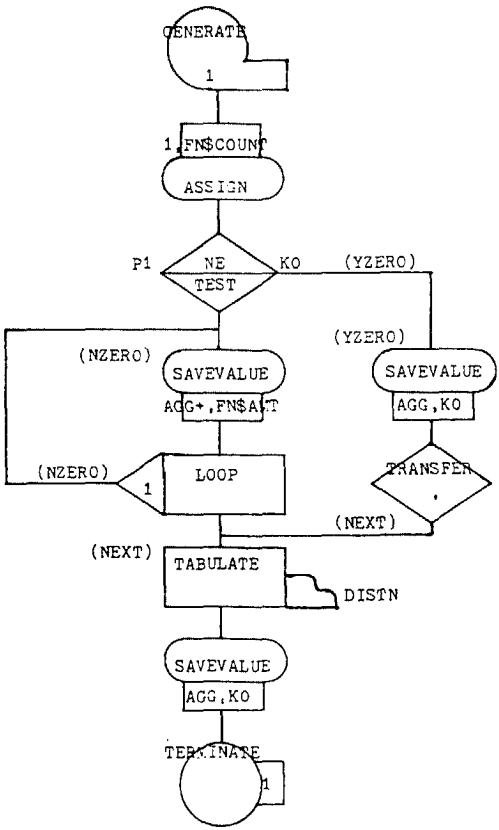
The static entities in GPSS are called equipment. The equipment represents the entities that are permanent during the simulation period. In simulating a one-line, one-server toll station, cars would be the transactions, and the ticket attendant would be the equipment (called a Facility in GPSS terminology). By analyzing simulation outputs, one obtains an understanding of behavior of the real system to be simulated. By changing the input parameters to the model, one can experiment with the system.

The GPSS Model

A block diagram of the GPSS model to simulate distribution of aggregate claims is presented in Exhibit 1. The input specifications are (i) claim count distribution, (ii) severity (claim amount) distribution, (iii) format for the frequency table of simulated aggregate claims, and (iv) number of aggregate claims to be simulated.

EXHIBIT 1

GPSS Block Diagram for Simulating Aggregate Claims



A transaction, here a counter, is created in every unit of time through the GENERATE block. This transaction has a numeric characteristic called Parameter 1. After entering the model, the transaction moves immediately from the GENERATE block to the ASSIGN block, where a value to its Parameter 1, P1, is assigned according to claim count distribution. At the ASSIGN block, the value assigned to P1 is chosen according to the function FN$COUNT. P1 denotes the number of claims. After the number of claims have been determined for the transaction, it moves to a TEST block. At the TEST block, P1 is tested for value of zero. If the claim count is zero, the transaction moves to the nonsequential block YZERO. YZERO is a SAVEVALUE block; its purpose is to store a value in the user-defined SAVEVALUE variable AGG (aggregate claims), the value of zero corresponding to the case of no claims. Then the transaction moves to the unconditional TRANSFER block, which directs it to block NEXT. If the claim count is not equal to zero, the transaction moves to the next sequential block NZERO, which is a SAVEVALUE block. Here a value from the severity distribution is added to the current value of SAVEVALUE variable AGG, *initially zero. A claim amount is determined that is added to the* existing value of the variable AGG. The claim amounts are generated by the function FN$AMT.

The transaction next moves to a LOOP block. At the LOOP block, the current value of P1 (number of claims) is decremented by one; then it is tested for value of zero. If P1 (current value) is zero, the transaction moves immediately to the next sequential block NEXT. NEXT is a TABULATE block, whose purpose is to record a computed value of AGG for the frequency table associated with aggregate claims. After leaving the TABULATE block, the transaction enters another SAVEVALUE block in order to reinitialize the value of variable AGG to zero for the next transaction entering the model. Finally, the transaction moves to the TERMINATE block, where it is removed from the model.

If the P1 (current value) is not zero, the transaction moves immediately to block NZERO. At block NZERO, another claim amount is generated and added to the existing value of SAVEVALUE variable AGG. Then the transaction moves to the LOOP block and proceeds from there. After looping a number of times, the

value of P1 is eventually decremented to zero, causing the transaction to exit the LOOP block for the next sequential block.

Applications

As a first illustration, we consider the example provided by Wooddy (1973). Here, the claim count distribution is Poisson with mean 2, and the severity values are 1, 2, 5, 7 and 9 with corresponding probabilities .15, .30, .10, .25, and .20. Exhibit 2a gives GPSS statements for Wooddy's example. Columns 1 and 4 of Exhibit 2b give the simulated values of the cumulative distribution function of aggregate claims. There is close agreement, using 3 decimal places, between the simulated cumulative distribution function values of Exhibit 2b, column 4, "Cumulative Percentage," and the theoretical values (using convolutions) derived by Wooddy (1973), on page 33, Table 5, column (9). The GPSS output also provides sample (simulated) values for the mean and variance of aggregate claims. The exact values of means and standard deviations are 9.6 and 8.037. These values correspond closely to the simulated values 9.632 and 8.062, respectively.

The second example considers the case where claim count is Poisson with mean 50, and the severity is Lognormal with parameters mu = 8.91, and σ = 1.78. This severity distribution was fitted to medical malpractice closed claim data; see Miccolis (1981). In this case, working out the theoretical distribution of aggregate claims would be extremely difficult as compared to the Wooddy example. Exhibit 3a gives the GPSS statements for the program, and Exhibit 3b provides the simulated cumulative function for the aggregate claims in this case. To estimate the probability that the aggregate claims in this case are less than or equal to 2,000,000, see the lower table of Exhibit 3b to find the value .713. Exhibit 3b has two output tables. The upper table provides the estimate of distribution of aggregate claims at specified values in the first column for a simulated sample of 1,500. The lower table gives the same estimates based on a simulated sample of 2,000. One's confidence in estimates of aggregate claims distribution is sharpened by considering different sample sizes.

EXHIBIT 2a

GPSS Statements For The Wooddy Example

```
*       SIMULATED AGGREGATE DISTRIBUTION USING GPSS
*         1973 SOA PART 5 STUDY NOTES BY JOHN C. WOODDY
*         SECTION 3.3, PAGE 31
*
*       GPSS INPUT REQUIREMENTS :
*         (1) SPECIFICATION OF CLAIM COUNT DISTRIBUTION
*         (2) SPECIFICATION OF CLAIM AMOUNT DISTRIBUTION
*         (3) PARAMETERS FOR THE AGGREGATE TABLE
*         (4) SIMULATED SAMPLE SIZE
*
        SIMULATE
*
 COUNT  FUNCTION    RN1,D10                CLAIM COUNT DISTRIBUTION
0.1353,0/.4060,1/.6767,2/.8571,3/.9473
0.9834,5/.9954,6/.9988,7/.9997,8/.9999,9
*
 AMT    FUNCTION    RN1,D5                 CLAIM AMOUNT ( SEVERITY ) DISTRIBUTION
0.15,1/.45,2/.55,5/.80,7/1,9
*
*
        GENERATE    1                      CREATE A COUNTER
        ASSIGN      1,FN$COUNT             DETERMINE THE NUMBER OF CLAIMS
        TEST NE     P1,K0,YZERO            TEST IF THE NO. OF CLAIMS IS NON-ZERO
 NZERO  SAVEVALUE   AGG+,FN$AMT            ACCUMULATE CLAIM AMOUNTS
        LOOP        1,NZERO                PROCEED WITH THE LOOP
 NEXT   TABULATE    DISTN                  TABULATE AGGREGATE CLAIM
        SAVEVALUE   AGG,K0                 REINITIALIZE AGGREGATE CLAIM AMOUNT
        TERMINATE   1                      REMOVE THE COUNTER
*
 YZERO  SAVEVALUE   AGG,K0                 IF CLAIM NO. IS 0 THEN AGG CLAIM IS 0
        TRANSFER    ,NEXT                  CONTINUE
*
*
 DISTN  TABLE       X$AGG,0,1,60           TABLE SPECIFICATION FOR AGGREGATE CLAIMS
*
        START       10000                  NUMBER OF SIMULATED AGGREGATE CLAIMS
        END
```

EXHIBIT 2b

Descriptive Simulated Statistics for the Wooddy Example

TABLE DISTN
ENTRIES IN TABLE 10000 MEAN ARGUMENT 9.632 STANDARD DEVIATION 8.062

UPPER LIMIT	OBSERVED FREQUENCY	PER CENT OF TOTAL	CUMULATIVE PERCENTAGE	CUMULATIVE REMAINDER
0	1338	13.37	13.3	86.6
1	421	4.20	17.5	82.4
2	894	8.93	26.5	73.4
3	251	2.50	29.0	70.9
4	254	2.53	31.5	68.4
5	358	3.57	35.1	64.8
6	156	1.55	36.7	63.2
7	822	8.21	44.9	55.0
8	248	2.47	47.4	52.5
9	1029	10.28	57.7	42.2
10	330	3.29	61.0	38.9
11	527	5.26	66.2	33.7
12	289	2.88	69.1	30.8
13	205	2.04	71.2	28.7
14	394	3.93	75.1	24.8
15	145	1.44	76.6	23.3
16	508	5.07	81.6	18.3
17	157	1.56	83.2	16.7
18	365	3.64	86.9	13.0
19	147	1.46	88.3	11.6
20	179	1.78	90.1	9.8
21	132	1.31	91.4	8.5
22	78	0.77	92.2	7.7
23	172	1.71	93.9	6.0
24	51	0.50	94.4	5.5
25	135	1.34	95.8	4.1
26	62	0.61	96.4	3.5
27	55	0.54	97.0	2.9
28	51	0.50	97.5	2.4
29	31	0.30	97.8	2.1
30	35	0.34	98.1	1.8
31	20	0.19	98.3	1.6
32	34	0.33	98.7	1.2
33	10	0.09	98.8	1.1
34	30	0.29	99.1	0.8
35	14	0.13	99.2	0.7
36	16	0.15	99.4	0.5
37	12	0.11	99.5	0.4
38	8	0.07	99.6	0.3
39	10	0.09	99.7	0.2
40	4	0.03	99.7	0.2
41	6	0.05	99.8	0.1
42	2	0.01	99.8	0.1
43	4	0.03	99.8	0.1
44	2	0.01	99.9	0.0
45	2	0.01	99.9	0.0
46	1	0.00	99.9	0.0
47	1	0.00	99.9	0.0
48	2	0.01	99.9	0.0
49	2	0.01	99.9	0.0
50	0	0.00	99.9	0.0
51	1	0.00	100.0	0.0

REMAINING FREQUENCIES ARE ALL ZEROS

EXHIBIT 3a

GPSS Statements For The Second Example

Claim Count Distribution : Poisson with mean 50
Severity Distribution : Lognormal with parameters μ = 8.91 and σ = 1.78

```
*       SIMULATED AGGREGATE DISTRIBUTION USING GPSS
*
*       GPSS INPUT REQUIREMENTS :
*         (1) SPECIFICATION OF CLAIM COUNT DISTRIBUTION
*         (2) SPECIFICATION OF CLAIM AMOUNT DISTRIBUTION
*         (3) PARAMETERS FOR THE AGGREGATE TABLE
*         (4) SIMULATED SAMPLE SIZE
*
        SIMULATE
*
 POI50 FUNCTION   RN1,C13              CLAIM COUNT DISTRIBUTION
0,0/.0125,34/.0881,40/.1144,41/.182,43/.3721,47/.6986,53/.7873,55
0.858,57/.9309,60/.9796,64/.9972,68/1,70
*
 LNORM FUNCTION   RN1,C27              CLAIM AMOUNT ( SEVERITY ) DISTRIBUTION
0,14.59/.0013,35.52/.0035,55.42/.0062,86.49/.0122,134.96/.0228,210.61
0.0401,328.65/.0668,512.86/.1056,800.31/.1587,1248.88/.2266,1948.86
0.3085,3041.18/.4013,4745.72/.5,7405.66/.5987,11556.45/.6915,18033.73
0.7734,28141.44/.8413,43914.48/.8944,68528.06/.9332,106937.30
0.9599,166874.60/.9772,260406.50/.9878,406361.60/.9938,634123.2
0.9965,989542.80/.9987,1544172.00/1,3760259.0
*
        GENERATE    1                 CREATE A COUNTER
        ASSIGN      1,FN$POI50        DETERMINE THE NUMBER OF CLAIMS
        TEST NE     P1,K0,YZERO       TEST IF THE NO. OF CLAIMS IS NON-ZERO
 NZERO  SAVEVALUE   AGG+,FN$LNORM     ACCUMULATE CLAIM AMOUNTS
        LOOP        1,NZERO           PROCEED WITH THE LOOP
 NEXT   TABULATE    DISTN             TABULATE AGGREGATE CLAIM
        SAVEVALUE   AGG,K0            REINITIALIZE AGGREGATE CLAIM AMOUNT
        TERMINATE   1                 REMOVE THE COUNTER
*
 YZERO  SAVEVALUE   AGG,K0            IF CLAIM NO. IS 0 THEN AGG CLAIM IS 0
        TRANSFER    ,NEXT             CONTINUE
*
 DISTN  TABLE       X$AGG,500000,500000,40   TABLE SPECIFICATION FOR AGGREGATE
*                                             CLAIMS
        START       500
        START       500
        START       500
        START       500
        END
```

EXHIBIT 3b

GPSS Partial Output For The Second Example

TABLE DISTN

ENTRIES IN TABLE	MEAN ARGUMENT	STANDARD DEVIATION
1500	1771258.000	962048.000

UPPER LIMIT	OBSERVED FREQUENCY	PER CENT OF TOTAL	CUMULATIVE PERCENTAGE	CUMULATIVE REMAINDER
500000	12	0.79	0.7	99.1
1000000	229	15.26	16.0	83.9
1500000	485	32.33	48.3	51.6
2000000	344	22.93	71.3	28.6
2500000	194	12.93	84.2	15.7
3000000	94	6.26	90.5	9.4
3500000	51	3.39	93.9	6.0
4000000	29	1.93	95.8	4.1
4500000	25	1.66	97.5	2.4
5000000	18	1.19	98.7	1.2
5500000	9	0.59	99.3	0.6
6000000	2	0.13	99.4	0.5
6500000	2	0.13	99.5	0.4
7000000	4	0.26	99.8	0.1
7500000	1	0.06	99.9	0.0
8000000	1	0.06	100.0	0.0

REMAINING FREQUENCIES ARE ALL ZERO

TABLE DISTN

ENTRIES IN TABLE	MEAN ARGUMENT	STANDARD DEVIATION
2000	1779362.000	986624.000

UPPER LIMIT	OBSERVED FREQUENCY	PER CENT OF TOTAL	CUMULATIVE PERCENTAGE	CUMULATIVE REMAINDER
500000	19	0.94	0.9	99.0
1000000	306	15.29	16.2	83.7
1500000	639	31.94	48.1	51.8
2000000	463	23.14	71.3	28.6
2500000	248	12.39	83.7	16.2
3000000	124	6.19	89.9	10.0
3500000	67	3.34	93.2	6.7
4000000	47	2.34	95.6	4.3
4500000	37	1.84	97.4	2.5
5000000	22	1.09	98.5	1.4
5500000	14	0.69	99.2	0.7
6000000	3	0.14	99.4	0.5
6500000	2	0.09	99.5	0.4
7000000	5	0.24	99.7	0.2
7500000	2	0.09	99.8	0.1
8000000	2	0.09	100.0	0.0

REMAINING FREQUENCIES ARE ALL ZERO

Conclusions

This presentation emphasizes the simulated approach to computing distribution of aggregate claims. By using a simulation language such as GPSS, one dispenses with programming efforts and concentrates on modeling and analysis aspects of the problem. The simulation facilitates sensitivity analysis, and computation n-fold convolutions of the severity distribution is not required, nor is summation of an infinite series. The GPSS processor is available for both mainframe computers and microcomputers. Although not attempted here, one can easily extend the GPSS model to price excess-of-loss reinsurance treaties. Because reasonably accurate estimates of the distribution of aggregate claims can easily be obtained by GPSS, more use of the aggregate claims distribution may be made in pricing insurance products.

References

1. Gordon, G. *The Application of GPSS V to Discrete System Simulation*, Englewood Cliffs, N.J.: Prentice-Hall, Inc., (1975).

2. Miccolis, R. S. "On the Theory of Increased Limits and Excess of Loss Pricing." *Proceedings* of the Casualty Actuarial Society, Vol. 64 (1977).

3. Panjer, H. H., and Willmot, G. E. "Models for the Distribution of Aggregate Claims in Risk Theory." *TSA*, XXXVI (1984).

4. Schriber, T. J. *Simulation Using GPSS*. New York: John Wiley & Sons, (1974).

5. Wooddy, J. C. "Study Note on Risk Theory," Society of Actuaries, (1973).

MR. AARON TENENBEIN: Johnson and Tenenbein (1981) presented general procedures of generating bivariate distribution involving two random variables whose marginal probability distributions and rank correlations coefficients are specified in advance. In this paper, a specific procedure, called the normal method, is discussed. An extension of this method to generating multivariate

distributions involving more than two random variables is briefly discussed. Applications of these methods to simulation models, which involve the determination of surplus reserves for life insurance companies, are discussed.

1. Introduction

A model, which can be used to describe the surplus requirements for a life insurance company, is as follows:

$$R_t = R_{t-1} + P_t + \sum_{j=1}^{k} X_{tj} \tag{1.1}$$

where

R_t = Surplus at the end of period t.

P_t = Planned Contribution to Surplus during period t.

X_{tj} = Random variable representing increment (or decrement) to surplus during period t for line of business j.

This model can be utilized to answer a number of questions such as:

1. What is the probability of ruin for given strategies regarding planned contributions to surplus?

2. How should the strategy for planned contributions to surplus be designed so as to keep the probability of ruin at a sufficiently low level?

In order to answer such questions, assumptions must be made concerning the joint distribution of the random variables $X_{t1}, X_{t2}, \ldots, X_{tk}$ in each period. The specification of a joint distribution is difficult for a number of reasons. First, there is usually limited information available on the increments for the different lines of business (such as bond investments, stock investments, individual insurance, and group pensions). Second there is a limited number of theoretical multivariate distributions which are available (see Johnston and

Kotz (1972)). At best we may be able to specify the marginal distributions of $X_{t1}, X_{t2}, \ldots,$ and X_{tk} and a measure of dependence between each pair of variables. These distributions and measures of dependence may remain the same or may change over time.

In this presentation we will present methodology for constructing multivariate distributions which have given marginal distributions and specified measures of dependence between pairs of variables. These multivariate distributions can then be generated by Monte Carlo Simulation for use in the model given by equation (1.1) or in any model involving the use of two or more non-independent variables.

2. Joint Multivariate Distributions Involving Two Variables

The problem can be stated as follows. We wish to construct a joint distribution of two variables, X_1 and X_2, which has the following 3 properties.

1. Marginal cumulative distribution function of X_1 is $F_1(x_1)$.
2. Marginal cumulative distribution function of X_2 is $F_2(x_2)$.
3. Spearman Rank Correlation Coefficient between X_1 and X_2 is $s(-1 \leq s \leq 1)$.

Procedures for generating a bivariate distribution with these three properties were discussed by Tenenbein and Gargano (1979) and Johnson and Tenenbein (1981). One such procedure, called the normal method, involves the following steps:

1. Generate 2 numbers from a standard normal distribution: u', v'

2. Let $u = u'$

   $v = cu' + (1-c)v'$

3. Let $x_1' = \Phi(u')$

   $x_2' = \Phi\left[\dfrac{v'}{\sqrt{c^2+(1-c)^2}} \right]$

4. Let a) a) $x_1 = F_1^{-1}(x_1')$; and

b) i) $x_2 = F_2^{-1}(x_2')$ or

ii) $x_2 = F_2^{-1}(1-x_2')$.

Here $\Phi(t)$ is the cumulative distribution function of the standard normal distribution.

$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-z^2/2}\, dz$$

The parameter c is chosen so as to yield a given value of the rank correlation s for $0 \leq s \leq 1$. Using 3b(i) yields a rank correlation of s, whereas using 3b(ii) yields a rank correlation of -s. The relationship between c and s is:

$$c = \frac{2\sin(\pi s/6)}{2\sin(\pi s/6)+\sqrt{1-4\sin^2(\pi s/6)}}$$

The values of c as a function of s are given in Table 1.

TABLE 1

Values of c for Fixed Values of s
for the Normal Method

Rank Correlation s	Value of c
0	0
.1	.095
.2	.176
.3	.248
.4	.314
.5	.377
.6	.440
.7	.507
.8	.583
.9	.684
1.0	1.000

NOTE:

1. For positive correlations use: $x_2 = F_2^{-1}(x_2')$

2. For negative correlations use: $x_2 = F_2^{-1}(1-x_2')$

The Spearman's Correlation Coefficient is sometimes called the rank correlation coefficient and is discussed by Kendall (1962). This rank correlation coefficient is, as the name suggests, the Pearson product moment correlation coefficient between the ranks of the two variables. The use of the Pearson product moment coefficient ρ, defined as:

$$\rho = \frac{cov(X_1, X_2)}{\sigma_1 \sigma_2} \qquad (2.1)$$

is not used because it will not remain constant over the various transformations in steps 2, 3, and 4 of the normal method. There are other disadvantages of the use of ρ, which are discussed by Johnson and Tenenbein (1981).

The normal method begins with generating random numbers from a normal distribution. Other procedures involve generating numbers from other distributions, such as the uniform, exponential and double exponential distributions. These methods are discussed by Johnson and Tenenbein (1981) along with the effects on the specifications of the resulting bivariate distribution.

3. An Example Involving Stock and Bond Returns

Suppose we wish to generate a bivariate distribution for stock and bond returns with the following properties.

1. The stock return variable, X_1, is lognormal, with $\ln(1+x_1)$ having a normal distribution with mean .12 and standard deviation .10. The mean and standard deviation of x_1 will be 13.31% and 11.36% respectively.

2. The bond return variable, X_2, is uniform over the interval .05 to .12. The mean and standard deviation will be 8.50% and 2.02% respectively.

3. The rank correlation coefficient s = -.4.

To generate the joint distribution of x_1 and x_2, we proceed as follows:

1. Generate 2 numbers from a standard normal distribution: u' and v'.

2. Let $u = u'$

   $v = cu' + (1-c)v'$

   (from Table 1, c = .314)

3. Let $x_1' = \Phi(u)$

   $x_2' = \Phi\left(\frac{v}{\sqrt{c^2+(1-c)^2}}\right)$

4. (a) Let $\ln(1+x_1) = .12 + .10\Phi^{-1}(x_1') = w$

   Let $x_1 = \exp(w) - 1$

   (b) Let $x_2 = .05 + .07(1-x_2')$

By repeating these steps, the joint distribution will have the three properties stated above.

4. Extensions to Joint Multivariate Distributions Involving More than Two Variables

The normal method can easily be extended to cover the case of k variables. In this case k different cumulative distributions and $\frac{k(k-1)}{2}$ rank correlations must be specified.

The steps would involve the generating of k normally distributed random variables. Then the random variables would be correlated to form a multivariate normal distribution with the specified rank correlation coefficients. After that, the variables would be transformed to conform to the specified marginal distributions.

References

1. Johnson, M. and Tenenbein, A. (1981). "A Bivariate Distribution Family with Specified Marginals," *Journal of the American Statistical Association*, Volume 76, Number 373, pages 198-201.

2. Johnson, Norman L., and Kotz, Samuel (1972). *Distributions in Statistics: Continuous Multivariate Distributions*, New York: John Wiley.

3. Kendall, Maurice G. (1962). *Rank Correlation Methods*, London: Charles Griffin.

4. Tenenbein, A. and Gargano, M. (1979). "Simulation of Bivariate Distributions with Given Marginal Distribution Functions," Chapter 18 of *Current Issues in Computer Simulation* (Editors Adam, N. and Dogramaci, A.). New York: Academic Press Inc., pages 271-282.

MR. SHAPIRO: This session was put together, in part at least, by the Committee on Relations with Statistical Associations of which I am the Chairman and Jed is the Vice-Chairman. We are very interested in any areas in statistics that people might like to see pursued and presented at these kinds of sessions. So if anyone has some interest we would be very happy to hear about it.

MR. THOMAS P. EDWALDS: I was curious as to what techniques are available for determining the optimal number of iterations that should be utilized in a simulation and does this depend on your hardware or particular software package?

MR. GUIAHI: I would look at the simulated distribution to see whether it stabilizes, but there might be other techniques.

MR. TENENBEIN: It depends on what your aim is. There are methods if you are for instance trying to do what Farrokh had been doing in estimating the average value. If he was estimating the average value of stock claims above a given level, then there are various confidence intervals that you can use.

For instance, you might be calculating the mean and also calculating the standard deviation of all those numbers, then using a sampling theory in order to determine how close you are to the "right answer" if you were able to simulate say millions upon millions of times. I was being very facetious when I said millions upon millions of times, but if you do it thousands upon thousands of times say for 2,000 simulations, how close are you to the theoretical average value you would have if you had unlimited money in your computer budget, unlimited time to wait, and boredom wasn't a factor? I don't know if that relates to your question, but that is basically within the realm of confidence intervals.

MR. STUART KLUGMAN: Jed, we've learned that you could successfully simulate the aggregate distribution. I could plot that distribution and physically extract the 95th percentile or whatever you might be looking for. I presume your approach is designed to get me to that point more efficiently. Can you give me some idea of what the advantage in efficiency is so that I'll want to learn how you do it?

MR. FREES: I call the 95th percentile approach the naive approach and I didn't go over that. I didn't want to waste your time too much on the naive approach. This is indeed the method that one should use for an initial start-up procedure to get you in the ball park. Specifically, I designed the algorithm in mind with a micro-computing environment where one does not have vast amounts of working storage and what you'll like to do is just like a Newton-Raphson technique to keep a few values around, in particular, your latest guess. Just like in Newton-Raphson you'd like to update things as they go, and if you are worried about storage capacity, then you would want to go with the stochastic Newton-Raphson technique. I would say that was the big advantage over the first approach that you outlined.

MR. ALBERT K. CHRISTIANS: I think we talked about this once before. When you go through your normal method, you seem to be imposing some kind of normality assumption on your joint distribution. I am concerned that, in the problems that we work with, that may be the most well behaved of the joint distributions, when in fact you want to evaluate what some upper limit on a proability of ruin is or where the variables are not well behaved. It seems

like you are imposing certain amounts of order on things, and I'm not sure you thought about this problem, but maybe you could explain whether it is inherent in what you've done. An alternative that I would like to suggest or evaluate in comparison would be to just generate the uniform values directly and then do some kind of either rejection/acceptance with one more random variable or just switching the pairs of u's and v's based on one additional uniform random variable.

MR. TENENBEIN: That's an excellent suggestion. I think I do remember that suggestion from before. Basically, with respect to the way the method started with two normal distribution random variables, it didn't matter what you started with. I could have started with two uniform random variables which I have in the Tenenbein-Gargano paper referred to in my presentation, which instead of normally distributed random variables starts with uniform; it takes the weight of linear combinations and goes through the process.

We went a step further and took a look at a whole family of ways of doing it. That is the subject of the 1981 paper, and we took a look at specific examples where you start with exponential variates, also with double exponential variates and work that out. With respect to the normality assumption, the fact that you start with a normal distribution doesn't impose that the distribution that you end up with is normal. As a matter of fact, the distributions weren't, and in the example I did, I took great pains to make sure that we weren't ending up with the normal. We ended up with a bivariate distribution whose marginals were lognormal and also uniform, and it could be anything. The normal distribution is just a way of getting to the marginal distributions.

You start with any distributions you care to name. There is no unique way of generating your distribution because the point is you're starting with limited information anyway so you know exactly what the joint distribution is. There shouldn't be any surprise that you end up with joint distribution which is not unique. So that's a problem that I have wrestled with quite a bit.

We have talked about how to discuss how the different results of your risk ruin model or the probability of ruin and how sensitive the results are to starting with the assumption of whether it's normal or uniform or double exponential or

any other distributions. The problem you have is that, when you get to more than two variates, trying to hypothesize a uniform gets to be very difficult. The idea that you mentioned about the rejection criterion I think might have some merit. I hadn't thought about it in quite that way, but that might have some merit in the way of generating the distributions. The problem we get to is how do we get back to the uniform distribution? That gets into very difficult integration problems where, for the normal distribution, it means just looking into the table of normal distribution random variables. But for uniforms and for the others you get into very complex integration problems. What was interesting about it is the types of integration problems that we had were so horrendous that we reported to a program called maxima which was developed by the Massachusetts Institute of Technology which actually goes through algebraic manipulation of integrals. That was basically the inherent problem that was involved, and it gets kind of horrendous.

I guess my own suggestion would be to, at least in the bivariate case, take a look at the different ways of doing it and see how sensitive your risk probability of ruin is to the specification of these different distributions. If it is fairly insensitive to the specifications, which we found was the case in the work I had done at Equitable, then we would hope when you get into the multivariate cases the same thing would carry through.

MR. CHARLES S. FUHRER: I just wonder if Mr. Guiahi could explain why it is such a good idea to use simulation methods at all? It just seems to me there are all kinds of good approximation methods for coming up with the final distributions given the initial unconvoluted ones.

MR. GUIAHI: Some of the approximation techniques might only evaluate the distribution of certain specified values. With the simulation you can get the approximation for the entire distribution. I would say in my case using GPSS would be to enlarge this model and solve bigger problems that you would be interested in as the insurer, this would be only part of it. If you can extend this to a bigger model, that solves other problems not just simulation of IDS claims because there are many ways to do this and one has to evaluate the merits of each one. As to model building this is a good approach.

MR. TENENBEIN: I found at least in work that I've been involved with that simulation was a very neat approach in the case of trying to hypothesize the joint distributions. You can't even get at the joint distributions. You can simulate them, but it is very difficult to try to get the joint distribution and do anything analytical with respect to it. I found it of great use to use simulation. I think with respect to the number of simulations you have to carry out, is that simulation, though theoretically should give you the exact answer, it will not, because practically you are limited to the number of simulations that you can carry out. I think the question you had raised was very important, and you should be concerned with how much error is involved in the simulation. I think that is quite an important thing to be concerned with, and unfortunately a lot of people aren't. But what you find here are the results; no attempt is made to get the errors. There is a very simple way to get at the errors especially if you're dealing with means. It goes back to sampling theory, and it is really confidence interval estimations. It is probably one of the most useful areas or applications of confidence interval estimations that I know of.

MR. SHAPIRO: I guess the question is the uniqueness of your result.