12 October 2007
STOCHASTIC MODELLING FOR DUMMIES

By Philippa Wild and Yehuda Hockman

ABSTRACT
Modelling is a core actuarial technique. Historically actuaries followed deterministic approaches when projecting future outcomes from a given set of data. Stochastic modelling allows for the random nature of the input data and then produces outputs that are similarly random in nature. Currently stochastic modelling is mostly used in the fields of health, short term insurance, life insurance, investments and asset liability matching.

The purpose of this paper is to explain stochastic modelling by means of a practical life insurance / mortality example. This example will provide non-technical readers with a better understanding of the work done by stochastic modelling practitioners. This paper also aims to explain the benefits of this actuarial technique in communicating risk and uncertainty. This paper provides a basic platform of stochastic modelling knowledge, enabling the reader to do further research as he/she deems appropriate.

This paper is specifically aimed at actuaries who can recall a time when the Institute and Faculty wrote different exams (some easier than others) and more recently qualified actuaries who passed subjects referred to as A, B, C & D.

CONTACT DETAILS
Wild, P, Short Term Insurance Actuary, Munich Reinsurance Company of Africa Limited, 47 Empire Road, Parktown, Johannesburg, South Africa, 2193; Tel/fax: (27) 11 242 2218; E-mail: pwild@munichre.com

Hockman, Y A, Life Actuary, Munich Reinsurance Company of Africa Limited, 47 Empire Road, Parktown, Johannesburg, South Africa, 2193; Tel/fax: (27) 11 242 2019; E-mail: yhockman@munichre.com

1 INTRODUCTION

In April 2000 the exam syllabus for the core technical subjects of the Institute and Faculty of Actuaries changed from subjects A - D, to subjects 101 - 109. For the first time, explicit stochastic modelling was incorporated into the skill set requirement for actuaries by means of subject 103: Stochastic Modelling. Anybody passing their core technical subjects before 2000 would therefore not have studied this key modelling technique.

Currently stochastic modelling is used by many students and actuaries in various fields including health insurance, short term insurance, life insurance, investments and asset liability matching. This paper is not intended for these practitioners. Instead, this paper aims to explain stochastic modelling by means of a practical life insurance / mortality example to those actuaries and students who were not fortunate enough to cover this aspect of modelling within their actuarial curriculum.

Stochastic modelling is a complex subject. As stated in section 2, the intention of this paper is not to address the many complexities of this subject, but rather to explain the basic principles and methodology by means of an example. The example will also illustrate the benefits of using stochastic modelling in practice.

This paper does this by defining stochastic modelling and random variation by means of a coin toss example in section 3. Section 4 explains why stochastic modelling is used in practice as opposed to a deterministic approach. Section 5 then discusses the techniques of random number generation, a pivotal item in the stochastic modelling toolkit, including the potential pitfalls / limitations of applying certain of these techniques.

After equipping the reader with basic random number generation skills, section 6 details a practical group life / mortality example. The mathematics of a stochastic exercise is explicitly shown, including example outcomes, which are then used to better understand the insurance risks associated with this group.

Section 7 covers determination of the optimal number of simulations, explaining how this challenge can be solved. This is followed by section 8 which informs the reader of additional, important aspects to consider when embarking on one's own stochastic modelling exercise.

After the conclusion in section 9, section 10 provides the reader with pointers on how to research this topic further.

2 DISCLAIMERS

This paper:
- Is not a statistical paper and doesn't cover the detail of the underlying statistics necessary for stochastic modelling;
- Contains no new research or findings with regards to the topic of stochastic modelling;
- Is not aimed at stochastic modelling practitioners; and

- Does not equip the reader to perform a rigorous stochastic modelling exercise as it merely scratches surface of what to consider when performing such an exercise.

The purpose of this paper is not to interrogate or explain random number generating methods in any great detail. This paper does not cover the various software solutions and economic scenario generators available for this purpose, but rather provides useful explanations and insights into the generation of random numbers by means of a computer.

The example given in section 6 was done in Excel. Later on, this paper elaborates on the appropriateness of Excel as an appropriate software package in which to do stochastic modelling. However, it should be noted from the outset that Excel is not ideal and in practice most companies/stochastic modelling practitioners use custom designed packages for their stochastic modelling exercises.

Throughout this paper, information is given about various aspects of stochastic modelling that need to be considered and appropriately addressed before embarking on one's own stochastic modelling exercise. The issues highlighted in this paper and specifically in section 8 are not comprehensive and the intention is not to provide concrete solutions, but rather an understanding of why certain issues are important.

3 WHAT IS STOCHASTIC MODELLING?

The 2002 Subject 103 notes contain the following paragraph in Unit 1: Principles of actuarial modelling:
"If it is desired to represent reality as accurately as possible the model needs to imitate the random nature of the variables. A stochastic model is one that recognises the random nature of the input components. A model that does not contain any random component is deterministic in nature."

Consider the outcome of an unbiased coin toss, which has two possible outcomes, heads or tails. If tossed in the same environment and under the same circumstances, the probability of the result being either heads or tails is exactly 50%. However, due to the random nature of the outcome, it is impossible to predict, with certainty, what the outcome will be, heads or tails. If the coin is tossed 1 million times, one would expect to see half a million heads and half a million tails, due to the statistical independence of each coin toss. This would be consistent with the known parameter of 50%. However, this expectation may not be borne out in practice due to the random nature of each coin toss.

Stochastic modelling attempts to reproduce such a real world scenario as realistically as possible. This is achieved by randomly simulating numerous possible outcomes and interpreting these results to give the modeller further insights into the underlying nature of the risk that is being modelled.

4 WHY DO IT?

Consider the dilemma of ensuring that an insurance company's assets exceed its liabilities, i.e. is solvent, now and in future years. Both the assets and liabilities are unknown entities as they are dependent on the actual future experience of a variety of

parameters including on the asset side: inflation and investment return; and on the liability side: claims inflation and the number of claims.

The deterministic approach involves determining the best estimate projections of assets and liabilities and subtracting the two to determine solvency. Graphically the results of this can be seen as follows:
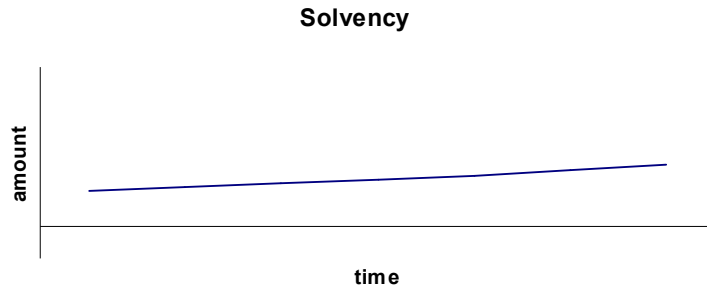
Solvency

In this example solvency is increasing gradually over time. The disadvantage of a best estimate is that it ignores the fact that there is uncertainty in these estimates and provides no further information about the level of risk associated with the best estimate.

Stochastic modelling allows the key parameters, such as interest and inflation in this example, to vary randomly in order to determine what potential outcomes of assets and liabilities and hence solvency, may arise. These key parameters are made to vary randomly between their possible future minimum and maximum values. For example the cost of a motor claim can vary randomly between the excess amount and the sum insured in a stochastic modelling exercise. Such modelling builds volatility and variability into the simulation and therefore provides a more accurate representation of reality. Graphically this can be seen as follows:
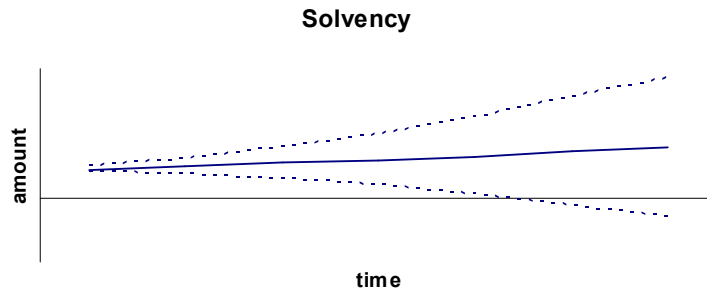
Solvency

The dotted lines indicate the relative maximum and minimum solvency amounts at each future point in time resulting from the stochastic modelling. The deviation from best estimate increases with time as the key parameters become more uncertain the further into the future they are being modelled.

It can be argued that the same results can be achieved by applying simple scenario testing within a deterministic model. This is true, but stochastic modelling goes a step further in that its outputs can be used to determine the likelihood of each outcome. Hence stochastic modelling can quantify answers to questions of the following nature: What is the probability of insolvency in each future time period? or What is the solvency value at the 99.5^{th} percentile for any given time period?

4

This is done by assessing the outcomes at any time t. Stochastic modelling enables the modeller to fit a distribution to these outcomes which is explained graphically as follows:

Solvency
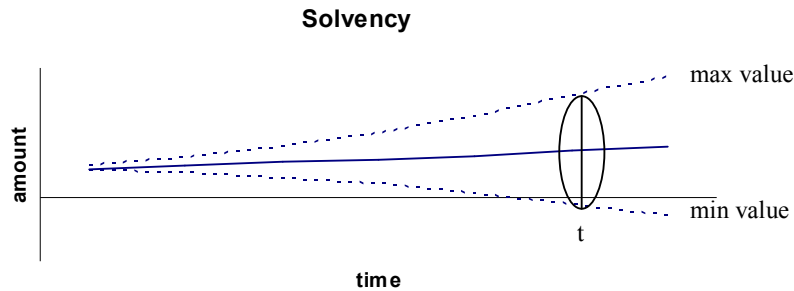
Firstly all outcomes at time t are extracted. These are found on the straight line within the oval in the above graph. Then a statistical distribution is fitted to these outcomes, derived from the number of times each outcome happens within the stochastic modelling exercise. An example of such a distribution is shown below.
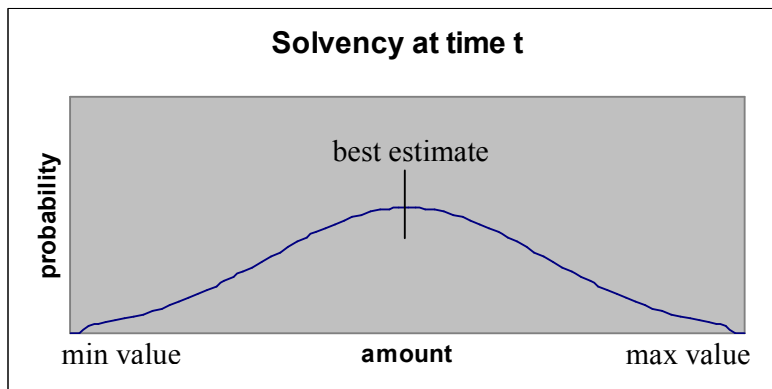
Solvency at time t

Having the distribution of outcomes at time t, the modeller can now determine the likelihood of insolvency, the 99.5^{th} percentile or any other chosen statistic.

As can be seen in this graph, stochastic modelling enables an actuary to graphically explain the associated risks and likelihood of a projected outcome, which in the above example is solvency. Such information can assist decision makers in determining an investment strategy that meets their optimal level of risk and return with greater certainty. Simultaneously it can provide greater insight into the liability profile being modelled, leading to more informed risk mitigation and management decisions.

Following from the above, stochastic modelling can produce more meaningful information than a deterministic approach when one is trying to predict an unknown future outcome. This is done by determining a distribution for the potential outcomes from which one can derive key statistics (e.g. mean, standard deviation, and percentiles) in order to assess the uncertainty/risk associated with those outcomes.

 In actuarial practice, stochastic modelling is applied in many fields to many problems, including:
• Dynamic Financial Analyses;

- Option valuation and pricing;
- Pension fund asset liability modelling; and
- Reinsurance pricing.

5. KEY ITEM IN TOOLKIT: MODELLING RANDOMNESS

A key item needed in any stochastic modelling toolkit is the ability to model randomness. True randomness is difficult, if not impossible to model. Generating a set of random numbers by means of a formula, in essence makes them non-random as they have become predictable. However, formulas are used in practice to generate "random" numbers, i.e. numbers that appear to be random in nature and don't exhibit any patterns. Such numbers are referred to as pseudo-random numbers. For the remainder of this paper, the word "random" should be understood to mean pseudo-random.

Of all the stochastic modelling methods, Monte Carlo simulation is most used within the actuarial profession due to its relative simplicity and ease of application. Monte Carlo simulation studies the statistical distribution of a stochastic quantity. This involves carrying out a large number of simulations, by means of a random number generating formula, which can be used to estimate probabilities and moments of distribution.

This section elaborates on the methods of generating pseudo-random numbers, forming the basis of Monte Carlo simulation, which is used in the example in section 6. The more technical explanations of this section are contained within text boxes and should be skipped over by those readers preferring a more general/overall understanding.

5.1 Linear congruential generators (LCGs)
LCGs are the most commonly used random number generators used today. LCGs generate a sequence of random numbers that are uniformly distributed between 0 and 1 (U(0,1)), i.e. on any given straight line between 0 and 1 and hence the term linear. The random numbers generated by this method are treated as an independent, identically distributed sequence of values uniformly distributed between 0 and 1.

Microsoft Office Excel has an LCG function which is detailed further below.

To understand the LCG formula lets first consider the Mod function in Excel:
This function returns the remainder after a number is divided by a divisor, as follows:
MOD(number, divisor), where:
Number   is the number for which you want to find the remainder;
Divisor   is the amount by which you want to divide the number.
e.g. MOD(n,d) = n – d* INT(n/d)
(INT is the integer function in Excel. It returns the integer portion of a number)
e.g. MOD (10, 8) = 10 – 8* INT(10/8) = 2

An LCG is specified by the choice of three positive integers:
a,   the multiplier;
c,   the increment;
m,   the modulus / divisor;
where m>a, m>c.
To generate the sequence of random numbers U1,U2,…,Un first generate a sequence of integers X1,X2,…,Xn in the range 0,1,2,…m-1 starting with an initial value, known as the seed.

The seed can be any number from 0 to m-1. The sequence of integers is then generated by the following formula:

$$X1 = Mod(a * seed + c, m)$$
$$X2 = Mod(a * X1 + c, m)$$
….
$$Xn = Mod(a * Xn-1 + c, m)$$

The random numbers are derived as follows:

$$U1 = X1/m,$$
$$U2 = X2/m$$
…
$$Un = Xn/m$$

One potential shortcoming of this method is that it cannot generate a random number equal to 1, only less than 1. However, this can be overcome, relatively speaking by choosing a large enough value for m.

Another potential shortcoming is that the random numbers are not exhaustive and can only be generated from i/m where i = 0,1,2,…,m-1. This means that the random numbers will start repeating themselves after a set number of runs depending on the original choice of a, c and m. Again, this can in most part be overcome by choosing your parameters appropriately.

The LCG can be derived by the Excel formula RAND(), which returns a random number between 0 and 1, as per the above formula where the parameters are the following:

a	=	1,140,671,485
c	=	12,820,163
m	=	2^24

The major shortcoming of this function is that it can only generate 16,777,216 unique numbers which is not always sufficient depending on what one is trying to simulate. The optimal number of simulations is discussed further on in this paper in section 7.

Of further interest:
  RAND()*(b-a)+a generates a random real number between a and b; and
  To recalculate the random number of the cell, press "F9" or "enter".

5.2 Generating random numbers from a specified distribution

In certain circumstances it is desirable to generate a random number from a specific distribution as opposed to a random number uniformly distributed between 0 and 1. This will depend on the actual random nature of the outcome one is trying to predict. This can be done by numerous methods. One example is the inverse transform method.

5.2.1   Inverse transform method

This method generates random numbers from a given distribution f(x), with cumulative distribution function F(x) by first generating a random number u from U(0.1). Then, if the inverse of F(x) is defined as $F^{-1}(x)$ it determines the random variable $x = F^{-1}(u)$

A shortcoming of this method is that an explicit expression for the inverse distribution of F(x) is needed, which may not always be mathematically possible. Hence this method can only be used in the discrete case or in the continuous case when the distribution function has a simple form. This is the reason why specific algorithms are needed for the generation of random numbers from the standard continuous normal distribution. (such as the Box-Muller algorithm)

This method is best explained graphically by considering the cumulative distribution function of the Pareto distribution, shown below:
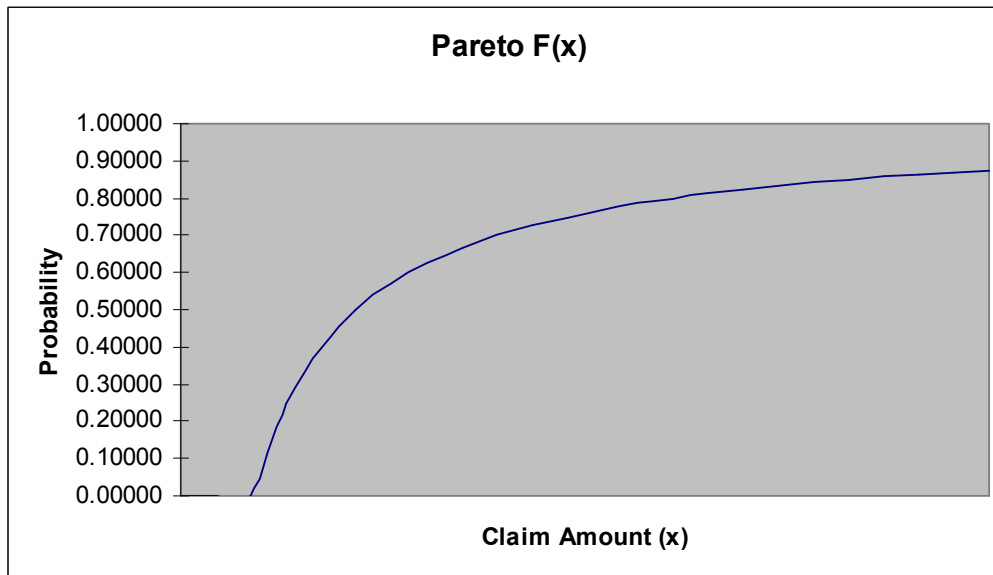
The Y-axis reflects the probability, between 0 and 1 of cumulative claim amounts being less than or equal to x, shown on the X-axis. As can be seen from the graph, each Y-value corresponds to an unique X-value. Hence, to generate a random number on the Pareto curve, the inverse transform method does the following:

- Generates a random number between 0 and 1;
- Takes this number to be a probability on the Y-Axis of the cumulative distribution function;
- Finds the corresponding x on the X-axis by solving for x using the inverse of the cumulative distribution function and the randomly generated number;
- This x is then deemed to be a randomly generated number from the Pareto distribution.

The mathematical workings of the inverse transform method are not covered in this paper but can be researched further, along with other methods. Please refer to section 10 for useful tips on further research.

6. A PRACTICAL GROUP LIFE EXAMPLE

The aim of this example is to model stochastically (via Monte Carlo simulation) the mortality experience of a defined group of lives. The example chosen is exceptionally simple and should be viewed in this light. This example aims to explain the actual calculation/workings of a stochastic model, including the application of the results enabling the reader to understand the practical benefits of doing a stochastic modelling exercise.

This example follows the following process:
- Explanation of the underlying assumptions and determination of the price of mortality cover for the group of lives over a 1 year time horizon (6.1);
- Discussion of the insurance risks associated with this book of business (6.2);
- Production of a stochastic model with numerous possible outcomes of the mortality results over this one year period of cover (6.3); and
- Interpretation of the results from the perspective of the insurer who will be insuring this group over the next year (6.4).

6.1 Assumptions

6.1.1   The group of lives
The group of lives has the following demographic profile, including mortality and premium rates:

Age group	Number of lives	Average sum insured	Average mortality	Premium
20 - 24	1	486,000	0.0023900	1,162
25 - 29	6	482,478	0.0021367	6,190
30 - 34	15	568,804	0.0020947	17,866
35 - 39	19	829,879	0.0024795	39,346
40 - 44	13	455,954	0.0034246	20,574
45 - 49	18	694,668	0.0050261	63,053
50 - 54	14	728,532	0.0077057	80,634
55 - 59	9	842,856	0.0105278	77,777
60 - 64	4	1,286,748	0.0207625	106,071
65 - 69	1	760,608	0.0267500	20,346
Total	100			433,019

The death benefit/sum insured is 3 times annual salary. All members are male. Please refer to Appendix A for more detail.

6.1.2   Mortality Table
The chosen mortality table is SA85-90 ultimate and is deemed appropriate for the selected group of lives. Please refer Appendix B for more detail.

6.1.3   Random Number generator
The Microsoft Office Excel 2003 function RAND(), will be used. However, this is not ideal as explained in section 5 (boxed area). It is assumed that for the purposes of our stochastic modelling exercise that this random number generator is suitable.

6.1.4 Simplicity

For the sake of simplicity, this example ignores the following:

- Rating factors, other than Age;
- HIV/Aids and its impact on mortality;
- Any additional insurance premium loadings including expenses, return on equity, etc. i.e. profit is simply equal to premiums less claims;
- Any policy conditions that could impact premiums, such as free cover limits and medical underwriting; and
- Any other factors not specifically mentioned here that could impact the mortality in any way.

6.2 Risks

Applying the mortality table to the group of lives results in a total premium of R433,019. Should the insurer choose to take on this business, there is a risk that the claims experience will be greater than the premium, thereby causing a financial loss.

This risk can be attributed to three underlying causes:

6.2.1 Parameter Risk

The parameters, i.e. the mortality rates of SA85-90, may be wrong. For the purposes of this example, it is assumed that there is no parameter risk. Please see section 8 for a more detailed discussion about parameter risk.

6.2.2 Accumulation Risk

The specific circumstances of this group may lead to additional accumulation risk that needs to be considered, e.g. a large majority of members travel frequently by means of the same mode of transport. For the purposes of this example, accumulation risk is assumed to be non-existent and each life is modelled independently.

6.2.3 Trend Risk

There is a risk that the underlying mortality rates change between the time of the data analysis and pricing exercise and the period to which the rates actually apply. This is referred to as trend risk and is usually more significant the further into the future one is pricing for. In this example trend risk is ignored.

6.2.4 Random Fluctuation Risk

Having eliminated other potential sources of risk, there is still a chance that the claims experience will be greater or smaller than R433,019. This is due to random fluctuation, as explained in the coin toss example in section 3. This risk is to be modelled stochastically.

6.3 Stochastic model

6.3.1 Generating the output

In this example, the stochastic model simulates the annual claims experience 10,000 times. Each simulation does this by randomly simulating whether each member died or not. i.e. each simulation contains 100 (number of members) random number generations. Therefore, there will be a total of 1,000,000 (10,000 simulations * 100 members) random numbers generated. This is done as follows:

Simulation 1:
Consider Member 1, with a sum assured of R486,000, mortality of 0.00239 and a premium of R1,161.54. He has a 239 in 100,000 chance of dying in the next year and a 99,761 in a 100,000 chance of surviving the next year. Worded a bit differently, 239 times out of 100,000 he will die and 99,761 times out of 100,000 he will live. The model randomly generates a number between 0 and 1. If this random number is less than 0.00239 (i.e. one of the 239 times), then it is decided that Member 1 died and the sum assured of R486,000 is paid. Conversely, if this random number is greater than 0.00239 (i.e. one of the 99,761 times), then it is decided that Member 1 survived and no claim is payable.

The Model then considers Member 2 and follows the same logic with Member 2's mortality, by generating another, different (due to the chosen random number generator) random number. The model does this again for member 3 and so on until it completes simulation 1 for all 100 members.

Simulation 2 and onwards:
The model re-runs the same process it completed for simulation 1, when deriving simulation 2, i.e. it randomly determines whether each of the 100 members survived or died in that given year. The model then continues until it has completed all 10,000 simulations, in a similar way.

This simulation coding can be seen in the Excel extract on page 12 below. The top table shows the actual Excel code. Each simulation is done in a separate row, starting in row 9 through to 12 and ending in row 10,008 (not included in the extract below). Each of the columns G and H and continued to column DB (not included in the extract below) represent a member. Rows 4 to 6 provide membership information (member number, mortality rate and sum assured) for each of the 100 members. The members' results of each simulation can be found in the area G9:H12. Columns C and D show the total number of deaths and cumulative claim amount of each simulation, i.e. each simulated year. The bottom table shows the actual result of the coding in the top table.

Unfortunately not all the stochastic modelling results could be shown due to the size of the spreadsheet, but this example can easily be simulated in Excel and similar results can be reproduced. It is important to note that the exact results will not be replicated as each random number generation will produce a different set of random numbers, i.e. different to the ones generated in this example.

> (TIP: Once the Rand() function formula has been copied to all the cells (G2:DB10,008) it is useful to either turn the "calculation" function off (Tools/Options/ in calculation tab move the dot from "automatic" to "manual") or to copy all the results and "paste special, values" in the same area. The reason for doing this is that each time you press "enter" in Excel or F9, the Rand() function will generate a new random number, which will change the results to some degree. Leaving the formula also takes up more computing space.)

2	C	D	E	F	G	H
3						
4		**Stochastic Model**		Member	1	2
5				Mortality rate	0.00239	0.00228
6				Sum Assured	486000	342000
7						
8	Number	Amount		Simulation		
9	=SUM(G9:DB9)	=SUMPRODUCT($G$6:$DB$6,G9:DB9)		1	=IF(RAND()<=G$5,1,"")	=IF(RAND()<=H$5,1,"")
10	=SUM(G10:DB10)	=SUMPRODUCT($G$6:$DB$6,G10:DB10)		2	=IF(RAND()<=G$5,1,"")	=IF(RAND()<=H$5,1,"")
11	=SUM(G11:DB11)	=SUMPRODUCT($G$6:$DB$6,G11:DB11)		3	=IF(RAND()<=G$5,1,"")	=IF(RAND()<=H$5,1,"")
12	=SUM(G12:DB12)	=SUMPRODUCT($G$6:$DB$6,G12:DB12)		4	=IF(RAND()<=G$5,1,"")	=IF(RAND()<=H$5,1,"")

2	C	D	E	F	G	H
3						
4		**Stochastic Model**		Member	1	2
5				Mortality rate	0.00239	0.00228
6				Sum Assured	486000	342000
7						
8	Number	Amount		Simulation		
9	2	1644588		1	0	0
10	1	931464		2	0	0
11	0	0		3	0	0
12	0	0		4	0	0

In simulation 1, indicated by cell F9, Member 99 and 100 died (not shown explicitly in above extract, refer to Appendix A). Their sums assured are R883,980 and R760,608 respectively, resulting in 2 deaths with total claim value of R1,644,588. (Cells C9 and D9) In simulation 2 Member 87 (refer to Appendix A) with sum assured of R931,464 died. In simulations 3 and 4 no members died.

Expanding the above table, results in 10,000 possible annual outcomes for the mortality experience of the group of lives. i.e. each life is modelled 10,000 times

Considering Member 1 again and assuming a truly stochastic model, the expectation is that Member 1 will die in 239 of 100,000 scenarios or in 23.9 of 10,000 scenarios as per the mortality rate. In this stochastic modelling example, Member 1 died 21 times in the 10,000 scenarios as opposed to the expected 23.9 times. This potentially indicates that the number of simulations could be insufficient and/or the random number generator chosen, is not ideal.

6.4 Interpreting the stochastic modelling results
The output of the stochastic model has resulted in 10,000 rows of information, with 100 member results for each row. This information/output is now combined and

summarised in different ways to enable the insurer to better understand the mortality risk associated with this group of lives. This analysis is detailed as follows:

6.4.1   Average Claim amount

The average claim amount is R430,717. In an ideal model, this would equal the premium of R433,019. If the stochastic modelling exercise was redone, it may in all likelihood produce difference results, due to the random number generation process. The reason for the difference between the average claim amount model output and premium is due to the chosen random number generation and the number of simulations produced. For more information, refer to section 7.

6.4.2   Number of deaths

The table below shows the distribution of the number of deaths based on the 10,000 simulations, i.e. there were approximately 5,800 simulations that resulted in no deaths.

Number of deaths in a year	Probability
0	58.0%
1	31.8%
2	8.7%
3	1.4%
4	0.2%
	100%

Graphically this is shown by:
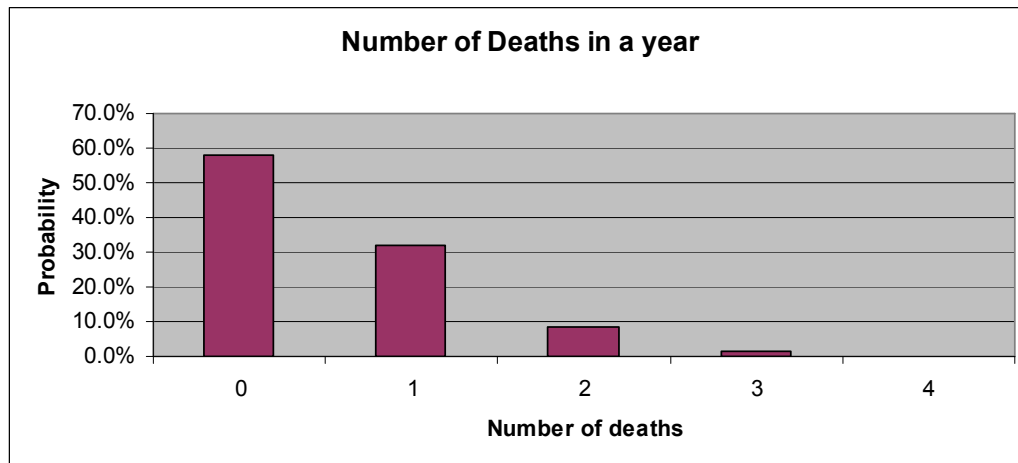
The results of the deterministic exercise results in an expected number of deaths of 0.54116. In contrast the stochastic exercise provides useful insight into the underlying distribution of the number of deaths enabling a better understanding of the risk.

Knowing the distribution of the number of deaths adds value, and even more insight can be gained into the underlying risk by examining the claim distribution.

6.4.3   Claim amount

The table below shows the distribution of the claim amounts based on the 10,000 simulations, i.e. there were approximately 5,800 simulations that resulted in no deaths and hence no claims. The remainder of the table shows the probability of the claim amounts being in the ranges shown in the left hand column.

Cumulative claim amount in a year	Probability
-	58.0%
< 200,001	0.0%
200,001 - 400,000	5.3%
400,001 - 600,000	6.9%
600,001 - 800,000	9.2%
800,001 - 1,000,000	8.5%
1,000,001 - 2,000,000	7.8%
2,000,001 - 3,000,000	3.2%
3,000,001 - 4,000,000	1.0%
4,000,001 - 5,000,000	0.1%
5,000,001 - 6,050,000	0.1%
	100.0%

Graphically this is shown by:
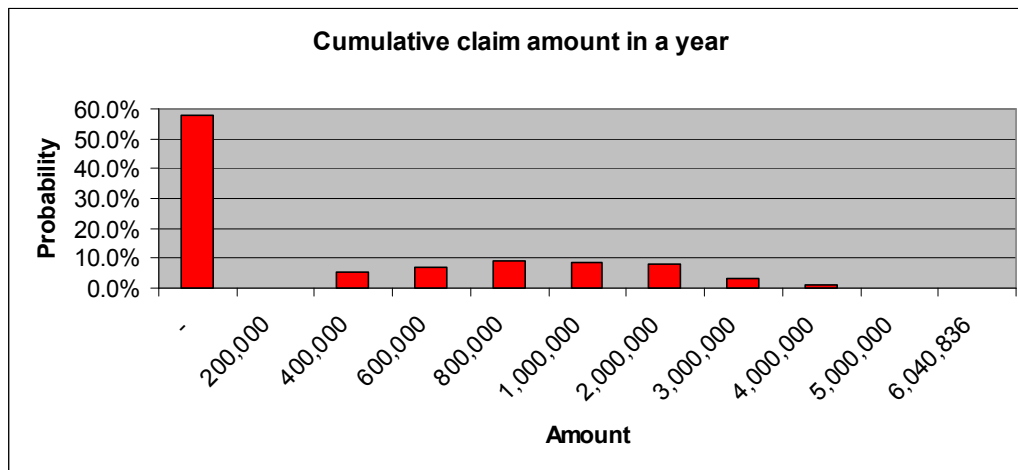
For the insurer, it is now useful to assess the likely profitability of this group by considering the premium charged in conjunction with this claims analysis.

6.4.4   Profitability

Similarly to 4.4.3 above, the table below shows the distribution of the profitability given the premium of R433,019 (derived from the deterministic approach).

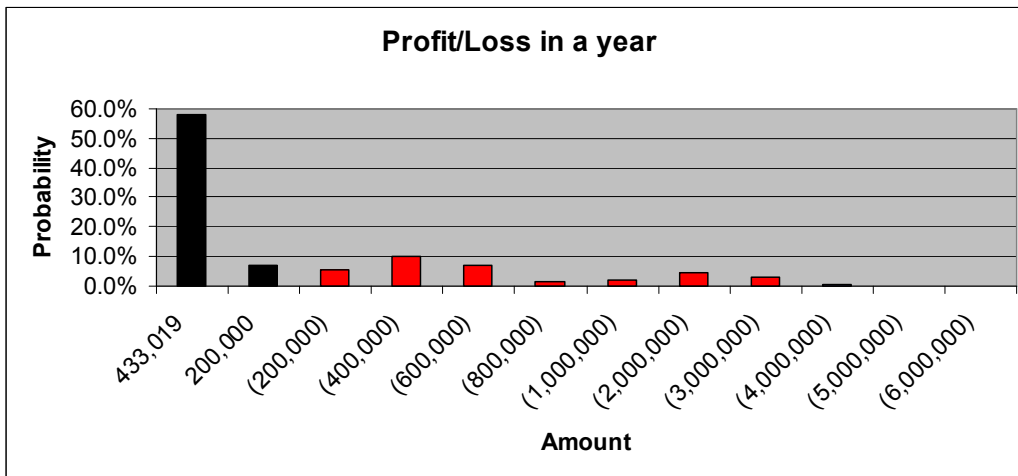| Profit/Loss in a year	Probability
433,019	58.0%
4330,18 - 200,001	7.1%
200,000 - 0	5.7%
0 - (200,000)	10.0%
(200,001) - (400,000)	7.2%
(400,001) - (600,000)	1.7%
(600,001) - (800,000)	2.2%
(800,001) - (1,000,000)	4.7%
(1,000,001) - (2,000,000)	2.8%
(2,000,001) - (3,000,000)	0.5%
(3,000,001) - (4,000,000)	0.0%
(5,000,001) - (5,610,000)	0.1%
	100.0%

Graphically this is shown by:

Note that the first two columns in this graph (in black) indicate profits and the remaining columns depict losses.

The above profit and claims distributions can now be used to derive other key statistics such as the standard deviation and any percentiles that need analysing. This provides the modeller and the recipient of the results with a better understanding of the risk being modelled, i.e. how the outcome behaves (the possible levels of claims and profitability) together with the associated probabilities.

7. HOW MANY SIMULATIONS?

One of the key questions that has arisen during this paper is: How many simulations are appropriate for any given stochastic modelling exercise? This can be derived by means of convergence testing.

This is done by choosing certain key statistics like the mean, standard deviation and various percentiles. These statistics are key in the sense that they adequately explain the underlying distribution and associated risk. The modeller then needs to test how each of these statistics changes by including additional simulations. This change in

the statistic due to the additional simulations is measured in percentage terms. When this percentage is sufficiently low, then the results of the stochastic modelling exercise are deemed to have converged, meaning that the number of simulations performed is sufficient. The box below explains a brief example.

This can be seen in the table below which shows the percentage change in two key statistics (the mean and 99.5th percentile) by adding 1000 simulations in each row.

The table below is just an example of the output from a stochastic model and is not related to the example found in section 6

CONVERGENCE TESTING

Number of simulations	Mean	99.5th percentile
8000	0.10%	2.14%
9000	0.03%	1.80%
10000	-0.02%	1.54%

The following is derived from the above:

7.1 Mean
The mean at 9,000 simulations changes by -0.02% by adding another 1,000 simulations. It can be concluded that the mean has converged

7.2 99.5th Percentile
The 99.5th percentile at 9,000 simulations changes by 1.54% by adding another 1,000 simulations. It can be concluded that this statistic has not converged.

As the statistic moves further away from the mean, it takes longer to converge. The reason for this is that the further away from the mean, the fewer the number of simulations that actually yield results in these outlier positions. Hence a lot more simulations are needed to provide convergent experience in these statistics.

Convergence testing is a critical part of any stochastic modelling exercise. If convergence testing is not done and the simulations in a stochastic modelling exercise are too few, the resulting distribution will be incorrect. This will invalidate the results and cause erroneous results and incorrect decision taking as a result.

8. ADDITIONAL THINGS TO CONSIDER

So far this paper has mentioned certain things that need to be considered as part of a stochastic modelling exercise. In practice the range of things to consider is a lot wider than those mentioned here. Some of these issues are elaborated on a bit further in this section.

8.1 Parameter Risk
Modelling parameter risk is a key component of stochastic modelling. In the mortality example in section 6, the key parameter: mortality, may be incorrect. Many stochastic models take parameter risk into account as this risk has a direct impact on the results.

To model parameter risk stochastically one needs to determine to what degree the parameter can vary (e.g. minimum and maximum value) and how the parameter varies within the specified range (randomly on a straight line or on a certain distribution or in some other way). The parameter itself is modelled stochastically, by means of a random generating function and then in addition, random fluctuation around this outcome also needs to be modelled.

Another aspect to consider is the correlation between this parameter and other parameters (whether stochastic or not) within the model. The appropriate level of correlation needs to be maintained throughout the whole stochastic modelling exercise. (see 8.2 directly below)

8.2 Modelling dependence
In the solvency example in section 4 the following parameters may have been modelled stochastically:
• On the asset side: Investment return; and
• On the liability side: Claims inflation.

It is important that the stochastic model correctly allows for the relationship between these parameters. For example an increase in investment return needs to be matched with an appropriate increase in claims inflation within the random simulation process. The results of the model need to be checked to ensure that parameter dependence is adequately maintained and allowed for.

8.3 Risk measures
The risk measures discussed above are not exhaustive. Probability of ruin is another key risk measure that can be assessed by means of stochastic modelling. The modeller needs to choose risk measure appropriate for the purpose of the exercise and the communication of the results.

8.4 Data
As with any modelling exercise the available data needs to be sufficiently interrogated and checked to ensure accuracy and quality. Where there is inconsistency and/or insufficiency the assumptions and impact of the assumptions, especially where critical (i.e. if they have a significant impact on the results), need to be clearly documented and discussed as part of the final results communication forum.

8.5 Distribution choice
In any statistical exercise involving fitting a distribution to a given set of data, standard, basic statistical principles need to be applied. This includes performing appropriate goodness-to-fit tests and documenting and communicating the reasons for and appropriateness of the chosen distribution.

In addition to the above, before embarking on a stochastic modelling exercise it is important to consider and address the following:
• The purpose of the exercise;
• The choice of random number generator and appropriate level of convergence;
• The deliverables/outputs required, including the communication strategy of the results;

- The budget, timing and available resources, specifically data, systems/computing power and systems support;
- Specification of the model and decision rules within the model;
- Adaptation of model to unexpected issues;
- Testing and reasonableness of results;
- Overall documentation of stochastic modelling exercise; and
- The costs involved in developing stochastic modelling expertise and capability.

9. CONCLUSION

Stochastic modelling is a useful tool to use when trying to understand the nature of the underlying risk and the potential outcomes that may arise. It provides a lot more information than a deterministic point estimate and in doing so can guide decision makers more appropriately in mitigating and/or eliminating risk.

As with all modelling it is important to understand the purpose of the modelling when choosing which tools (e.g. random number generator) to use and it is essential to understand the shortcomings of the chosen tools and methodology and clearly include these in any communication of the stochastic modelling results.

Stochastic modelling is now very much an integral part of the actuarial toolkit. It is a complex task and as a prerequisite it requires the modeller to have detailed statistical knowledge. However, understanding the process and utilising the results appropriately is well within the grasp of all actuaries and students.

In addition, it assists the actuary in one of the most important aspects of an actuary's work: communication. Stochastic modelling is a useful tool to enhance this fundamental skill as it can visually and graphically illustrate the uncertainty associated with future outcomes.

10. FURTHER RESEARCH

Further research can be done in numerous ways, such as:
- Internet
    - Search engines such as www.google.com
      (searching "stochastic modelling", "asset liability matching", "dynamic financial analysis", "option pricing", "stochastic modelling books")
    - Wikipedia at http://en.wikipedia.org/

- Actuarial notes for Subjects: CT4 Models; and CT6 Statistical Methods

- Professional publications can be found on the Actuarial websites:
    - www.actuaries.org.uk
    - www.assa.org.za
    - www.casact.org

- Software: www.riskmind.com/software.php

- Asking stochastic modeling practitioners for examples and advice

References

Dorrington R E & Rosenburg S B, *Graduation of the 1985-90 assured life mortality experience: A question of maintaining standards.* Presented at the 1996 ASSA convention

Guijarro P, White M *Practical stochastic modelling for life insurers* Presentation at Glasgow Moat House (1 December 2003)

Hockman Y A *Stochastic Modeling...applications for Group Business* Presentation at Munich Reinsurance Company of Africa (8 August 2006)

McCullough B D *The unreliability of Excel's statistical procedure* Foresight (Issue 3 February 2006) (http://forecasters.org/foresight/documents/glance0106.pdf)

The Actuarial Education Company, *Subject 103, Stochastic Modelling* IFE: 2002 Examinations

http://www.actuaries.org.uk/
- *Institute and Faculty of Actuaries exam curriculum*

http://www.assa.org.za/
- *SA85-90 Mortality Tables*

Appendix A: Group life example membership detail

Member	Age	Annual salary	Death Benefit	Mortality rate	Premium
1	24	162,000	486,000	0.00239	1,161.54
2	25	114,000	342,000	0.00228	779.76
3	26	230,400	691,200	0.00220	1,520.64
4	27	204,120	612,360	0.00213	1,304.33
5	28	187,596	562,788	0.00208	1,170.60
6	28	84,000	252,000	0.00208	524.16
7	29	144,840	434,520	0.00205	890.77
8	30	216,468	649,404	0.00204	1,324.78
9	30	232,800	698,400	0.00204	1,424.74
10	31	166,824	500,472	0.00205	1,025.97
11	31	208,020	624,060	0.00205	1,279.32
12	31	220,379	661,137	0.00205	1,355.33
13	32	225,120	675,360	0.00208	1,404.75
14	32	211,932	635,796	0.00208	1,322.46
15	32	83,435	250,305	0.00208	520.64
16	32	191,400	574,200	0.00208	1,194.34
17	33	96,096	288,288	0.00212	611.17
18	33	131,292	393,876	0.00212	835.02
19	33	217,152	651,456	0.00212	1,381.09
20	34	315,600	946,800	0.00217	2,054.56
21	34	181,944	545,832	0.00217	1,184.46
22	34	145,560	436,680	0.00217	947.60
23	35	283,872	851,616	0.00225	1,916.14
24	35	107,676	323,028	0.00225	726.81
25	36	221,424	664,272	0.00233	1,547.75
26	36	326,904	980,712	0.00233	2,285.06
27	36	184,667	554,001	0.00233	1,290.82
28	36	526,680	1,580,040	0.00233	3,681.49
29	36	331,896	995,688	0.00233	2,319.95
30	37	315,660	946,980	0.00244	2,310.63
31	37	122,676	368,028	0.00244	897.99
32	38	330,000	990,000	0.00256	2,534.40
33	38	248,592	745,776	0.00256	1,909.19
34	38	164,688	494,064	0.00256	1,264.80
35	38	158,256	474,768	0.00256	1,215.41

Member	Age	Annual salary	Death Benefit	Mortality rate	Premium
36	38	221,028	663,084	0.00256	1,697.50
37	38	236,952	710,856	0.00256	1,819.79
38	38	223,200	669,600	0.00256	1,714.18
39	39	403,656	1,210,968	0.00272	3,293.83
40	39	291,624	874,872	0.00272	2,379.65
41	39	556,451	1,669,353	0.00272	4,540.64
42	40	127,188	381,564	0.00286	1,091.27
43	41	133,176	399,528	0.00304	1,214.57
44	41	124,200	372,600	0.00304	1,132.70
45	42	128,676	386,028	0.00325	1,254.59
46	42	103,200	309,600	0.00325	1,006.20
47	43	225,096	675,288	0.00350	2,363.51
48	43	81,768	245,304	0.00350	858.56
49	43	134,100	402,300	0.00350	1,408.05
50	43	175,800	527,400	0.00350	1,845.90
51	44	201,923	605,769	0.00377	2,283.75
52	44	236,940	710,820	0.00377	2,679.79
53	44	216,696	650,088	0.00377	2,450.83
54	44	87,036	261,108	0.00377	984.38
55	45	238,667	716,001	0.00408	2,921.29
56	45	552,000	1,656,000	0.00408	6,756.48
57	46	143,603	430,809	0.00443	1,908.49
58	46	94,667	284,001	0.00443	1,258.13
59	46	184,764	554,292	0.00443	2,455.51
60	47	175,356	526,068	0.00482	2,535.65
61	47	190,872	572,616	0.00482	2,760.01
62	47	204,000	612,000	0.00482	2,949.84
63	47	115,920	347,760	0.00482	1,676.20
64	48	327,168	981,504	0.00526	5,162.71
65	48	132,420	397,260	0.00526	2,089.59
66	48	254,280	762,840	0.00526	4,012.54
67	48	137,712	413,136	0.00526	2,173.10
68	49	284,064	852,192	0.00574	4,891.58
69	49	254,868	764,604	0.00574	4,388.83
70	49	318,240	954,720	0.00574	5,480.09

Member	Age	Annual salary	Death Benefit	Mortality rate	Premium
71	49	306,000	918,000	0.00574	5,269.32
72	49	253,404	760,212	0.00574	4,363.62
73	50	166,236	498,708	0.00628	3,131.89
74	50	240,011	720,033	0.00628	4,521.81
75	50	322,980	968,940	0.00628	6,084.94
76	51	87,024	261,072	0.00686	1,790.95
77	52	196,176	588,528	0.00750	4,413.96
78	52	238,020	714,060	0.00750	5,355.45
79	52	128,220	384,660	0.00750	2,884.95
80	53	271,992	815,976	0.00820	6,691.00
81	53	148,595	445,785	0.00820	3,655.44
82	53	131,412	394,236	0.00820	3,232.74
83	53	271,800	815,400	0.00820	6,686.28
84	54	433,608	1,300,824	0.00896	11,655.38
85	54	592,128	1,776,384	0.00896	15,916.40
86	54	171,612	514,836	0.00896	4,612.93
87	55	310,488	931,464	0.00979	9,119.03
88	55	271,476	814,428	0.00979	7,973.25
89	55	136,163	408,489	0.00979	3,999.11
90	55	957,072	2,871,216	0.00979	28,109.20
91	55	169,116	507,348	0.00979	4,966.94
92	56	136,080	408,240	0.01070	4,368.17
93	57	249,264	747,792	0.01170	8,749.17
94	57	166,776	500,328	0.01170	5,853.84
95	57	132,132	396,396	0.01170	4,637.83
96	63	251,304	753,912	0.02027	15,281.80
97	63	865,668	2,597,004	0.02027	52,641.27
98	63	304,032	912,096	0.02027	18,488.19
99	64	294,660	883,980	0.02224	19,659.72
100	66	253,536	760,608	0.02675	20,346.26

Appendix B: SA85-90 mortality table

Age	SA85-90 Mortality Rate	Age	SA85-90 Mortality Rate	Age	SA85-90 Mortality Rate
15	0.00189	45	0.00408	75	0.05894
16	0.00240	46	0.00443	76	0.06411
17	0.00295	47	0.00482	77	0.06974
18	0.00330	48	0.00526	78	0.07588
19	0.00335	49	0.00574	79	0.08259
20	0.00302	50	0.00628	80	0.08992
21	0.00283	51	0.00686	81	0.09792
22	0.00266	52	0.00750	82	0.10664
23	0.00251	53	0.00820	83	0.11613
24	0.00239	54	0.00896	84	0.12645
25	0.00228	55	0.00979	85	0.13764
26	0.00220	56	0.01070	86	0.14976
27	0.00213	57	0.01170	87	0.16286
28	0.00208	58	0.01280	88	0.17699
29	0.00205	59	0.01401	89	0.19220
30	0.00204	60	0.01536	90	0.20842
31	0.00205	61	0.01684	91	0.22580
32	0.00208	62	0.01847	92	0.24440
33	0.00212	63	0.02027	93	0.26425
34	0.00217	64	0.02224	94	0.28538
35	0.00225	65	0.02440	95	0.30784
36	0.00233	66	0.02675	96	0.33161
37	0.00244	67	0.02932	97	0.35671
38	0.00256	68	0.03211	98	0.38311
39	0.00272	69	0.03513	99	0.41078
40	0.00286	70	0.03840	100	0.43966
41	0.00304	71	0.04192	101	0.46966
42	0.00325	72	0.04572	102	0.50067
43	0.00350	73	0.04980	103	0.53255
44	0.00377	74	0.05419	104	0.56513