RECORD OF SOCIETY OF ACTUARIES
1977 VOL. 3 NO. 1

STOCHASTIC APPROACHES TO CORPORATE PLANNING

Teaching Session

FRANK S. IRISH

There are a large number of planning problems in which risk is an important element, although the importance of risk is not always recognized in our planning procedures. Most planning systems require many projections and estimates to be made; these are combined into a picture of the future which is presented as if it were precise, although everyone knows that it is not. Not often enough is the imprecision quantified into probability statements that could convey to the users of the models, forecasts, and plans just how uncertain the outcome is.

As an example, let us start with project analysis, which is the procedure of making cost-benefit estimates for the purpose of determining whether the project should be undertaken. Frequently, in an insurance company, such projects involve the development of computer systems to do something more efficiently. Cost studies produce an estimate of the development cost (investment in the project), and of the recurring savings once the project is implemented. Usually, the savings arise from a reduction in the unit cost of handling items, and an estimate of the number of items (volume of activity) is also important.

The process of combining these data estimates into an evaluation of the project usually treats the estimates as if they were certainties, whereas they are in reality the means of probability distributions. Let us first review some of the favored methods for evaluation:

(1) Pay back period. The length of time required for the accumulated savings to equal the development costs; in other words, the development costs divided by the expected annual savings.

(2) Discounted cash flow. Using an appropriate rate of interest, the outflow represented by the development costs, and subsequent inflow of savings arising from reduced operating costs, are both discounted to produce a present value of net savings from the project.

(3) Rate of return. Using the same type of technique as discounted cash flow, except that the present value of net savings is set at zero, and the interest rate is solved for.

In each of these methods it is possible to produce the answer in probabilistic terms. In doing so, we will be providing decision makers with more information about the nature of the data on which they will base their decision than they would otherwise have. In order to accomplish this, it is necessary to start with probability distributions for each of the estimates that went into the evaluation. Then we must combine these distributions into a single distribution that represents project value.

This combining process is usually sufficiently complicated to make the usual analytic approaches of probability theory inappropriate. It will probably be necessary to use stochastic simulation.

Obviously we are paying a price for insisting on a probabilistic assessment of project value. We have got to ask whether it is worth it to do the extra work that is involved. We have already said that the decision makers will be provided with better information thereby, but what are they to do with it? How does one make the choice between a low risk project of 12% expected return and a high risk project of 18% expected return?

One extreme point of view would have it that the higher rate of return is always preferable, unless we are dealing with a decision of such magnitude that there could be an impact on solvency. But most decision makers obviously don't subscribe to this theory; they accept risky ventures with reluctance. In other words, risky ventures are penalized, either explicitly or implicitly, by being required to bear a higher rate of return. In most cases, however, the judgment on risk is an instinctive one. The transition to thinking in explicitly probabilistic terms is a difficult one.

How then should we establish the link between risk and rate of return? The question is one for which easy answers are not available. The current situation of the company will influence the answer, for it must be remembered that the company is bearing a large and diverse set of risks at the same time, and every additional risk puts a strain on the company's ability to absorb risk. We in the insurance business ought to be particularly aware of this, because we are in the business of accepting risks. In a very real sense, the capacity to absorb risk is a resource, with the classic characteristics of a resource, namely that it has a limited supply, which is very difficult to expand, and it must be allocated among different uses. Thus the "risk premium" that management feels it must obtain on risky ventures is very much a function of the degree to which the company can absorb additional risk, and becomes, in essence, an allocation device in choosing among projects.

Up until now, we have used the word "risk" in a general sense to mean, roughly, possible variations in outcome. There is a distinction to be made, however, within the broad spectrum of our lack of knowledge about outcomes, between the truly random effects, and the variability that is a function of our ignorance about the conditions which will affect the results. Some authors on these subjects make the distinction in the following way. Risk, they say, describes the situation where there is a variability of outcomes, but where the probabilities of various outcomes can be calculated with fair degree of accuracy. Uncertainty exists where the probability distribution itself is not well known.

Mathematically, the distinction is not evident. Probability distributions about future events contain varying proportions of true randomness and ignorance, and both serve to increase the variance of the distribution. The theorists of personalistic probability have made it quite clear that our own lack of knowledge can be treated in a probability distribution quite effectively. But operationally, the distinction is an important one, for we can do nothing about reducing the inherent randomness of a situation, but we can do a lot about reducing our ignorance of a situation.

The planner should not be averse to risk. The acceptance of risk is usually profitable, and as long as the total amount of risk does not exceed the capacity of the company to absorb risk, he should be willing to accept risk and deal with it openly. But it is part of the planner's job to reduce uncertainty (or "ignorance" if you prefer the term), and in the present context this means sharpening up his knowledge of the probability distributions.

This goes right back to the beginning of the planning problem we stated at the outset. It raises the question of how to get good probability distributions for future events. Unfortunately, the "experts" who can give us the best information about particular situations are not usually very good at framing their knowledge in probabilistic terms. Very few people are, as a matter of fact. Thus, the techniques of eliciting probability statements become very important.

To provide examples of the difficulties of eliciting probability statements, some very small surveys were made within my own company. Sales projections were used in the example, because they are well understood and because people are usually willing to give an opinion on the subject.

For instance, Product A is a minor line of business that was completely revamped at the beginning of this year. Respondents were asked to make estimates of the percentage increase in sales for this year over last year. The request was phrased in very precise terms, with the aim in mind not of producing an authoritative "projection" as it usually is thought of, but of producing a probability distribution that recognizes the randomness and uncertainty of sales projections. Respondents were first asked to estimate the most likely result. Then they were asked to make a range estimate, which was couched in terms of personalistic odds; they were asked to quote a number such that they would bet three to one that sales would be at least that number.

You will recognize that these latter questions were designed to elicit the quartiles of the distribution. The concept of phrasing the question in terms of a personal bet is one that has often been recommended in this type of survey, and it seems to work well. The question about "most likely" value, if taken literally, means the mode of the distribution. It seems probable, however, that some of the participants responded with the median or mean of their personal distribution, and in this respect the survey could have been better designed.

Respondents were also asked for a frank evaluation of their degree of knowledge about this particular product, relative to others in the company. Their answers were weighted 1, 2, or 3, depending on the response to this question.

If we were going to use the results of this survey as input to a stochastic model, we would, at this point, wish to fit a distribution to the weighted average results, which were 34% (most likely), 17% (lower quartile), and 76% (upper quartile). Note the wide spread of the quartiles which reflects the willingness of the respondents to admit to a high degree of uncertainty in the outlook. Also, the distribution is highly skewed, which seems to be quite natural in a situation like this.

In fitting this kind of distribution, a modification of the normal distribution is often used. One distribution which is particularly useful in portraying this kind of situation is the lognormal:

$$f(x) = \frac{1}{\sigma x \sqrt{2\pi}} \exp\left[-\frac{(\log_e x - \mu)^2}{2\sigma^2}\right] \qquad x > 0$$

Note particularly the x in the denominator which permits the logarithmic transform to preserve the cumulative distribution function. The lognormal is particularly useful in portraying distributions which are skewed to the right and which are essentially non-negative. It happens, unfortunately, that the lognormal is a two parameter distribution, while here we are looking for a three parameter distribution, but this can usually be handled without significant harm to the data by a simple shift transformation $y = (x - A)$.

If you wish to get more deeply into the technical aspects of distributions, you may wish to look into the Gamma and Beta distributions. These are non-negative, and very flexible as far as skewness is concerned, and they do have their uses, but probably not in stochastic sampling, because of the difficulties in deriving the numerical values for the cumulative distribution functions. There are also some very unsophisticated distributions which would be acceptable in this particular example; such as the very simple triangular distribution, which works better than one would think.

Another survey of the same type was taken on what we shall call Product B. This is a completely new product, of a type with which the company has had only a little previous experience. Respondents were asked to frame their answers in terms of a certain type of "sales unit" rather than percentage increase. The weighted average answer was 14 units for most likely, and quartiles of 6 units and 30 units. Under most assumptions about what a "reasonable" distribution is, it would be impossible to fit a reasonable distribution to this data.

The wide spread of the quartiles for Product B, and the diversity of answers received on both Product A and Product B present some very interesting data for the statistician, but the situation should be a little disquieting to the planner. If these data were to be used as a basis for any important decisions they would appear to be inadequate. It should be worth our while to do what we spoke of earlier -- "sharpening up the probability distributions". In the case of survey estimates of future events, the best known technique for improving consensus is called the "Delphi method". The essence of this method is holding several rounds of estimating with the same group of experts. In the second round, respondents are told of the average response and asked if they wish to change their answer in the light of it; or, if not, to defend their answer in writing. This may lead to further rounds in which those who staunchly defend their answers convince the others, or are convinced, to change their answers. The validity of the method in producing a more accurate consensus than existed at first has been empirically demonstrated to the satisfaction of most who have studied it, although the technique does have its detractors. It is not as easy a technique as it first appears, and most of its failures, say its proponents, are due to improper technique.

Let us turn now to another type of planning problem; the problem of common stock policy making, as an example of stochastic approaches. The situation of life insurance companies with regard to common stock investments is rather unique among investing institutions, and we must reflect some of these unusual features in our model. For example, common stocks are valued at market whereas most of the other investments of the company are valued at book value. There is a mandatory securities valuation reserve which is much more of a burden on common stock investments than on other types of investments. But the mandatory reserve also serves as a buffer between the appreciation of common stocks and their effect on surplus. As a result, common stocks produce, in most years, a very low rate of return, in terms of "available" funds; but occasionally produce a very large chunk of (positive or negative) surplus contribution. Thus, common stocks present a very good problem for stochastic modelling, because the long run impacts of common stock policy decisions on corporate risk are probably impossible to derive by analytic methods.

To add to this we have the fact that the tax situation of common stocks vis-a-vis other investments are rather strange. Dividends are taxed at a low rate. Capital gains are taxed at a fixed rate which may be higher or lower than the marginal rate on other types of investment return. Therefore taxes also make up an important part of the model.

In the creation of such a model, one must construct quantitative relationships that reflect all these variables. The only random variable in one model that I know of was the common stock performance, which was expressed as a percentage appreciation. It was also found necessary to introduce pricing policy equations, since it was obvious that no management was simply going to let its surplus go up and down without some reaction in terms of pricing. Other than that, the insurance activities of the company were modelled extremely simply, in one or two equations, which told us how much cash flow would be available for investment.

In operating a stochastic time-process model such as this one, the basic principle in handling the random variable is to use a random number table or generator to determine the value of the random variable in each time period. Typically the events in each time period are dependent in part on the events of the preceding time periods; for some, these are the characteristics that define a stochastic process, although I use the term more broadly. In such a case, it is necessary to cycle through the model as many times as the number of periods being projected, determining the random variable at each cycle.

A question that frequently comes up is how many projections to make in order to get a valid spread of results. Beware of easy answers to this question; there are none. In part, the answer will depend on what statistic it is that you are interested in. If it is the mean or expected value of the results that is of interest, the number of runs needed might be relatively small. Unfortunately, most stochastic modelling is not done for the purpose of finding a mean value.

In the present case, for example, we are interested in the amount of fluctuation in surplus that can be expected to occur as a result of common stock operations; and, in particular, we are interested in the probability of a serious downward fluctuation. Let us say, for example, that the statistic we are interested in is the probability of surplus dropping below 3% at any time in a 20 year projection. The number of projections we might have to do to get a dependable value for this probability could be quite large. But how do we know when we have a dependable value? One approach is to track the value of the statistic of interest as more and more runs are done, to determine when the value of the statistic begins to stabilize.

In most cases, random number generators used for a model like this produce numbers which are distributed according to a rectangular distribution,(evenly distributed between zero and one) and it is necessary to translate that distribution into the one chosen for the random variable. This is often done by a percentile method - that is, the percentiles of the distribution are listed so that the rectangular distribution can be read directly into the desired distribution. Listing the percentiles is cumbersome, but there are some packages which do this for the normal distribution; as a matter of fact there are random number generators that produce random normal deviates directly. Thus it is often of value to use a normal distribution, or one of the distributions that can quickly be derived from it, such as the lognormal or the series expansion.

The question of the proper distribution to use for common stock appreciation is one that has received much attention in the literature. There is general agreement that the logarithm of the rate of appreciation produces a fairly symmetric distribution, and a fairly good fit may be obtained with the lognormal distribution. There is evidence of some slight leptokurtosis (positive fourth moment), however.

Another question that is properly brought up in connection with any model, stochastic or otherwise, is sensitivity testing. We are always in a position to say, with regard to the probability distributions and the quantitative relationships that are used in the model, that further research could produce a better model. If we run the model with several different types of variations in the assumptions to determine which variations may have an effect on the ultimate decision (in other words, to determine which assumptions the model is "sensitive" to), this can give us guidance as to whether further research is worthwhile and what direction that further research should take.

Another type of stochastic model that involves randomized investment assumptions is an investment guarantee model. The purpose of the model is to determine the amount of risk involved in granting investment guarantees of varying types. Here we are principally concerned about fixed dollar returns, and therefore about predicting the future course of long term interest rates (or, more accurately, specifying a probability distribution for future long-term interest rates). Here is an area where there is little agreement.

The picture is complicated if, as is frequently the case, the customer has the right to some kind of withdrawal at book value, which can cause losses if the market rate of interest happens to be high at the time the withdrawal right is exercised. Thus the model is sensitive to the kind of withdrawal rights granted, and to the pattern as well as the trend of interest rates.

There is a large school of thought which holds that random-walk is as applicable to interest rates as it is elsewhere (most notably in stock market returns). One presumes that they admit serial correlation and cycles into the outlook, because interest rates obviously have long "runs" of ups or downs. Then there are those who react to the inflationary bias in our economy and foresee a rapid and irregular upward trend (the ratchet phenomenon). On the other side are those who have studied the causes of past movements, and feel we are about to fall off the top of an interest rate cycle (the Kondratieff cycle). The most that we can do in such an uncertain situation is to inform ourselves of the impact of these assumptions on the financial results of our decisions, and hope that our decision making is improved. To do less might result in a ludicrously inaccurate assessment of the risk.

As a final topic, we are going to take just a brief look at some techniques that open up a whole new way of dealing with uncertainty. The subject of Bayesian inference is based upon our old friend Bayes' theorem, which is in Chapter One of almost every first year course on probability. Most of us learned it there, and then forgot about it, for it seemed to have no relation to what came after.

$$P\{A_i | B\} = \frac{P\{B | A_i\} P\{A_i\}}{\sum_i P\{B | A_i\} P\{A_i\}}$$

We are, however, aware that in recent years, Bayes' theorem has come to have some significance in the analysis and graduation of mortality statistics. But it also has some wider uses.

Consider, for example, that the A_i in the above equation are possible (mutually exclusive) future events, and that B is information gained about the future as a result of research. The probability of A_i in the right-hand side of the equation is known as the "a priori" distribution and represents the state of our knowledge before we receive the information B. The left side, or "a posteori" distribution represents the sum of our knowledge after learning B. By using Bayes' theorem, we can combine data from disparate sources to produce a tighter probability distribution. We can even make an estimate of whether it is worth doing further research to improve our distribution.

The term "Bayesian" has broad applicability. Any time that we embrace, in our decision making procedures, the concept that the "world out there" is stochastic and that the situation can be described in probabilistic terms, we call the procedures Bayesian. In its simplest form, a Bayesian strategy is that strategy which maximizes

$$\sum U(S, A_i) P\{A_i\}$$

where U is the utility of the strategy S if A_i occurs. As before, $P(A_i)$ is our prior estimate of the probabilities of an event, a function frequently called the "plausibility function" to emphasize that in practical applications it is usually not derived from classical statistical procedures.

If we are about to do the research that will provide us with new information B, we anticipate that the new distribution of A_i may change our choice of optimum strategy. This can be expressed as:

$$\omega_i = \sum U(S', A_i) P\{B | A_i\}$$

which is the weighted value of the optimum strategy S' (whose choice now depends on B) if A_i occurs, and

$$V = \sum_i W_i P\{A_i\}$$

which is the Bayesian value, using our current estimate of the distribution of A_i. The procedure is important for two reasons; for not only have we decided how to choose a new strategy in the light of new information, but we have made an estimate of the value of getting the new information (the excess of the new value, V, over the original maximum U).

A brief bibliography:

Aitchison, John Choice Against Chance
Addison - Wesley (1970)

Burington and May: Handbook of Probability and Statistics with Tables
McGraw Hill (2nd Ed. 1970)

Fama, Eugene F. "Behavior of Stock Market Prices"
Journal of Business, January, 1965

Hertz, David B. "Risk Analysis in Capital Investment"
Harvard Business Review, Jan.-Feb., 1964

Linstone and Turoff: The Delphi Method
Addison - Wesley, 1975

Starr, Martin K. Management: A Modern Approach
Harcourt Brace Jovanovich, 1971