Section B
DESCRIPTION OF STOCHASTIC MODELS

[B1]
WHAT IS A STOCHASTIC MODEL?

Section 2B of the Supplementary Introduction to Volume 1 gives a general description of reserving methodology. In that description, the process of arriving at an estimate of future payments is described as one of constructing a model, fitting it to some set of past observations, and using it to infer results about the future — in this case, the future events we are interested in are the payment of claims. Several distinctions are made between different types of model, including those between deterministic and stochastic models.

Deterministic reserving models are, broadly, those which only make assumptions about the <u>expected</u> value of future payments. Stochastic models also model the <u>variation</u> of those future payments. By making assumptions about the random component of a model, stochastic models allow the validity of the assumptions to be tested statistically, and produce estimates not only of the expected value of the future payments, but also of the variation about that expected value.

All the methods in Volume 2 could be described as stochastic to a greater or lesser extent. One can distinguish between them a little, since the methods described in Sections D1, D4, D5, D6 and D7 all allow the user to make estimates of the variation about the expected future payments. The methods described in sections D2 and D3, however, simply involve the fitting of curves to sets of data. The curves are then used to predict future payments, but do not allow the modeller to make estimates of the variation of these payments.

A further distinction can be made between those models based on individual claims, and those which project grouped claims data. This distinction is most commonly found amongst stochastic methods, although the only methods presently in Volume 2 which model individual claims information are those explained in Sections D4 and D7.

<>

[B2]
WHAT ARE THE ADVANTAGES / DISADVANTAGES OF A STOCHASTIC MODEL?

This section briefly highlights some of the perceived advantages and disadvantages of stochastic models, to give the reader some idea of their strengths and weaknesses.

Section 2B of the Supplementary Introduction to Volume 1 observed that deterministic models may often be applied without a clear recognition of the assumptions one is making. One of the main benefits of a stochastic model is that it is totally explicit about the assumptions being made. Further, it allows these assumptions to be tested by a variety of techniques. Because it models the random variation of future payments, estimates may be made of the likely variability of the estimated future payments.

This allows one to monitor whether the predictions of a model are within the bounds one would expect. For example, a deterministic model simply makes a point estimate of the expected future payments in a given period. The one sure thing one can say about these expected payments, is that the actual payments will be different from expected. Deterministic models do not give you any idea as to whether this difference is significant. Stochastic models enable the modeller to produce a band within which the modeller expects payments to fall with a certain level of confidence, and can be used as an indication as to whether the assumptions of the model hold good.

The strengths of stochastic models can also be their weaknesses.

A stochastic reserving method models an immensely complex series of events with a few parameters. Hence, as with any model, stochastic or otherwise, it is open to the criticism that its assumptions are far too simple and hence unrealistic. Because stochastic models are quite clear and rigid, there is very little scope for incorporating judgement, or extraneous factors into the model.

Finally, stochastic models can be computationally quite complex to perform, and may require a more in-depth statistical and computational ability than some of the more simple deterministic models. This in turn can mean that the results are more difficult to communicate than some of the more simple deterministic models.

<>

[B3]
WHAT MAKES A GOOD STOCHASTIC MODEL?

To appreciate what makes a good stochastic model, it is necessary to understand why one constructs a model in the first place.

Take as an example a set of data on (say) motor claims. This may consist of tens of thousands of claim payments, extending over a number of years. If we simply record each of those amounts individually on reams of paper, the human mind simply cannot grasp the essential characteristics of the data, or discern any pattern, let alone use the data to make sensible predictions.

To understand the data in any meaningful way, therefore, requires the formulation of a pattern that in some way represents the data. In this way, the important characteristics of the data can be represented by a limited number of terms that can be relatively easily understood.

Further, when considering any set of data over time, there will be some systematic influences affecting the claims experience, such as the inflation in the cost of car repairs in our example. There may also be some random influences, such as the variation in the frequency of cars having accidents. To understand the data effectively, one needs to differentiate between systematic influences and random variation.

It is this need to reduce complexity and to separate systematic influences from random variations that leads to a stochastic model. A stochastic model allows the modeller to replace the individual data values by a summary that both describes the essential characteristics of the data by a limited number of parameters, and distinguishes between the systematic and random influences underlying the data.

The parameters of a model are chosen to "fit" the data as closely as possible. The fit can be made better and better by having more and more parameters. However, this then becomes self-defeating, as a model with hundreds of parameters provides no real reduction in complexity from the raw data, and allows the user only a limited ability to grasp the key characteristics of the data.

An essential requirement of a good model, therefore, is that it has enough parameters to describe the characteristics of the data, but not so many that its descriptive power becomes limited. Additionally, as described in the first paper in Volume 2, as you increase the number of parameters of the model, you decrease its predictive power. That is, the model begins to adhere more and more closely to the raw data. Small changes in those data can then lead to large changes in the parameters of the model, making any predictions produced by the model unstable.

A good stochastic model should also enable one to appreciate the systematic influences underlying the data, together with the random influences. Some data points may be subject to considerable random variation, so the model should ensure that it is not unduly affected by such isolated values. A good stochastic model should therefore be capable of testing the underlying assumptions. By applying such tests, the modeller will gain a greater understanding of the characteristics of the data and, hence, have better control over the projected values.

The above points are, of necessity, fairly general in nature, as any sort of modelling is as much an art as a science. To this end, it is worth observing what it takes to be a good modeller.

The first and most important requirement is to appreciate that all models are “wrong” to some extent. They are not “reality”; they are just a simplified representation of reality, enabling the user to make practical projections of the data. As a consequence, there is no one “right” model, and many different models may be more or less equally applicable.

So, the second requirement is that a good modeller should consider many different models, trying to recognise all those that might be useful, rather than whether they are “right” or “wrong”.

A final requirement of the modeller is that they should check the fit of a model. The object of this exercise is to understand the past data, and to infer useful results about the development of those data. This cannot be done rigorously if the modeller does not understand where the model fits or deviates from the data.

◇