The Actuary's Toolkit: A View from EMB

P.D. England

EMB Consultancy LLP, Saddlers Court, 64-74 East Street, Epsom, KT17 1HB.
peter.england@emb.co.uk
http://emb.co.uk

ABSTRACT

In this short paper, some of EMB's software tools that have been developed to help solve business problems are described, from the perspective of the new and emerging statistical and risk management techniques that have been incorporated. In addition, the challenges of developing actuarial software, and actuarial education, are discussed.

KEYWORDS

Bayesian, Bootstrap, Demand Modelling, Dynamic Financial Analysis, Generalised Linear Models, Internal Capital Models, Markov Chain Monte Carlo, Parameter Uncertainty, Predictive Modelling, Price Optimisation, Pricing, Simulation, Spatial Smoothing, Stochastic Reserving.

1. INTRODUCTION

Just over twenty years ago, a typical desktop computer used by actuarial trainees had 1Mb RAM, a 40Mb hard disk, and a clock speed of 12MHz. Basic spreadsheets were in use, and largely responsible for radically overhauling working practices in all areas of finance. Today, actuarial trainees employed by EMB use desktop computers with dual-core processors, 4Gb RAM, 2 * 250Gb hard disks, and clock speeds counted in GHz. The increase in computer power over the last two decades has simply been staggering, and this can only be expected to continue.

At the same time, actuaries and statisticians have found increasingly innovative ways of using computers to tackle business problems. For example, twenty years ago, motor business was priced using a series of one-way tables, with various *ad hoc* adjustments. Motor business is now priced routinely using generalised linear models with multiple explanatory variables for frequency and average severity separately for each peril type, before combining (this is known as predictive modelling in the USA). Data sets typically run into millions of observations. The pricing models can be combined with other models that measure price-demand elasticity and the results used to numerically optimise profit or growth.

Another area where the power of computers has been harnessed to great benefit is simulation modelling for assessing risk. When assessing the need for reinsurance to cover natural catastrophes such as windstorms, floods, and earthquakes, and to price that reinsurance and investigate risk aggregations, simulation models are routinely used to simulate the frequency and severity of catastrophe events. The use of simulation models is not even questioned

when assessing the impact of such rare events. Although it has taken longer for simulation based methods to be accepted in other areas, they are now routinely used for: reinsurance analysis (especially for complex reinsurance programmes); assessing the impact of different states of the economy on asset and liability portfolios (using Economic Scenario Generators); and capital modelling, where financial simulation models of an insurance enterprise are built that provide simulated future cash flows, enabling future financial statements to be simulated, and the probability of liabilities exceeding assets (ruin) to be assessed.

Before the advent of simulation based methods, many actuaries (particularly those of an academic persuasion) invested vast amounts of effort deriving mathematical formulae to solve given problems, thereby spawning the field of "risk theory". However, many of the problems that were tackled analytically were far simpler than those encountered in practice, and the problems could very quickly become very complicated, and soon become intractable to solve analytically. Thankfully, progress can often still be made using simulation based approaches, allowing far harder and more realistic problems to be tackled. As Daykin, Pentikainen and Pesonen (1996) state:

> *"Modern computer simulation techniques open up a wide field of practical applications for risk theory concepts, without the restrictive assumptions, and sophisticated mathematics, of many traditional aspects of risk theory".*

2. EMB's ACTUARIAL SOFTWARE

EMB Consultancy LLP is primarily an actuarial and management consulting firm. However, EMB also develops software to help tackle the business problems that arise, and develops the software in such a way that it can be marketed and used by its clients. This demands added rigour in the software development process. EMB has pioneered the development of actuarial techniques in several areas, and has harnessed the power of computers to make those techniques available to practitioners.

EMB currently has 8 software products:

Igloo Professional	Financial simulation framework for risk modelling
ExtrEMB	Dynamic parameterisation for risk modelling
ResQ	Loss reserving software
EMBLEM	Pricing software for personal and small commercial lines
Rate Assessor/Optimiser	Business decision support for pricing
Classifier	Geographical classification of risk
RePro	Visualisation and MI system for Excess of Loss Reinsurance
PrisEMB	Pricing software or commercial insurers and reinsurers

The purpose of this paper is not to describe each product in detail, but simply to highlight the technological developments that have been included or utilised in developing the software. Some of these developments for some of EMB's software products are described below.

2.1 Igloo Professional and ExtrEMB

Igloo Professional is a simulation framework, primarily used for simulating future cash-flows and financial statements for use in capital modelling and pricing exercises. It is a framework, not a model: that is, models are built using Igloo Professional for tackling complex business problems. The fundamental philosophy is that every company is different, and every problem is different, so a model should be built to fit the problem. A problem should not be expected to fit a model, so the concept of providing a fixed model that attempts to solve all problems is not a concept that EMB follows. Igloo Professional provides the functionality to allow simulation models to be built quickly and easily, allowing the user to focus on the business problem, and not, for example, the technicalities of how to write an algorithm to simulate from a particular distribution.

Igloo Professional provides forecasts conditional on given parameters for the statistical distributions. A large number of statistical distributions are available, including extreme value distributions (see, for example, Embrechts et al (1997) and Klugman et al (2004)) When licensed with Igloo Professional, ExtrEMB provides additional functionality allowing the parameters to be estimated from data, completing the path from data to forecast. Parameterisation is performed primarily using maximum likelihood techniques. Where possible, this is achieved analytically, but often numerical maximisation techniques are required. ExtrEMB includes a variety of maximisation algorithms that are useful in different circumstances, and have various features, for example, maximisation subject to constraints. The maximisation routines are general, so they can be used to maximise any objective function, not just log-likelihoods. For example, they can be used to optimise (or simply improve) reinsurance programmes or business mix in the context of a capital model by maximising return on capital. Numerical optimisation in the context of simulation modelling has only been feasible recently, since the demands on computer power are high.

Igloo Professional can apply dependencies between simulated variables in several ways. One standard way is simply relating variables causally through a mathematical formula. An alternative is to use copula theory, which is a natural candidate in insurance modelling. Basically, a copula separates a joint distribution into the underlying marginal distributions and a dependency structure. For an introduction to the use of copulas, see Frees and Valdez (1998). Igloo Professional supports several copulas in a very easy to use way, including copulas that allow tail dependency. The use of copulas is becoming increasingly popular in finance and actuarial science.

When building models to assess capital sufficiency, three fundamental components are required: a risk profile, a risk measure and a risk tolerance criterion. A popular risk profile is distribution of economic net assets at some future time horizon, although other alternatives are commonly used, for example distribution of profit over some time period. Once a risk profile has been obtained using simulation techniques (often by building a complex model of an insurance enterprise), it is straightforward to apply a risk measure at a given risk tolerance level to assess capital requirements. The choice and characteristics of risk measures has received considerable attention in recent years. A popular risk measure is “Value at Risk” (VaR), which is simply a percentile in statistical terminology. The weaknesses of VaR are well known, and a popular alternative is to use “Tail Value at Risk” (TVaR) which is an example of a coherent risk measure, popularised by Artzner et al (1999). In a simulation world, coherent risk measures can be applied simply as a weighted average across all simulations, where a particular risk measure dictates the definition of the weights used. For

example, TVaR weights are equal above a given centile, and zero below, whereas for Wang's proportional hazards transform (see Wang(1995)), each simulation is given a different weight, where the weights increase with the underlying values. Igloo Professional supports a variety of risk measures.

Once capital requirements have been assessed, it is common to think about allocating capital back to lines of business, insurance contracts or risk types. There is a variety of allocation methods that can be applied, which are appropriate in different circumstances. For example, methods that are appropriate for risk management may not be appropriate for pricing. One advantage of coherent risk measures is that they are also a natural candidate for use in a capital allocation exercise for risk management, since allocated capital sums naturally to total capital (see for example Artzner (1999)). Igloo Professional supports a variety of methods to allocate capital.

Capital models require many other components, for example, Economic Scenario Generators (ESGs), catastrophe modelling facilities, reinsurance default modelling, reserve risk modelling and so on. All of these can be supported, built and calibrated using Igloo Professional, although a description is beyond the scope of this document.

An additional feature of ExtrEMB that is worth mentioning, however, is the ability to investigate parameter uncertainty using Bayesian techniques (for a discussion of parameter uncertainty in insurance, see Cairns (2000)). ExtrEMB includes Gibbs sampling (Geman & Geman (1984)), a Markov Chain Monte Carlo (MCMC) technique for obtaining simulated distributions of parameters for multiple parameter problems. Gibbs sampling is used in ExtrEMB in conjunction with Adaptive Rejection Sampling (see Gilks & Wild (1992)) and Adapative Rejection Metropolis Sampling (see Gilks et al (1995)). Simulation based MCMC techniques for investigating parameter uncertainty are required since alternatives based on the asymptotic distribution of parameters from likelihood analysis are often inadequate. In addition to parameter estimation using maximum likelihood techniques, ExtrEMB includes prebuilt modules for investigating parameter uncertainty for the commonly used loss distributions. ExtrEMB also includes prebuilt components for investigating reserving risk using Bayesian techniques, again using Gibbs sampling.

2.2 ResQ Professional

ResQ is EMB's loss reserving tool, primarily for use in analysing reserving run-off triangles. Input data might be, for example: paid claims, incurred claims, reinsurance payments, expense payments, numbers of claims, premium amounts and so on. A number of standard exhibits are included for viewing data graphically, and quickly manipulating data into different formats. A range of standard deterministic techniques for obtaining a best estimate of outstanding liabilities is also included, such as: Development Factor (Link Ratio) methods, Bornhuetter-Ferguson, Cape Cod, Berquist & Sherman, Average Cost Per Claim, Payment per Claim Finalised, Fisher-Lange and so on. Once models are fitted, there is a variety of diagnostic checks that can be performed for assessing model adequacy.

The technical developments incorporated into ResQ that are noteworthy concern methods to investigate the uncertainty in the estimates of outstanding liabilities, and in particular, methods to simulate predictive distributions of the outstanding liabilities, for use in capital modelling. Interest in stochastic reserving techniques has increased steadily over the last 20

years, although it is only relatively recently that it has developed into a discipline in its own right, and has spawned several PhD theses, and is even included in the actuarial exam syllabus in the UK and the US. The main driver of the recent interest has been the availability of simulation based methods that are easy to implement with the help of a computer, and do not require an understanding and implementation of complex analytical formulae.

ResQ Professional includes a Bootstrap module, which can be used in association with Development Factor Methods to provide a simulated distribution of outstanding liabilities. Bootstrapping can be applied to two underlying models: the over-dispersed Poisson model (ODP, with a single scale parameter or varying scale parameters), and Mack's model. Mack (1993) proposed an analytic formula for the standard deviation of the outstanding liabilities associated with the chain-ladder model. EMB has developed this further, demonstrating how to bootstrap Mack's model to provide a simulated predictive distribution of the outstanding liabilities (see England & Verrall (2006)). Bootstrapping provides not just distributions of outstanding liabilities for each origin year and in total, but also distributions of the cash-flows, which can then be used to provide distributions of the discounted outstanding liabilities. ResQ Professional includes the facility for discounting the distributions of cash flows using a term structure of interest rates.

Although bootstrapping provides distributions by line of business, companies are often interested in a distribution of the aggregate outstanding liabilities across multiple lines. ResQ Professional includes the facility to aggregate the simulated results, imposing dependencies (correlations) between lines of business using a Gaussian copula.

From a computing perspective, a key benefit of ResQ Professional is that it supports COM, so it is straightforward to interface the results with a management information reporting system, and also to automate the whole of the quarterly reserving process, from data entry to modelling and reporting.

2.3 EMBLEM

EMBLEM forms the key component of EMB's personal and small commercial lines pricing suite. It is essentially a data mining/modelling tool, and fits Generalised Linear Models (GLMs) very quickly to extremely large data sets. The types of models that are routinely fitted now were unimaginable 20 years ago. In fact, the use of GLMs for motor premium rating only became routine after publication of Brockman & Wright in 1992 (Brockman being the "B" of EMB).

GLMs have been used in statistical modelling for some time (see for example, McCullagh & Nelder (1983) and (1989)), and the algorithm for fitting GLMs is well known. Several mainstream statistical packages are capable of fitting GLMs.

So what makes EMBLEM different? The answer lies in the volumes of data that can be modelled in EMBLEM, and the speed with which they can be fitted. The algorithms used for fitting GLMs and memory management have been highly optimised to allow data sets with millions of observations to be modelled in seconds (or minutes with extremely large data sets). Most standard statistical software used for fitting GLMs is simply incapable of

modelling the size of data sets typically used in a pricing exercise. Even when it can, EMBLEM can fit models 40 times faster than competing software.

In addition, EMBLEM has been designed with an easy to use graphical user interface, allowing the modeller to focus on the business problem, not the technical difficulties.

EMBLEM includes the standard set of link functions and error structures that are used with GLMs, together with some non-standard ones. EMBLEM also includes a set of macros that materially speed up the modelling process.

In terms of additional technical developments, EMBLEM also has the facility to fit regression splines easily, with interactive knot placement. Although it is not well known, the underlying GLM algorithm can also be used to fit non-linear models (models that are non-linear in the *parameters*), with a bit of extra information provided by the modeller. EMBLEM has been extended to allow non-linear models to be fitted.

Typically in a pricing exercise, claims are split by peril type. For example, motor claims may be split into: property damage, bodily injury, theft, windscreen and fire. For each peril type, the frequency and average severity of claims are modelled separately and a parsimonious model selected for each component. Finally the sum across peril types of the product of the frequency and average severity components is formed to give the "pure risk premium".

GLMs can also be used to model the price demand elasticity (using logistic regression; that is, using a Binomial error structure). If the change in price relative to last year's premium is recorded, and the competitiveness relative to the market can be ascertained, then the probability of renewal (or conversion for new business) can be modelled using a variety of additional policyholder factors.

From a computing perspective, a further innovation in EMBLEM is the inclusion of *Visual Basic for Applications* (VBA). This allows macros to be written or recorded to automate a variety of tasks, and even implement new methods and extend the graphical user interface. This is an extremely powerful feature that has many uses.

2.4 Rate Assessor/Optimiser

Rate Assessor is a rate-setting decision support tool used for pricing personal and small commercial lines to explore how pricing strategies can impact market share, retention, new business and current/future year profits. It is designed to be used in association with EMBLEM, although data can be imported from a wide variety of systems and models. The main use of Rate Assessor is for moving from a "pure risk premium" to a street price.

The power of Rate Assessor is maximised when it is used in conjunction with Optimiser. Optimiser combines results from the pure risk premium analysis and the price/demand elasticity models to provide highly sophisticated price optimisation at the policy and portfolio level. Optimiser finds optimal prices by systematically combining cost models, demand models and market prices. It recommends policy level price revisions across an entire portfolio, to maximise profit and/or growth.

Essentially, Optimiser employs numerical maximisation routines to find the recommended policy level price revisions. Again, this type of analysis was inconceivable a few years ago, and has only been made possible by the availability of increased computer power.

2.5 Classifier

Classifier provides detailed categorisation and assessment of risk by geographical location (post code/zip code), which has been a notoriously difficult challenge in the past. It is used in combination with pricing models, such as those created by EMBLEM, to deliver a holistic predictive model reflecting both geographic and non-geographic factors.

Classifier assesses geographic risk differentials using colour coded maps. Of course, it is hoped that all of the socio-demographic information that may impact geographic risk has already been included in the underlying EMBLEM models, resulting in a map with a uniform colour in Classifier. However, this is rarely the case, and some geographical risk usually remains. Classifier is designed to identify and analyse that residual risk.

Apart from the use of mapping software, the main technical developments included with Classifier are spatial smoothing techniques, using distance-based and adjacency-based methods, such as enhancements to Boskov & Verrall (1994). These are analogous to image processing techniques, and enable the "signal" to be separated from the "noise" more effectively, to enable the geographic effects to be seen more clearly. Significant competitive advantage can be achieved by analysing how geographic location affects claims experience.

3. DISCUSSION

The task of developing easy to use actuarial software that keeps up with the pace of developments in computing, and the associated developments in statistical techniques, should not be under-estimated. EMB has learnt that although there are some actuaries who are very good programmers (or some programmers who happen to be actuaries), on the whole, it is better to use professional programmers and software developers when developing software for actuarial use, which has suitable levels of reliability and quality assurance. In addition to programmers and developers, EMB employs professional testers to test software to destruction as it is developed. New software products and software upgrades go through a rigorous quality assurance procedure before being used by EMB consultants internally. It is only made available to EMB clients after it has had sufficient use by consultants.

In the early days of computing, when computer power was minimal, users of computers had to think carefully about efficiency when writing computer programs. On the whole, computers are now more than adequate for most actuarial tasks, so users have in some ways become lazy when developing spreadsheets and simple programs, since it is rare to reach hardware limits. When faced with more complex problems involving modelling large data sets, or simulation models generating large datasets, users are again sometimes faced with hardware limits, and increasingly innovative ways need to be found to overcome those limits. In addition to the requirement for creative thinking when developing the software itself, creative thinking is sometimes required when using the software. Students coming into the profession are now young enough never to have known life without computers for everyday

tasks, and are sometimes surprised when a computer cannot do what they want due to hardware constraints. Hopefully some of these constraints will be overcome with the advent of 64bit computing, but users will only try ever more complex tasks.

There is a continual challenge to educate actuaries and insurance professionals in new techniques as they arise, and that education process is often facilitated with the availability of easy-to-use software. In fact, pioneering techniques often only become mainstream when suitable software becomes available, or the techniques are explained in an easy to follow manner. In his Guest Editorial in the Annals of Actuarial Science, Angus Macdonald asks the question "Actuarial Education: Does it Compute?", and provocatively discusses the relevance of exams written using pen and paper for the practising actuary. Quoting Macdonald (2006):

> *"The student who is introduced in the classroom to the bootstrap, or simulation, or GLMs, or MCMC estimation, gets the chance to play with them in the computer lab … Those actuarial professions whose qualifications are rigidly based on written examinations have not made the leap into the modern world."*

It is not the purpose of this paper to comment on the suitability or otherwise of written exams; it is simply interesting to note that all of the technical developments in Macdonald's view of the "modern world" have been incorporated into EMB software, and described in this paper. The advent of desktop computers and the rapid increase in computer power has radically changed the day-to-day work of most general insurance actuaries, so much so, that we could say, like L.P. Hartley (1953):

"The past is a foreign country; they do things differently there."

REFERENCES

Artzner, P. (1999). Application of Coherent Risk Measures to Capital Requirements in Insurance. *North American Actuarial Journal* Vol 3 No 2.

Artzner, P., Delbaen, F., Eber, J.M. and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9, 3, 203-228.

Boskov, M. and Verrall, R.J. (1994). Premium Rating by Geographic Area using Spatial Models, *ASTIN Bulletin* Vol 24, No 1.

Brockman, M.J. and Wright, T.S. (1992). Statistical Motor rating: Making Effective Use of Your Data. *Journal of the Institute of Actuaries* 119, III, 457-543.

Cairns, A.J.G (2000). A discussion of parameter and model uncertainty in insurance. *Insurance: Mathematics and Economics* Vol 27 313-330.

Daykin, C.D., Pentikainen, T. and Pesonen, M. (1996). *Practical Risk Theory for Actuaries.* Chapman and Hall.

Embrechts, P., Kluppelberg, C., and Mikosch, T (1997). *Modelling Extremal Events.* Springer.

England, P.D. and Verrall, R.J. (2006). Predictive Distributions of Outstanding Liabilities in General Insurance. *Annals of Actuarial Science* Vol 1 Part II, 221-270.

Frees, E.W. and Valdez, E.A. (1998). Understanding relationships using copulas. *North American Actuarial Journal*, 2, 1, 1-25.

Geman, S. and Geman, D. (1984). Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **6**: 721–741.
Gilks, W.R. & Wild, P. (1992). Adaptive Rejection Sampling for Gibbs Sampling. *Applied Statistics*, **41 (2)**, 337-348.
Gilks, W.R., Best, N.G. & Tan, K.K.C. (1995). Adaptive Rejection Metropolis Sampling within Gibbs Sampling. *Applied Statistics*, **44 (4)**, 455-472.
Hartley, L.P. (1953). *The Go-Between*.
Klugman, S.A., Panjer, H.H. and Wilmot, G.E. (2004). *Loss Models: from data to decisions*. (2nd Edition). Wiley.
Macdonald, A.S. (2006). Actuarial Education: Does it Compute? *Annals of Actuarial Science* Vol 1 Part II.
Mack (1993). Distribution-free calculation of the standard error of chain-ladder reserve estimates. *ASTIN Bulletin*, **23**, 213-225.
McCullagh, P. & Nelder, J. (1983, 1989). *Generalized Linear Models*. Chapman and Hall, London.
Wang, S. (1995). Insurance pricing and increased limits ratemaking by proportional hazards transforms. *Insurance: Mathematics and Economics*, 17, 43-54.