1 Liabilities over long

1.1 Introduction

General insurance was in the preceding chapter handled on on a year-to-year basis. Premium was charged annually and the reserve re-calculated each year. A longer view is sometimes needed too. We might, for example, want to examine the amount of capital needed to take an insurance company through a certain number of years with minimal danger of a cash refill (the 'ruin' problem) or evaluate the long-term effect of re-insurance. Such issues were initiated in Section 3.5 through the recursion ($k = 1, 2 \ldots$)

$$\mathcal{V}_k = \mathcal{V}_{k-1} + \underbrace{\mathcal{R}_k\mathcal{V}_{k-1}}_{\textit{financial income}} + \underbrace{\Pi_k}_{\textit{premium}} - \underbrace{\mathcal{O}_k}_{\textit{overhead}} - \underbrace{\mathcal{X}_k}_{\textit{liabilities}},$$

where $\mathcal{V}_k$ is asset surplus over liabilities (or **solvency margin**) at time $t_k = kh$. The time increment h is often one year. Income (financial $\mathcal{R}_k\mathcal{V}_{k-1}$ and premium Π_k) are added and expenses (overhead costs $\mathcal{O}_k$ and insurance claims $\mathcal{X}_k$) subtracted. Financial return has a more natural home among life and pension insurance in Part III and is here only included as a fixed rate of interest $\mathcal{R}_k = r$. If $\mathcal{R}_k = 0$, the solvency margin reduces to the **underwriting result** where financial side is ignored.

Following $\{\mathcal{V}_k\}$ over many years could be a complex enterprise. Features not included in the recursion above are taxes and dividend to owners, several portfolios (not just one) and re-insurance (often present). Section 11.5 deals with the process of setting up simulation schemes that integrate many contributions of that kind. There is also additional complexity. Underneath the main variables we often find other time-varying factors, such as planned growth of the portfolio, inflation affecting losses, claim frequencies oscillating with the business cycle or risk in systematic growth or decline. Some of these issues raise new problems of **dynamic** modelling that weren't there in a short-term perspective. The most striking case may be claim numbers varying and fluctuating over time, a topic treated from two different angles in Sections 11.3 and 11.4.

With the features mentioned (and others!) combined in a simulation program we have a tool for solving the ruin problem with an ease and in a generality other computational methods simply can't match. *Traditional* ruin theory through the random walk is reviewed in Section 11.6.

1.2 Simple calculations

Introduction

Financial return as fixed rates of interest r allows the main features of the distributions of the solvency margins $\mathcal{V}_k$ to be examined through simple mathematics. The defining recursion now becomes

$$\mathcal{V}_k = (1+r)\mathcal{V}_{k-1} + \Pi_k - \mathcal{O}_k - \mathcal{X}_k, \qquad k = 1, 2 \ldots \qquad \text{for} \qquad \mathcal{V}_0 = v_0. \tag{1.1}$$

Here Π_k and $\mathcal{O}_k$ are taken as fixed quantities. Normally premium is paid in advance so Π_k covers the period ahead (the initial v_0 contains premium for the first period). For the aggregated claims $\mathcal{X}_k$ introduce the ordinary representation

$$\mathcal{X}_k = \sum_{i=1}^{\mathcal{N}_k} Z_{ki} \qquad \text{where} \qquad \mathcal{N}_k \sim \text{Poisson}(\lambda_k).$$

It is here assumed that the counts $\mathcal{N}_k$ and individual losses $Z_{1k}, Z_{2k}, \dots$ are stochastically independent for all k. All parameters are given as fixed values. That applies to the Poisson parameters λ_k and expectation, standard deviation, skewness (denoted ξ_{zk}, σ_{zk} and ζ_{zk}) of the individual losses. Some of of these conditions will be modified in later sections.

Net asset surplus at time t_k

The preceding assumptions mean that the stream of liabilities $\{\mathcal{X}_k\}$ ais stochastically independent. This is the prerequisite for simple results since it permits variances and third order moments to be broken down on sums of simple contributions. With expectations such decompositions are available regardless. Indeed,

$$E(\mathcal{V}_k) = (1+r)E(\mathcal{V}_{k-1}) + (\Pi_k - \mathcal{O}_k) - E(\mathcal{X}_k), \qquad E(\mathcal{X}_k) = \lambda_k \xi_{zk}, \tag{1.2}$$

true *generally* and

$$\text{var}(\mathcal{V}_k) = (1+r)^2 \text{var}(\mathcal{V}_{k-1}) + \text{var}(\mathcal{X}_k), \qquad \text{var}(\mathcal{X}_k) = \lambda_k(\xi_{zk}^2 + \sigma_{zk}^2) \tag{1.3}$$

$$\nu_3(\mathcal{V}_k) = (1+r)^3 \nu_3(\mathcal{V}_{k-1}) - \nu_3(\mathcal{X}_k), \qquad \nu_3(\mathcal{X}_k) = \lambda_k(\zeta_{zk}\sigma_{zk}^3 + 3\xi_{zk}\sigma_{zk}^2 + \xi_{zk}^3), \tag{1.4}$$

which is valid because of *independence.* The recursions run over for $k = 1, 2, \dots$, starting at

$$E(\mathcal{V}_0) = v_0, \qquad \text{var}(\mathcal{V}_0) = 0 \qquad \nu_3(\mathcal{V}_k) = 0. \tag{1.5}$$

All relationships are fairly obvious consequences of (1.1); see Section 11.6 for detailed proofs. Their implementation in the computer is straightforward. Skewness is computed through

$$\text{skew}(\mathcal{V}_k) = \frac{\nu_3(\mathcal{V}_k)}{\text{var}(\mathcal{V}_k)^{3/2}}. \tag{1.6}$$

This scheme provides us with an approximate view on the distribution of any $\mathcal{V}_k$ under any scheme where the time-variation is defined as fixed changes of the parameters.

When risk is constant

Closed expressions can be derived when risk parameters do not change with time. Let

$\Pi_k = J\pi, \ \ \mathcal{O}_k = Jo$	$\lambda_k = J\mu h$	$\xi_{zk} = \xi_z, \ \ \sigma_{zk} = \sigma_z, \ \ \zeta_{zk} = \zeta_z.$
premium and overhead	*claim frequency*	*losses per event*

where J is portfolio size (same for all k) and π and o are premium income and overhead per policy holder. Mean, standard deviation and skewness of $\mathcal{X}_k$ are now the same for all k with expressions given in Section 10.7. It is under these circumstances verified in Section 11.6 that

$$E(\mathcal{V}_k) = v_0(1+r)^k + J(\pi - o - \mu h \xi_z)\frac{(1+r)^k - 1}{r}, \tag{1.7}$$

$$\text{sd}(\mathcal{V}_k) = \text{sd}(\mathcal{X})\sqrt{\frac{(1+r)^{2k} - 1}{(1+r)^2 - 1}}, \tag{1.8}$$

$$\text{skew}(\mathcal{V}_k) = -\text{skew}(\mathcal{X})\frac{(1+r)^{3k} - 1}{\{(1+r)^{2k} - 1\}^{3/2}} \frac{\{(1+r)^2 - 1\}^{3/2}}{(1+r)^3 - 1}. \tag{1.9}$$

Here

$$\mathrm{sd}(\mathcal{X}) = \sqrt{J\mu h(\xi_z^2 + \sigma_z^2)} \qquad \text{and} \qquad \mathrm{skew}(\mathcal{X}) = \frac{\zeta_z\sigma_z^3 + 3\xi_z\sigma_z^2 + \xi_z^3)}{(J\mu h)^{1/2}\,(\xi_z^2 + \sigma_z^2)^{3/2}};$$

see (**??**) and (**??**). If you employ l'Hôpital's rule, it is easy to prove that

$$\mathrm{sd}(\mathcal{V}_k) \to \mathrm{sd}(\mathcal{X})\sqrt{k} \quad \text{and} \quad \mathrm{skew}(\mathcal{V}_k) \to -\frac{\mathrm{skew}(\mathcal{X})}{\sqrt{k}} \qquad \text{as} \qquad r \to 0.$$

The skewness coefficient suggests that $\mathcal{V}_k$ is becoming more and more Gaussian as the time horizon goes up. Skewness is unsensitive to variation in the rate of interest r; try a few values of r for a fixed value of skew($\mathcal{X}$); see Exercise 11.2.2.

Deductions and observations

A prerequisite for an insurance business making profit in the long run is is that the *second* term in (1.7) is positive; i.e. that

$$\underbrace{\pi - o - \mu h\xi_z > 0}_{\textit{premium in arrears}} \qquad \text{or rather} \qquad \underbrace{(1+r)\pi - o - \mu h\xi_z > 0.}_{\textit{premium in advance}}$$

In practice premium in property insurance is always paid in advance, and we must include the interest-earning contribution on the right. Longer delays in the settlements of claims (as in Chapter 12) makes the effect of financial income larger; see also Section 11.4.

In the very long run the power terms $(1+r)^k$ and $(1+r)^{2k}$ in (1.7) and (1.8) dominate the rest. Hence both mean and standard deviation grows *approximately* as

$$E(\mathcal{V}_k) \doteq a_1(1+r)^k \qquad \text{and} \qquad \mathrm{sd}(\mathcal{V}_k) \doteq a_2(1+r)^k$$

where

$$a_1 = \{v_0 + J(\pi - o - \mu h\xi_z)\}/r \qquad \text{and} \qquad a_2 = \sqrt{\frac{J\mu h(\xi_z^2 + \sigma_z^2)}{(1+r)^2 - 1}}.$$

The growth is the same as for an ordinary bank account. Don't use these relationships for numerical approximations! Skewness is little influenced by r, as was noted above.

It isn't common to use this approach to determine the amount of initial capital (you see why in Exercise 11.2.3). But Normal and Normal Power approximations to the percentiles $q_{\epsilon k}$ of $\mathcal{V}_k$ *are* possible, and the accuracy is somewhat improved over those in Section 9.2 because the number of independent terms behind are higher. The form of the approximations are

$$\underbrace{q_{\epsilon k}^{\mathrm{No}} = E(\mathcal{V}_k) + \phi_\epsilon \mathrm{sd}(\mathcal{V}_k)}_{\textit{normal approximation}} \qquad \text{and} \qquad \underbrace{q_{\epsilon k}^{\mathrm{NP}} = q_{\epsilon k}^{\mathrm{No}} + \mathrm{sd}(\mathcal{V}_k)\mathrm{skew}(\mathcal{V}_k)\frac{1}{6}(\phi_\epsilon^2 - 1);}_{\textit{normal power approximation}} \qquad (1.10)$$

exactly the same as in Section 10.2.

1.3 Modelling dynamic risk: The hidden process

Introduction

Liability risk $\mathcal{X}_k$ at time t_k depends on a Poisson parameter λ_k and a distribution describing individual losses Z_k. Those are in turn influenced by other parameters and input quantities that may well vary with time. Regression is used as a tool in the next section; here the viewpoint is **latent** or **hidden** risk. The idea was developed in Section 8.5 as a random intensity μ buried in the number of claims $\mathcal{N}$ we actually see. Over longer time horizons experience suggests that claim intensity μ_k at time t_k might be linked to the state of the economy, more accidents in boom times! If so, the entire sequence $\mu_1, \mu_2, \dots$ will be auto-correlated, and a stochastic process is needed to describe it.

This section demonstrates how it can be done. The issue is a challenging one. Our target is a process $\{\mu_k\}$ hidden in another process $\{\mathcal{N}_k\}$ (the counts of the actual claims). Such hierarchically positioned time series are common in many fields of science; see references at the end of the chapter. A major problem is how the model for the inner series is determined from historical data. Unlike (say) rates of interests frequencies of claims are not observed directly! The treatment given is here is a rudimentary one that does not provide justice to the subject, but at least it is simple enough for it to be implemented on your own. Other cases of hidden risk is discussed among the exercises; see also Chapter 13.

Claim intensity as log-normal, stationary process

Let μ_0 be the claim intensity today and $\mu_1, \mu_2, \dots$ those of the future. They are all random with mean $\xi = E(\mu_k)$ and standard deviation $\sigma_\mu = \mathrm{sd}(\mu_k)$. That's how far Section 8.5 took us, and we now seek a way to impose fluctuations similar to business cycles. Such models are *stationary.* If that concept is unfamiliar, consult Section 5.7 before you go further (and look up first-order auto-regression too). About the simplest piece of modelling imaginable is to take ($k = 1, 2, \dots$)

$$\mu_k = \xi \exp(-\frac{\tau^2}{2} + \tau Y_k), \qquad \text{where} \qquad Y_k = aY_{k-1} + \sqrt{1 - a^2}\,\varepsilon_k. \tag{1.11}$$

Here τ and a are parameters ($\tau \geq 0$ and $|a| < 1$), and the sequence $\varepsilon_1, \varepsilon_2, \dots$ independent and $N(0,1)$. We are dealing with a recursion starting at

$$Y_0 = \frac{\tau}{2} + \frac{\log(\mu_0/\xi)}{\tau} \qquad (\mu_0 \text{ the initial intensity}), \tag{1.12}$$

which is deduced from (1.11) left by solving for Y_0.

The construction is similar to the Black-Karisinsky model for interest rates in Section 5.7. If the model is simulated, the paths of μ_k will gradually spread out from the common start until their fluctuations have stabilized; see Figure 11.1. At that point Y_k is $N(0,1)$, and the standard formulae for mean and standard deviations of log-normals give us

$$\xi = E(\mu_k) \quad \text{and} \quad \sigma_\mu = \mathrm{sd}(\mu_k) = \xi\sqrt{\exp(\tau^2) - 1} \qquad \text{(initial state forgotten);}$$

look them up in the introduction to Section 9.3 in case you have forgotten. There are three parameters, ξ, τ and a.

A more general viewpoint

Dymamic models for claim intensities can also be set up by combining auto-regression and inversion sampling. This is in reality the copula approach of Section 6.7 which allows us to *choose* an arbitrary family of distributions for μ_k. Let $F_0(s)$ be the distribution function of a positive random variable S *with mean one* and define $(k = 1, 2\ldots)$

$$\mu_k = \xi F_0^{-1}(U_k), \qquad U_k = \Phi(Y_k), \qquad Y_k = aY_{k-1} + \sqrt{1-a^2}\,\varepsilon_k. \tag{1.13}$$

Here $F_0^{-1}(u)$ is the percentile function of $F_0(s)$ and $\Phi(y)$ the standard normal integral. The driver process $\{Y_k\}$ is the same as before. After a while it has become $N(0,1)$ making U_k uniform. At that point μ_k has the familiar representation $\mu_k = \xi S_k$ where $S_k = F_0^{-1}(U_k)$, by inversion F_0-distributed. The scheme is run from

$$Y_0 = \Phi^{-1}\{F_0(\mu_0/\xi)\} \qquad\qquad (\mu_0 \text{ initial intensity}), \tag{1.14}$$

which follows from (1.13).

A favourite selection for $F_0(s)$ would be the standard Gamma distribution $G_\alpha(s)$ where α defines the shape. Then

$$\xi = E(\mu_k) \quad \text{and} \quad \sigma_\mu = \text{sd}(\mu_k) = \xi/\sqrt{\alpha} \qquad\qquad \text{(initial state forgotten)},$$

and now the three parameters are ξ, α and a. If $F_0(s)$ is specified as the log-normal, we are back to the Black-Karisinsky version above; see Exercise 11.3.5.

Claim numbers as stochastic process

The link to the actual claims is through

$$\mathcal{N}_k|\mu_k \sim \text{Poisson}(\lambda_k) \qquad \text{where} \qquad \lambda_k = J_k\mu_k. \tag{1.15}$$

Here the number of policies $\{J_k\}$ is known (they might well be planned growth). It is also natural to assume that

$$\mathcal{N}_1, \mathcal{N}_2, \ldots \qquad \text{conditionally independent given} \qquad \mu_1, \mu_2, \ldots; \tag{1.16}$$

i.e. stochastic dependence in the series $\{\mathcal{N}_k\}$ is created by $\{\mu_k\}$ only. This condition is implicit in Algorithm 11.1 below.

What realized series look like is indicated in Figure 11.1 where Monte Carlo realizations (30 replications) are plotted jointly. They were generated from Algorithm 11.1 under the model scenario of the figure caption. The initially intensity $\mu_0^* = 4.2\%$ was common to all. Under the model assumed that is on the low side, and the reversion to mean effect in the inner series $\{\mu_k^*\}$ (see Section 5.7) makes the simulated claim numbers $\{\mathcal{N}_k^*\}$ grow from an expectation of 420 in the beginning to fluctuations around their long-term average at 560 later. Note that the transition is *slower* for the *higher* value of the coefficient a on the right. If you think the variation in the simulated series is higher on the left, that is an illusion. The theoretical one is exactly the same under both models.

Utilizing historical data

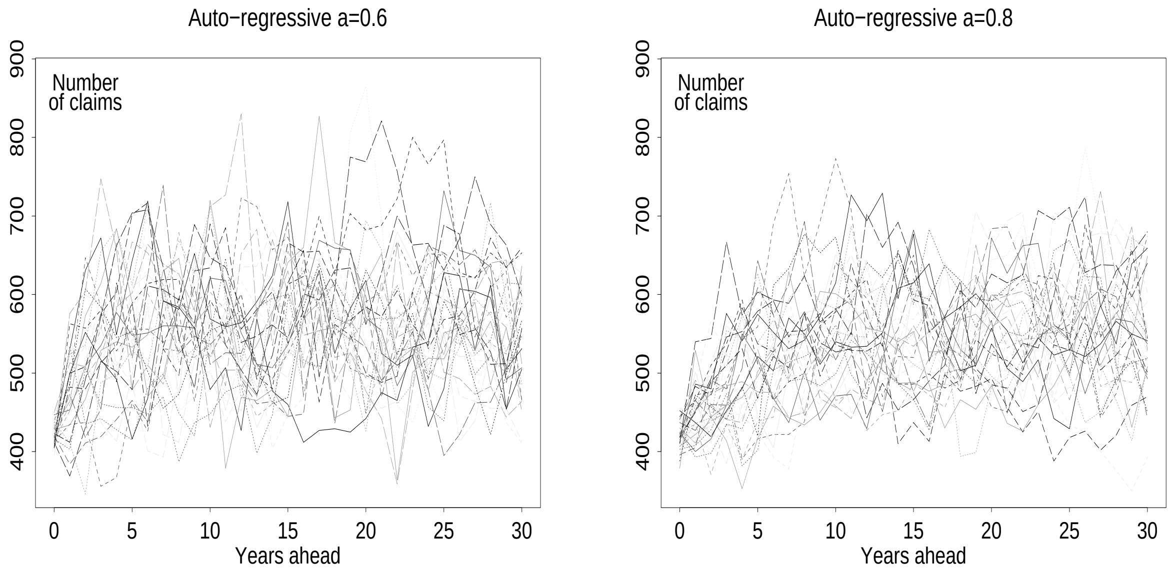

Figure 11.1 Simulated claim frequencies for 10000 *policies under log-normal hidden risk. Parameters:* $\xi = 5.6\%$, $\sigma_\mu = 0.7\%$ *and* $a = 0.6$ *(***left***) and* $a = 0.8$ *(***right***). Common, initial* $\mu_0^* = 4.2\%$.

We must address how the parameters are determined from the historical counts $\{n_k\}$. Consider the series of elementary estimates $\{\hat{\mu}_k\}$ where $\hat{\mu}_k = n_k/J_k$. This can be taken as our data, but the parameters we seek apply to a different series $\{\mu_k\}$. Can that be ignored? Not for small portfolios where the uncertainty make the two series quite deviating, but for large ones: Yes. The best way (catering for all possibilities) is to use likelihood estimation, but you won't find that in standard software, and it is too complicated numerically to be included here (references in Section 11.8). A simple scheme that works for sufficiently large portfolios runs as follows.

Start by estimating ξ and σ_μ by (**??**) and (**??**) in Section 8.5; i.e. take

$$\hat{\xi} = w_1\hat{\mu}_1 + \ldots + w_n\hat{\mu}_n \qquad \text{where} \qquad w_k = \frac{J_k}{J_1 + \ldots + J_n}, \tag{1.17}$$

$$\hat{\sigma}_\mu^2 = \frac{\sum_{k=1}^n w_k(\hat{\mu}_k - \hat{\xi})^2 - C}{1 - \sum_{k=1}^n w_k^2} \qquad \text{where} \qquad C = \frac{(n-1)\hat{\xi}}{J_1 + \ldots + J_n}. \tag{1.18}$$

Note that the historical data extend n periods back. The justification in Section 8.7 is still valid. Next convert $\hat{\sigma}_\mu$ to estimates of the shape of the distributions through

$$\underbrace{\hat{\tau} = \sqrt{\log(1 + (\hat{\sigma}_\mu/\hat{\xi})^2)}}_{\textit{log-normal}} \qquad \text{and} \qquad \underbrace{\hat{\alpha} = (\hat{\xi}/\hat{\sigma}_\mu)^2,}_{\textit{Gamma}} \tag{1.19}$$

which follows from the formulae for σ_μ. Finally reconstruct $(k = 1, \ldots, n)$ the process $\{Y_k\}$ by

$$\underbrace{\hat{Y}_k = \frac{\hat{\tau}}{2} + \frac{\log(\hat{\mu}_k/\hat{\xi})}{\hat{\tau}}}_{\textit{log-normal}} \qquad \text{and} \qquad \underbrace{\hat{Y}_k = \Phi^{-1}\{G_{\hat{\alpha}}(\hat{\mu}_k/\hat{\xi})\}}_{\textit{Gamma}} \tag{1.20}$$

and take

$$\hat{a} = \frac{\sum_{k=1}^{n} \hat{Y}_k \hat{Y}_{k-1}}{\sum_{k=1}^{n} \hat{Y}_k^2}. \tag{1.21}$$

Justification: $\hat{Y}$ is the solution of (1.11) and (1.13) (estimates inserted for parameters), and $\hat{a}$ is the routine estimate of auto-regressive coefficients.

Numerical experiment

The preceding argument is so heuristic that we simple have to check that it works. A representative model scenario (annual parameters) is

$$\xi = 5.6\% \quad \sigma_\mu = 0.7\% \quad a = 0.6 \qquad \text{and} \qquad n = 20, \quad J_1 = \ldots = J_{20} = 10000.$$

The log-normal model was run $m = 1000$ times and the parameters reconstructed from the simulations. Average and standard deviations of the estimates were

Mean	sd	Mean	sd	Mean	sd
5.6%	0.3%	0.63%	0.15%	0.36	0.21,
for ξ		*for* σ_μ		*for* a	

which shows that a is the most difficult parameter to estimate (large standard deviation, large bias with the average estimate (0.36) much smaller than the true one (0.60). This state of affairs owes much to the shortness of the time series (20 years), and didn't change much when the number of policies where ten-doubled. Results under the Gamma model were similar.

Monte Carlo implementation

It is (as usual) good training to summarize how the model is simulated. The following algorithm has been detailed for the log-normal with the Gamma version as commentaries:

Algorithm 11.1 Claims from log-normal intensities %*Gamma too*

0 Input: ξ, τ, a, $\sigma = \sqrt{1-a^2}$ and $\{J_k\}$ %*Gamma:* α *instead of* τ

1 $\mu^* \leftarrow \mu_0$ %*Initial* μ_0*, comments below*

2 $Y^* \leftarrow \tau/2 + \log(\mu^*/\xi)/\tau$ %*Gamma:* $Y^* \leftarrow \Phi^{-1}\{F_0(\mu^*/\xi)\}$

3 For $k = 1, 2, \ldots$ do

4 Draw $\varepsilon^* \sim N(0,1)$ and $Y^* \leftarrow aY^* + \sigma\varepsilon^*$

5 $\mu^* \leftarrow \xi \exp(-\tau^2/2 + \tau Y^*)$, and $\lambda^* \leftarrow J_k \mu^*$ %*Gamma:* $\mu^* \leftarrow \xi G_\alpha^{-1}\{(\Phi(Y^*))\}$,

6 Draw $\mathcal{N}_k^* \sim \text{Poisson}(\lambda^*)$

7 Return $\mathcal{N}_1^*, \mathcal{N}_2^*, \ldots$

How should the algorithm be started? One strategy is to depart from the stationary distribution. Line 1 is then skipped and on Line 2 you take $Y^* \sim N(0,1)$. The study of estimation errron above was run in this manner. Alternatively, the first μ^* may be the current value μ_0. If only available with error (which you may want to take into account) implement the conclusion in Exercise 11.3.6.

1.4 Dynamic risk II: Using regression

Introduction

A simple extension of the model of the preceding section is

$$\mu_k = \xi_k S_k \qquad \text{where} \qquad E(S_k) = 1. \tag{1.22}$$

Here $\{S_k\}$ is a stochastic part as before, but the time-dependent means $\{\xi_k\}$ is a new feature. Their role is to portray fluctuations in risk that are systematic and foreseeable. A case in point is the Norwegian claim frequency data of Figure 8.1 where there was a a steady increase in accident risk over a $3-4$ year period and monthly oscillations linked to the time of the year. These are fixed effects that might be learned from historical data through regression. The purpose of this section is to outline how that is done. We shall draw on Poisson regression from Section 8.4 and even extend so that it applies to random background conditions. The Norwegian automobile data will serve as illustration. These are monthly claim numbers from around 200000 policies over a period of 45 months extending from January 1997 up to September 2001.

Poisson regression with time effects

Historical claim data consists of individuals attached explanatory variables of the type presented earlier. With time added, we may envisage an arbitrary row of a data matrix as

$$\underset{\textit{individual}}{j} \quad \underset{\textit{time}}{k} \quad \underset{\textit{claim number}}{n_{jk}} \quad \underset{\textit{ordinary covariates}}{x_{j1},\dots,x_{jv_x},}$$

extending over all j and k. Suppose first that all background uncertainty is removed by taking $S_k = 1$ in (1.22). This yields a pure Poisson model; i.e.

$$n_{jk} \sim \text{Poisson}(\lambda_{jk}) \qquad \text{where} \qquad \lambda_{jk} = T_{jk}\xi_{jk}. \tag{1.23}$$

Here T_{jk} is the exposure of individual j in period k (often $T_{jk} = 1$). The detailed modelling is in terms of ξ_{jk}. A very common approach in such situations is a *multiplicative* specification, i.e.

$$\xi_{jk} \quad = \quad \underset{\textit{policy}}{\xi_{j\cdot}} \cdot \underset{\textit{time}}{\xi_{\cdot k}} \qquad \text{where} \qquad \log(\xi_{j\cdot}) = \sum_{i=1}^{v_x} b_i x_{ji}.$$

Here $\xi_{j\cdot}$ represents the individual and is described by an ordinary log-linear relationship. The new feature is the time factor $\xi_{\cdot k}$ for which a *second* log-linear relationship is a natural possibility. Suppose there are v_s different seasonal components (with $s = 12$ if they are months and $s = 4$ if quarters and so forth). Introduce

$$\begin{aligned} s_k(i) &= 1, \quad \text{for season } i \text{ at time } t_k \\ &= 0 \quad\ \text{otherwise,} \end{aligned} \qquad \text{and} \qquad \log(\xi_{\cdot k}) = b_d t_k + \sum_{i=2}^{v_s} s_k(i) b_s(i).$$

Here b_d accounts for a systematic drift upwards (or downwards), and $b_s(2),\dots,b_s(v_s)$ define the seasonal contributions. The coding is the same as in Section 8.4 with $b_s(1) = 0$. Note that $\log(\xi_{\cdot k}) = b_d t_k + b_s(i)$ for season i. Combining all these relationships leads to the regression equation

$$\log(\xi_{jk}) \quad = \quad \underset{\textit{covariates}}{\textstyle\sum_{j=1}^{v_x} b_k x_j} \quad + \quad \underset{\textit{drift}}{b_d t_k} \quad + \quad \underset{\textit{seasonal}}{\textstyle\sum_{i=2}^{v_s} s_k(i) b_s(i)}, \tag{1.24}$$

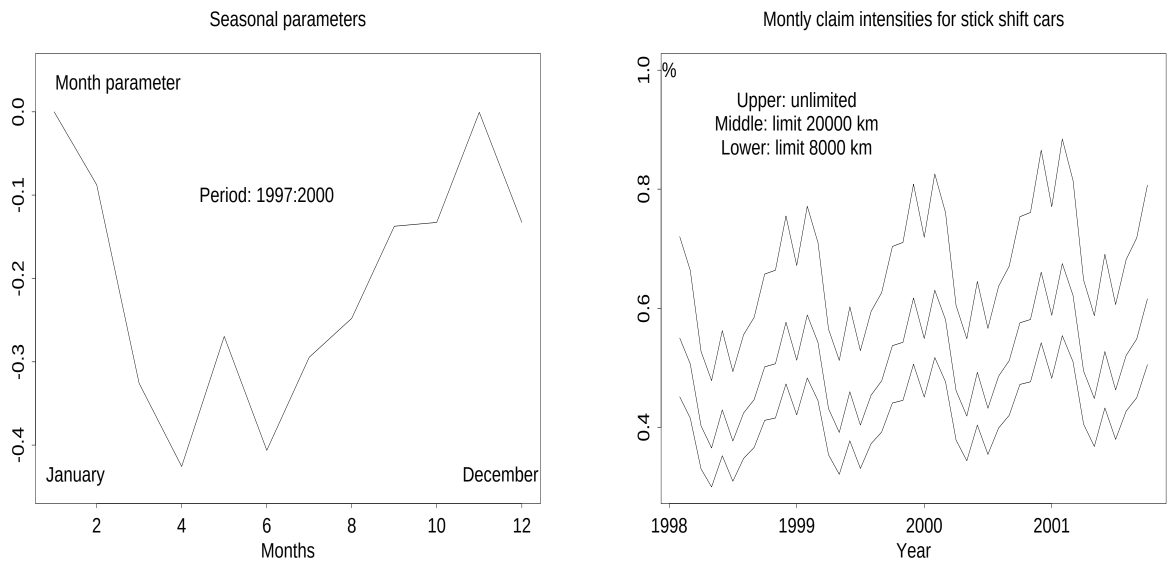

Figure 11.2 Left: Regression coefficients for seasonal effects. Right: Estimated claim intensities for three groups of policies

which can be fitted the historical data $\{n_{jk}\}$ by Poisson regression as outlined in Section 8.4.

Example: Automobile insurance
To see how the model works consider claim numbers from the Norwegian car insurance portfolio where the underlying intensity both oscillates over the year and has grown systematically over the period in question. There are covariates too. Those selected (purely for the purpose of illustration) were

- *car type* (two categories, transmission either manual or automatic),
- *driving limit on policy* (seven categories, definitions in Table 11.1).

The dynamic part consists of a drift term and twelve seasonal contributions (one for each month). This yields a regression equation of the form (1.24) with estimated coefficients listed in Exercise 11.4.1.

The monthly components $\hat{b}_s(i)$ have been plotted in Figure 11.2 left. For January $\hat{b}_s(1) = 0$ by design. The seasonal pattern is a strong one (as expected for a country of Northern Europe) with negative values for the summer months (accidents less frequent)[1]. The fitted model becomes

$$\hat{\xi}_{jk} = \underbrace{\exp\left(\sum_{j=1}^{v_x} \hat{b}_k x_j\right)}_{\textit{client contribution}} \cdot \underbrace{\exp\left(\hat{b}_d t_k + \sum_{i=2}^{v_s} s_k(i)\hat{b}_s(i)\right)}_{\textit{dynamic contribution}},$$

which has been plotted agains t_k in Figure 11.2 right for three different categories of customers. What emerges is an increasing pattern modulated by seasonal effects and pushed up and down in

[1]The local top for May is real and inherently Norwegian due to people (for local reasons) driving a lot during that month.

parallel according to characteristics of the policy holder.

Random effect regression

Background conditions may easily be random in a case like the preceding one and is accomodated through the conditional model

$$n_{jk}|\lambda_{jk} \sim \text{Poisson}(\lambda_{jk}) \qquad \text{where} \qquad \lambda_{jk} = T_{jk}\xi_{jk}S_k. \tag{1.25}$$

Here S_k is a random term influencing (in period k) all policies equally. You may not easily find software for such regression models, yet a reasonably simple method *is* developed in Section 11.7 when $S_k \sim \text{Gamma}(\alpha_k)$ for $k = 1, 2, \ldots$ and all S_k independent. This is **negative binomial regression** with $\alpha_1, \alpha_2 \ldots$ being responsible for background randomness. We are back to ordinary Poission regression if all $\alpha_k \to \infty$ (then all $S_k = 1$).

The Gamma parameters must in practice be estimated from historical data (that's part of the method in Section 11.7), and usually we do not want many of them. Possible simplifications are

$$\underbrace{\alpha_k = \alpha}_{\textit{constant variability}} \qquad \text{and} \qquad \underbrace{\alpha_k = \alpha(i) \quad \text{for season } i \text{ at time } t_k.}_{\textit{variability depending on season}}$$

There are on the right v_s parameters $\alpha(1), \ldots, \alpha(v_s)$ for v_s seasons. Both formulations are captured by the method in Section 11.7. It is based on a slightly complicated likelihood function (that's why details have been deferred), but we do not need its precise form to to understand what we might get out of it.

Automobile insurance: A new round

The monthly description in Figure 11.1 may be more detailed than necessary. We shall now attempt a rougher decription with only two seasonal components. The year is divided into 'winter' and 'summer' with random background conditions (due to weather) present. Such variability may easily be stronger in winter. That is taken care of by separate Gamma parameters $\alpha(1)$ and $\alpha(2)$ for the two seasons. With two covariates present the model now reads

$$\mu_{jk} = \xi_{jk}S_k, \qquad \text{where} \qquad \log(\xi_{jk}) = b_0 + b_i x_{1j} + b_2 x_{2j} + b_d t_k + s_k(2)b_s(2)$$

and

$$S_k \sim \text{Gamma}(\alpha_k) \qquad \text{for} \qquad \alpha_k = \{1 - s_k(2)\}\alpha(1) + s_k(2)\alpha(2).$$

Here the zero-one indicator $s_k(2)$ is responsible for the right seasonal parameters at time t_k. The model is fed the likelihood function as outlined in Section 11.7.

In detail, fitting was still done on a monthly time scale with the same two covariates (car type, driving limit) as earlier. Seasons were reduced to 'winter' (October to April) and 'summer' (May to September), and there is a trend term. Estimated parameters have been listed in Table 11.1 and compared with those obtained when background uncertainty is ignored. The two sets are not that different, although there are certain discrepancies. Above all, monthly drift has gone up around 25% (from 0.57% to 0.71%) which would have repercussions on long time projections. As expected background uncertainty is a more important feature in winter (lower α); see Figure 11.3 below where simulated liability scenarios have been plotted.

[a]Model with fixed parameters for season. [b]Random parameters for season.

			Car types (according to gear)		
	Fixed[a]	Random[b]		Fixed[a]	Random[b]
Intercept	-5.407	-5.407	Manual	0	0
Monthly drift	0.0057	0.0071	Automatic	-0.342	0.385
Seasonal mean			Distance limits (in 1000 km)		
	Fixed[a]	Random[b]		Fixed[a]	Random[b]
Summer	0	0	8	0	0
Winter	0.166	0.149	12	0.097	0.111
			16	0.116	0.124
Seasonal α			20	0.198	0.198
Summer	-	20.6	25	0.227	0.224
Winter	-	4.0	30	0.308	0.322
			no limit	0.468	0.483

Table 11.1 Estimated (monthly) coefficients when seasonal effects are described by **fixed** *and* **random** *regimes.*

1.5 Building simulation models

Introduction

We shall in this section examine net assets $\{\mathcal{V}_k\}$ as defined in (1.1). Though a simple version contributions could still be many, for example different cash flows hiding in $\{\mathcal{X}_k\}$ (possibly influenced by inflation too), time-varying (even stochastic) parameters, time-varying volume of business and market fluctuations in premia. We have to learn how these things are integrated into statements of risk. That means Monte Carlo. Simulating $\{\mathcal{V}_k\}$ is sometimes called Dynamic, Financial Analysis (**DFA**). It could be in practice be a huge enterprise in modelling and implementation. Rather than demonstrating this on a grand scale this section tries to show how such programs are being constructed. You should above all proceed step-by-step or **hierarchically**. For example, suppose liabilities $\{\mathcal{X}_k\}$ consist of several sub-classes $\{\mathcal{X}_{jk}\}$. If there are J of them, then

$$\underbrace{\mathcal{X}_k = \mathcal{X}_{1k} + \ldots + \mathcal{X}_{Jk}}_{\textit{real loss}} \qquad \text{and} \qquad \underbrace{\mathcal{X}_k^* = \mathcal{X}_{1k}^* + \ldots + \mathcal{X}_{Jk}^*}_{\textit{simulated loss}},$$

and all we have to do is to add output from J related implementations that would differ in detail. Subsequently this sum is fed the rest of the program. There is more on this line of thinking in Chapter 15.

A second issue is the re-use of existing computer programs, possibly modified to deal with additional aspects. Consider, for example, a cedent buying re-insurance. The solvency margin recursion (1.1) now changes to

$$\mathcal{V}_k = (1+r)\mathcal{V}_{k-1} \quad + \quad \underbrace{(\Pi_k - \Pi_k^{\mathrm{re}})}_{\textit{net premium received}} \quad - \quad \underbrace{(\mathcal{O}_k + O_k^{\mathrm{re}})}_{\textit{total overhead}} \quad - \quad \underbrace{(\mathcal{X}_k - \mathcal{X}_k^{\mathrm{re}})}_{\textit{net loss}},$$

for $k = 1, 2 \ldots$. Here Π_k^{re}, O_k^{re} and $\mathcal{X}_k^{\mathrm{re}}$ are premium, overhead and reimbursement due to re-insurance, not difficult to work into a given implementation (often O_k^{re} would be too small to matter). Several examples of such extra terms are discussed below. The list is not exhaustive, yet comprises a lot of what you will encounter in practice.

Integrating hidden risk
Time-varying parameters are simply entered into schemes like Algorithms 3.5 and 3.6. If fixed and given they are simply read from file, but stochastic versions must be simulated. How such commands are organized is indicated in Algorithm 11.2 which combines Algorithm 11.1 for hidden, log-normal claim intensity risk with the underwritng Algorithm 3.5:

Algorithm 11.2 Underwriting result under log-normal claim intensitites

0 Input: For claim intensity model ξ, τ, a, $\sigma = \sqrt{1-a^2}$, $\{J_k\}$, π, o.

1 $\mathcal{V}_0^* \leftarrow v_0$, draw μ_0^*, $Y^* \leftarrow \tau/2 + \log(\mu_0^*/\xi)$ *%About μ_0^*: See comments*
2 For $k = 1, \ldots,$ do *%following Algorithm 11.1*

3 Draw $\varepsilon^* \sim N(0,1)$ and $Y^* \leftarrow aY^* + \sigma\varepsilon^*$
4 $\mu^* \leftarrow \xi \exp(-\tau^2/2 + \tau Y^*)$ and $\lambda^* \leftarrow J_k \mu^*$ *%Current Poisson parameter*
5 Generate $\mathcal{X}_k^*$ *%λ^* as input to Algorithm 3.1.*
6 $\mathcal{V}_k^* \leftarrow \mathcal{V}_{k-1}^* + J_k(\pi - o) - \mathcal{X}_k^*$

7 Return $\mathcal{V}_1^*, \mathcal{V}_2^*, \ldots$

The crucial point is that the liabilitities (Line 5) are generated *conditionally* given the current Poisson parameter. Several of the exercises discuss similar constructions. It is useful to separate volume (i.e the size of the portfolio) from market and expenses. A convenient way is through

$$\Pi_k = J_k \pi \qquad \text{and} \qquad O_k = J_k o$$

where π and o is premium and overhead cost per policy.

The relevance of hidden risk
As an example of how the preceding algorithm can be put to use consider a situation with claims influenced by random background conditions. Now, this *has* important repercussions on risk (as has been argued repeatedly), but it isn't equally clear what the consequences would be of neglecting auto-regressive dynamics. A small experiment was run with claim intensity being first-order auto-regressive and log-normal. Three different model scenarios were compared; i.e.

$\sigma_\mu = 0$	$\sigma_\mu = 0.7\%$ and $a = 0$	$\sigma_\mu = 0.7\%$ and $a = 0.6$,
fixed μ	*stochastic μ, independence*	*stochastic μ, auto-correlation*

with expectation $\xi = 5.6\%$ as before. Other conditions were

$Z = \exp(-\tau^2/2 + \tau\varepsilon)$, $\tau = 0.5$	$J_k = 100000$	$\pi - o = 0.0612$
claim model per event	*portfolio size*	*net premium per policy*

Note that the net premium exceeds the pure one (0.056) by 10%. The underwriting result $\{\mathcal{V}_k\}$ was followed five years through $m = 10000$ simulated scenarios. The mean of $\mathcal{V}_5$ was 2820 (same for all scenarios), but standard deviations were far from equal:

$\mathrm{sd}(\mathcal{V}_5) = 180$	$\mathrm{sd}(\mathcal{V}_5) = 1600$	$\mathrm{sd}(\mathcal{V}_5) = 2360$
fixed μ	*stochastic μ, independence*	*stochastic μ, auto-correlation*

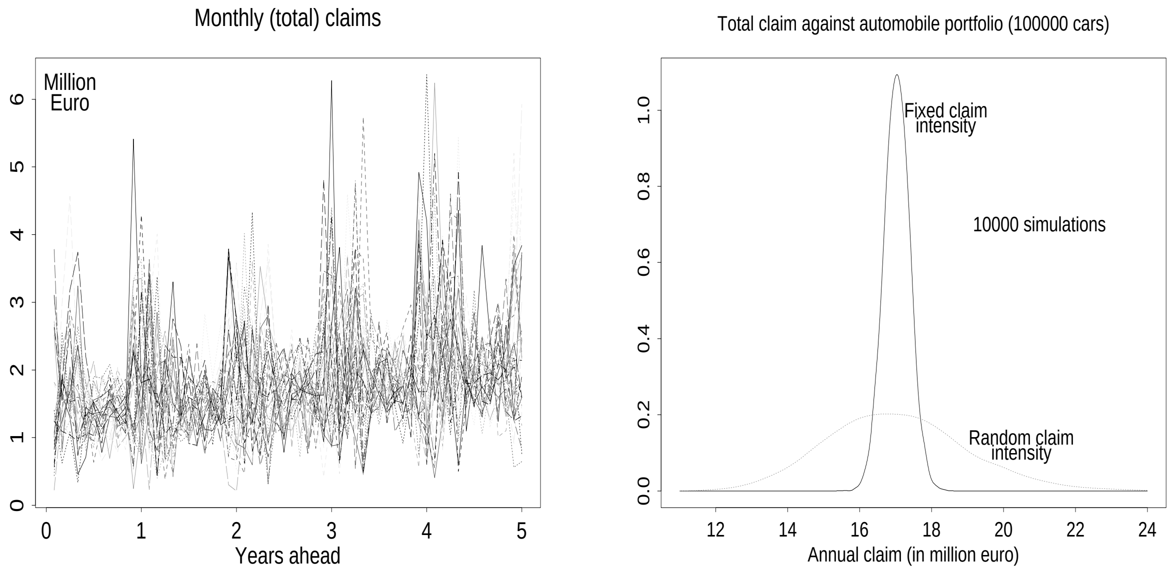

Figure 11.3 **Left**: *Monthly portfolio loss (30 replications) for the random regime seasonal model of Table 11.1.* **Right**: *One-year total loss compared with the fixed regime model*

A positive result from underwriting is virtually certain when μ is fixed, but not under the other model scenarios. Auto-correlation *is* significant and should not be ignored.

Trend and seasonal fluctuations
As a second example consider the trend/seasonal model for claim intensity in Table 11.1. Time scale is now monthly and for an annual simulation Algorithm 11.2 must be run twelve steps. There was no auto-regressive element (take care of that by inserting $a = 0$) whereas expectations depended on k, and on Line 4 ξ must be replaced by ξ_k. Gamma distributions were used rather than log-normal ones. For the required modifications, see the comments to Algorithm 11.1.

This model has in Figure 11.3 been coupled with a claim size model based on the empirical distribution of 6764 losses on Norwegian car accidents. Their mean and standard deviation were 3.018 and 3.636 thousand Euro respectively. Monthly payments over a period of five years have been plotted on the left for a portfolio of $J_k = 100000$ policies. High background uncertainty makes the results highly variable, much more so in the winter months. There was an increasing trend in the accident rate. This is strongly felt and has huge repercussions if prolonged over the entire five year period. The uncertainty in annual results for the first year is shown on the right in Figure 11.3. As we have seen many time earlier, the variability is much smaller when randomness in the background conditions is ignored.

Dealing with inflation
Future claims are likely to be influenced by price increases which over longer time horizons could have huge consequences. Inflation is treated more throroughly in Part III; here is a simple way to add it. Suppose premia, expenses and losses are all subjected to the same price changes. The

recursion (1.1) for the net asset surplus then becomes ($k = 1, 2, \ldots$)

$$\mathcal{V}_k = (1+r)\mathcal{V}_{k-1} + Q_k\{\Pi_k - \mathcal{O}_k - \mathcal{X}_k\}, \qquad \text{where} \qquad Q_k = (1+I_k)Q_{k-1},$$

starting at $Q_0 = 1$ (and $\mathcal{V}_0 = v_0$). Here Q_k is the price level and I_k the rate of inflation in period k.

Taxes and dividend

When premia fluctuate

In real life annual premium π per policy is no constant quantity. Indeed, strong oscillations have been far from uncommon (see Daykin, Penticäinen and Pesonen, 1994). One way to attack the issue is to distinguish between the pure premium π^{pu} that exactly covers the loss on average and the amount π actually paid. Their relationship is

$$\pi_k = (1+\gamma_k)\pi^{\text{pu}},$$

where the sequence of loadings $\{\gamma_k\}$ capture **market impact**. Note that the pure premium per policy is constant, though our assessment of it may be uncertain.

There are several ways to build models that try to imitate how market forces affect pricing; for references see Section 11.8. Perhaps the most natural one is through the loadings. A simple approach is to let the state of a national or global insurance market be reflected by the solvency margin of the company in question. This links γ_k to a preceding value $\mathcal{V}_{k-d}$ where d is a time lag a company needs to accomodate desired changes in pricing. Daykin, Penticäinen and Pesonen (1994) argues that $d = 2$ years (at least), and following what is essentially their line, we might take

$$\gamma_k = \xi_\gamma e^{-b(v_{k-2}-v_0)} \qquad \text{where} \qquad v_k = \frac{\mathcal{V}_k}{\Pi_k}. \tag{1.26}$$

What counts under this model is the solvency ratio v_k and its deviation from a target value v_0. Discrepancies takes later loadings up or down from the fixed ξ_γ. The third parameter $b \geq 0$ reflects sensitivity to the market condition. Implementation is discussed in Exercise 11.5.3

Simulations of such a system under the auto-regressive and log-normal claim intensity with log-normal losses (see above) are shown in Figure 11.4. Detailed *additional* assumptions are

$$\underbrace{\xi_\gamma = 0.25,\ b = 1,\ v_0 = 0.6}_{\textit{premium model}} \qquad \underbrace{o = 0.2\pi^{\text{pu}},}_{\textit{expenses}} \qquad \underbrace{\mathcal{V}_0 = 20J_k,}_{\textit{initial capital}}$$

which leads to the fluctuations in the premium received (left) and the solvency ratio (right). The solvency ratio is low in the beginning, but increases to (strong!) oscillations around the target value of 60%. Note the reversion to mean effect, caused by the underlying claim intensity model. Real life situations have this flavour, though the fluctuations here are (perhaps) a bit exaggerated.

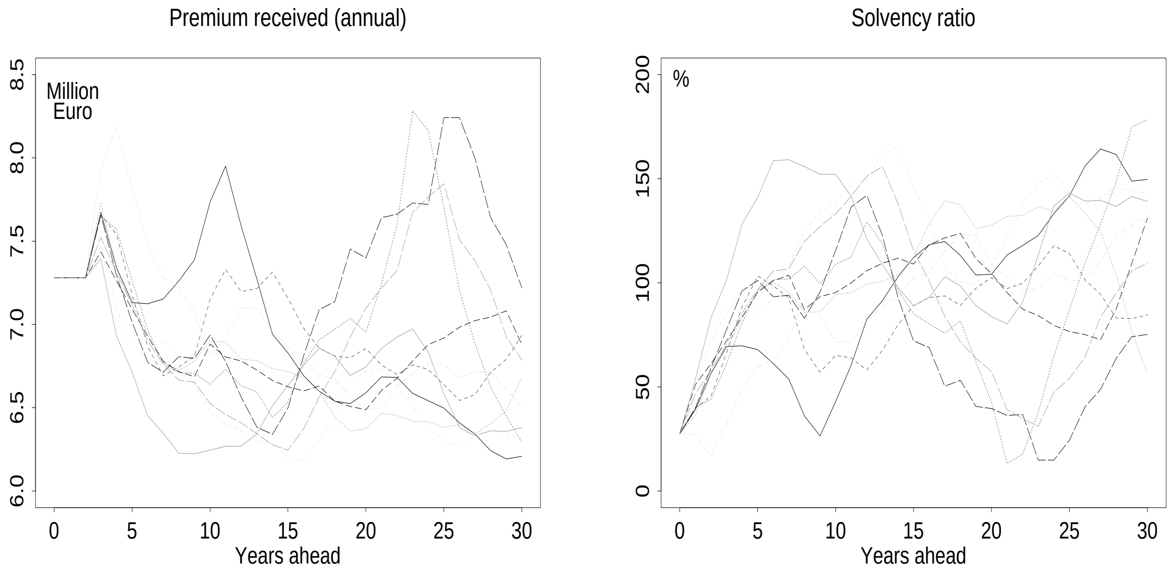

*Figure 11.4 Fluctuations (10 scenarios) in premia received (***left***) and solvency ratio (***right***) under the model described in the text.*

1.6 Ruin theory under the random walk

Introduction

The preceding section dealt with the distribution of $\mathcal{V}_k$ at fixed k. We may also be concerned with the risk of the portfolio running out of money at some undefined point in the future. If so, it becomes necessary to study the *joint* distribution of all $\mathcal{V}_k$. This is the ruin problem, introduced in Chapter 3 and tackled there through Monte Carlo. The purpose of this section is to attack the same issue through simplified conditions which permit simple, *general* statements on how the risk processes behave. Finanacial earnings will now be ignored, and the portfolio does not change with time, neither in composition nor in risk per policy. The process $\{\mathcal{V}_k\}$ follows under these circumstances a random walk (Section 5.6), defined through

$$\mathcal{V}_k = \mathcal{V}_{k-1} - Y_k \qquad \text{where} \qquad Y_k = \mathcal{X}_k - J(\pi - o)h. \tag{1.27}$$

Here πh and oh are (as in the preceding section) premium income and overhead cost per policy and Y_k is *net loss*. This is the traditional convention in actuarial science and makes *negative* values those we want!

The sequence $\{Y_k\}$ is under the conditions laid out a stochastically independent one. All Y_k follow the same probabilty distribution. Their mean $\zeta = E(Y_k)$ and standard deviation $\tau = \text{sd}(Y_k)$ follow from those of $\mathcal{X}_k$; expressions are

$$\zeta = J\{\mu\xi_z - (\pi - o)\}h \qquad \text{and} \qquad \tau = \sqrt{J\mu h(\xi_z^2 + \sigma_z^2)}. \tag{1.28}$$

Probabilistic arguments will now be invoked to descibe the performance of the process $\{\mathcal{V}_k\}$ in the long run.

Deductions from the central limit theorem
If $\mathcal{V}_{\prime} = v_0$ is the initial reserve, then

$$\mathcal{V}_k = v_0 - (Y_1 + \ldots + Y_k),$$

for which

$$E(\mathcal{V}_k) = v_0 - k\zeta \qquad \text{and} \qquad \mathrm{sd}(\mathcal{V}_k) = \sqrt{k}\tau.$$

This is no more than a special case of some of the formulas of the preceding section and by the central limit theorm leads to ruin at time k according to a probability for which the mathematical expression is approximately

$$\Pr(\mathcal{V}_k \leq 0|v_0) \doteq \Phi(d_k) \qquad \text{where} \qquad d_k = \frac{-v_0 + k\zeta}{\sqrt{k}\tau}. \tag{1.29}$$

Here $\Phi(d)$ is the normal integral. Suppose $\zeta > 0$. From (1.28) left this means that

$$\pi < \mu\xi_z + o, \tag{1.30}$$

and premium, is on average, not sufficient to cover payments and overhead expenses. This is not what we want. Indeed, $d_k \to \infty$ as $k \to \infty$ so that $\Phi(d_k) \to 1$, and eventually the account is certain to be out of money. This conclusion holds even when $\zeta = 0$ (so that π balances expected payments and cost exactly). The argument is now more elaborate; see, for example Karlin and Taylor (199?). Such processes were simulated in Exercise 5.6.? and their very slow change while extending over the entire real axis was duely noted.

In practice the inequality (1.30) would mean that risk has been assessed wrongly. The normal situation is that $\zeta < 0$. Now the portfolio account $\mathcal{V}_k$ is systematically pushed upwards, but the scheme may still run out of money by accident. From now on it is assumed that $\zeta < 0$.

Kramér-Lundberg theory
Ruin theory was developed during the first part of the twentieth century, largely by Swedish actuaries and mathematicians. The criterion is similar to (**??**) in Section 3.5, except that the time horizon now is *infinite*. In mathematical terms

$$p^{\mathrm{ru}}(v_0) = \Pr\{\min(\mathcal{V}_1, \mathcal{V}_2, \ldots) < 0|\mathcal{V}_0 = v_0\}, \tag{1.31}$$

and the issue is how this quantity behaves. It is easier to run out of money the longer the time horizon, an (1.15) is therefore upper bound on ruin in *finite* time. The formulation (and proof) of the results make use of the the moment generating function

$$M(s) = E\{\exp(sY)\} \qquad \text{for } s > 0, \tag{1.32}$$

where Y is a variable in the sequence $\{Y_k\}$. Clearly $M(0) = 1$, and when differentiated

$$M'(s) = E\{Y\exp(sY)\} \qquad \text{and} \qquad M''(s) = E\{Y^2\exp(sY)\},$$

so that

$$M'(0) = E(Y) = \zeta \qquad \text{and} \qquad M''(s) > 0.$$

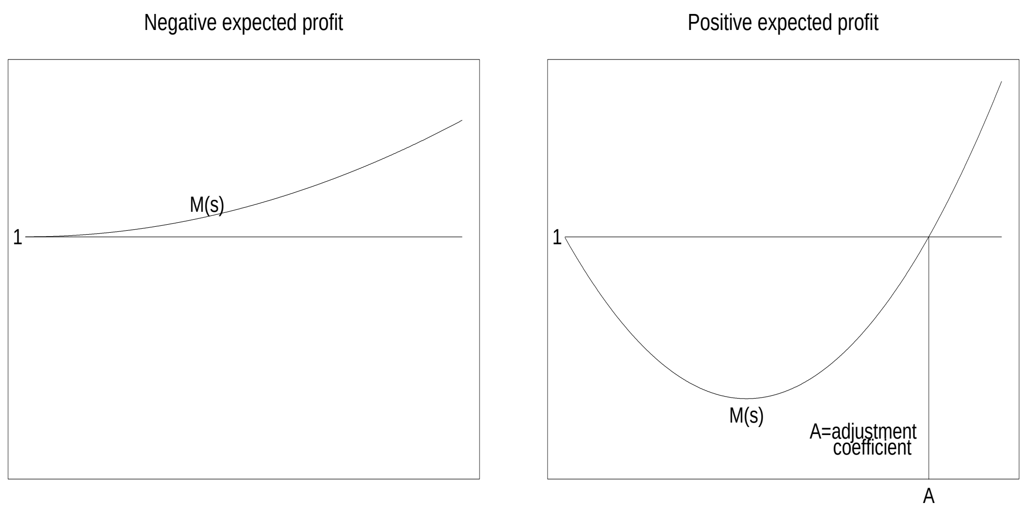

Figure 11.5. The moment generating function of $\{Y_k\}$ *under* **positive** *and* **negative** *drift.*

Moment generating functions are thus convex functions, and they have one of the two shapes in Figure 11.1 according to whether ζ is positive or negative. It is the second case which interests us. The derivative of $M(s)$ is then negative at the origin, but the curve $M(s)$ must eventually rise again since it become infinite as $s \to \infty$. It follows (Figure 11.1 right) that the equation $M(s) = 1$ has a single posive root A. This quantity is known as the **adjustment** coefficient and defines the **Lundberg** bound through

$$p^{\mathrm{ru}}(v_0) \leq \exp(-Av_0) \qquad \text{Condition: } \zeta < 0. \tag{1.33}$$

The proof is given in Section 11.6. It is possible to use this to evaluate solvency in the long run; see Exercise 11.2.1.

Lundberg's inequality shows that the ruin probability becomes zero at exponential rate as the initial capital v_0 becomes infinite. A stronger result is

$$p^{\mathrm{ru}}(v_0) \sim C\exp(-Av_0) \qquad \text{as} \qquad v_0 \to \infty \tag{1.34}$$

where the mathematical notation signifies that the ratio of the two sides is equal to one in the limit. This is due to the Swedish mathematician Harald Cramér. The constant C is

$$C = \frac{(\pi - o)/(\mu h) - \xi_z}{A\int_0^\infty z\exp(Az)\{1 - F(z)\}dz}$$

where $F(z)$ is the distribution function of the claim size.

Heavy-tailed claim size distributions

The Kramér-Lundberg theory assumes the moment generating function to *exist*, i.e. to be finite for positive values of s close to the origin. This requires the right tail of the claim size distribution to be of the Gamma type or lighter. Neither Pareto nor log-normal distributions are included, and for those there is an alternative, much more recent theory; see Asmussen (2000). It must then be assumed that there exists an $\alpha > 0$ and a constant C so that

$$\lim_{z\to\infty} \frac{1-F(z)}{Cz^{\alpha}} = 1.$$

Such functions, decaying *polynomially* as $z \to \infty$, are in mathematics known as **slowly varying**. Exercises ?? and ?? discuss them for Pareto and log-normal distributions. The replacement for the former limit (**??**) is then

$$p^{\text{ru}}(v_0) \sim \frac{\int_{v_0}^{\infty} 1-F(z)dz}{\mu h/(\pi - o) - \xi_z} \qquad \text{as} \qquad v_0 \to \infty$$

which means that the ruin probability approaches zero as $v_0 \to \infty$ at *polynomial* rather than exponential speed (i.e. *much* slower).

1.7 Mathematical arguments

Section 11.2

Basic recursions. To establish (1.4), (1.5), (1.6) the main point is to pass mean, variance and third order moments over the recursion

$$\mathcal{V}_k = \mathcal{V}_{k-1} + \Pi_k - \mathcal{O}_k - \mathcal{X}_k, \qquad k = 1, \ldots, K$$

and use appropriate formulas for such operations. For example,

$$E(\mathcal{V}_k) = E\{(1+r)\mathcal{V}_{k-1}\} + (\Pi_k - O_k) - E(\mathcal{X}_k) = (1+r)E(\mathcal{V}_{k-1}) + (\Pi_k - O_k) - E(\mathcal{X}_k)$$

which is (1.4). The derivations of the formulas for the variance and the third order moment are is similar, except that $\mathcal{X}_k$ must now be stochastically independent of $\mathcal{V}_{k-1}$; i.e. of all preceding liabiltities $\mathcal{X}_1, \ldots, \mathcal{X}_{k-1}$. Then

$$\text{var}(\mathcal{V}_k) = \text{var}\{(1+r)\mathcal{V}_{k-1}\} + \text{var}(\mathcal{X}_k) = (1+r)^2\text{var}(\mathcal{V}_{k-1}) + \text{var}(\mathcal{X}_k)$$

and

$$\nu_3(\mathcal{V}_k) = \nu_3\{(1+r)\mathcal{V}_{k-1}\} - \nu_3(\mathcal{X}_k) = (1+r)^3\nu_3(\mathcal{V}_{k-1}) - \nu_3(\mathcal{X}_k),$$

which are (1.5) and (1.6).

Formulas under constant risk The formulas (1.7, (1.8) and (1.9), valid when all risk parameters, premium income, overhead cost and portfolio size are the same for all k, makes use of the result that

$$v_k = av_{k-1} + b, \quad k = 1, 2, \ldots \qquad \text{yields} \qquad v_k = v_0 a^k + b\frac{a^k - 1}{a-1}$$

if $a \neq 1$. It can be proved by induction. If the expression for v_k is valid for at $k-1$, then

$$\begin{aligned} av_{k-1} + b &= a\left(v_0 a^{k-1} + b\frac{a^{k-1}-1}{a-1}\right) + b \\ &= v_0 a^k + b\left(a\frac{a^{k-1}-1}{a-1} + 1\right) = v_0 a^k + b\frac{a^k - 1}{a-1}) = v_k, \end{aligned}$$

and the expression is valid for v_k too. To verify the expression (1.10) for the expectation take

$$a = 1 + r \qquad \text{and} \qquad b = (\Pi_k - \mathcal{O}_k) - E(\mathcal{X}_k) = J(\pi - o - \mu\xi_z).$$

The other two are even simpler. Now $v_0 = 0$, and $a = (1+r)^2$ (for the variance) and $a = (1+r)^3$ (third order moment) whereas b must be identified with $\text{var}(\mathcal{X}_k)$ and with $\nu_3(\mathcal{X}_k)$ respectively.

Section 11.4
Negative binomial regression
The model in Section 11.4 reads

$$n_{jk}|\lambda_{jk} \sim \text{Poisson}(\lambda_{jk}) \qquad \text{where} \qquad \lambda_{jk} = T_{jk}\xi_{jk}S_k.$$

Here T_{jk} are the exposure rates, ξ_{jk} contained the regression terms and the sequence $S_1, S_2, \ldots$ independent with $S_k \sim \text{Gamma}(\alpha_k)$. We shal determine the log likelihood function of the observed counts $\{n_{jk}\}$. The assumptions mean that n_{jk} and $n_{j'k'}$ are stochastically *dependent* for $k = k'$ and *independent* otherwise. In other words, we may examine joint density functions $f(n_{1k}, \ldots, n_{Jk})$ and add their logarithms at the end.

Introduce the total number of claims in period k as

$$n_{\cdot k} = n_{1k} + \ldots + n_{Jk},$$

which yields the factorization

$$f(n_{1k}, \ldots, n_{Jk}) \;=\; \underbrace{f(n_{1k}, \ldots, n_{Jk}|n_{\cdot k})}_{\textit{multinomial}} \quad \cdot \quad \underbrace{f(n_{\cdot k})}_{\textit{negative binomial}}.$$

The interpretation of the two factors is due to results from Chapter 8. First consider the total number of claims $n_{\cdot k}$. It is by assumption conditionally Poisson with parameter

$$\sum_{j=1}^{J} \lambda_{jk} = \left(\sum_{j=1}^{J} T_{jk}\xi_{jk}\right) S_k.$$

But $S_k \sim \text{Gamma}(\alpha_k)$ and the unconditional distribution is therefore negative binomial with parameters

$$\alpha_k \qquad \text{and} \qquad p_{\cdot k} = \frac{\alpha_k}{\alpha_k + \sum_{j=1}^{K} T_{jk}\xi_{jk}};$$

see Section 8.5 and (**??**) in particular. For the other factor it is from Section 8.3 known that

$$n_{1k}, \ldots, n_{Jk}|n_{\cdot k} \sim \text{multinomial}(n_{\cdot k}, q_{1k}, \ldots, q_{Jk}) \quad \text{where} \quad q_{jk} = \frac{T_{jk}\xi_{jk}}{\sum_{j=1}^{K} T_{jk}\xi_{jk}}.$$

Here the random factor S_k has canceled in the ratios defining q_{jk}.

In summary the likelihood function becomes

$$\mathcal{L} \;=\; \underbrace{\mathcal{L}_1}_{\textit{multinomial}} \quad + \quad \underbrace{\mathcal{L}_2}_{\textit{negative binomial}}$$

where

$$\mathcal{L}_1 = \sum_{k=1}^{K}\sum_{j=1}^{J} n_{jk}\log(q_{jk}) \qquad \text{for} \qquad q_{jk} = \frac{T_{jk}\xi_{jk}}{\sum_{j=1}^{J} T_{jk}\xi_{jk}}$$

and

$$\mathcal{L}_2 = \sum_{k=1}^{K}\{\log\{\Gamma(n_{\cdot k}+\alpha_k)\} - \log\{\Gamma(\alpha_k)\} + \alpha_k\log(p_{\cdot k}) - n_{jk}\log(1-p_{\cdot k})\}$$

Any model for $\{\xi_j k\}$ and $\{\alpha_k\}$ can be fed this scheme. Numerical methods is necessary to optimize.

Section 11.6

Lundberg's inequality The proof based on a simple idea taken from Sundt (1991). Let $g(y)$ being the density function of Y and introduce $g_A(y) = \exp(Ay)g(y)$ where A is the adjustment coefficient. Since by Figure 11.4

$$M(A) = \int_{-\infty}^{\infty}\exp(Ay)g(y)dy = 1 \qquad \text{and} \qquad M'(A) = \int_{-\infty}^{\infty} y\exp(Ay)g(y)dy > 0$$

it follows that $g_A(y)$ is a density function with positive mean. Ruin under such models is certain.

Let $S_k = Y_1 + \ldots + Y_k$ and note that ruin at time k means that

$$B_k = (S_1, \ldots, S_{k-1} < v_0 \;\text{ and }\; S_k \geq v_0).$$

These events are disjoint so that

$$p^{\text{ru}}(v_0) = \Pr(B_1) + \Pr(B_2) + \ldots.$$

But

$$\Pr(B_k) = \underset{B_k}{\int \ldots \int} g(y_1)\cdots g(y_k)dy_1\ldots dy_k$$

which after inserting $g(y) = \exp(-Ay)g_A(y)$ becomes

$$\Pr(B_k) = \underset{B_k}{\int \ldots \int} \exp\{-A(y_1+\ldots+y_k)\}g_A(y_1)\cdots g_A(y_k)dy_1\ldots dy_k.$$

Since $y_1 + \ldots + y_k \geq v_0$ on B_k, this is bounded above by

$$\Pr(B_k) \leq \exp(-Av_0)\underset{B_k}{\int \ldots \int} g_A(y_1)\cdots g_A(y_k)dy_1\ldots dy_k = \exp(-Av_0)\Pr_A(B_k),$$

where the subscript A refer to the model $g_A(y)$. Adding over all k on both sides yields

$$p^{\text{ru}}(v_0) = \sum_{k=1}^{\infty}\Pr(B_k) \leq \exp(-Av_0)\sum_{k=1}^{\infty}\Pr_A(B_k) = \exp(-Av_0)p_A^{\text{ru}}(v_0) = \exp(-Av_0)$$

since ruin under the *second model* $g_A(y)$ is certain. This is the inequality (1.33).

		Car types (according to gear)	
Intercept	-5.407 (0.010)	Manual	0
Monthly drift	0.0057% (0.0002%)	Automatic	-0.340 (0.005)
Seasonal		Distance limits (in 1000 km)	
January	0	8	0
February	-0.088 (0.010)	12	0.097 (0.006)
Mars	-0.323 (0.011)	16	0.116 (0.007)
April	-0.426 (0.012)	20	0.198 (0.008)
May	-0.270 (0.011)	25	0.227 (0.019)
June	-0.406 (0.012)	30	0.308 (0.012)
July	-0.294 (0.011)	No limit	0.468 (0.019)
August	-0.248 (0.011)		
September	-0.137 (0.011)		
October	-0.133 (0.012)		
November	-0.001 (0.011)		
December	-0.132 (0.012)		

Table 11.2 Estimated (monhtly) coefficents in Poisson regression for the automobile data with standard deviation in parathesis

1.8 Further reading

Kaminsky, K. (1987). Prediction of IBNR claim counts by modeling the distribution of report lags. *Insurance: Mathematics & economics.* 6, 151-159. belongs to the class of **non-linear state space models**, much studied in engineering and statistics; see (for example) Durbin and Koopmans (2001).

1.9 Exercises

Exercise 9.1
Let Z_l be the size of a claim settled l years after its inception. A *simple* description of a delay effect in the model for Z_l is to assume

$$Z_l = Z_0 \exp(\theta_l)$$

where θ_l is a parameter. If $\theta_0 = 0$ (as will be assumed), Z_0 is the claim size of an incident settled during the first year.

a) How do mean and standard deviation of Z_l relate to $E(Z_0)$ and sd$(Z_0$?

Suppose a simulation algorithm of Z_0 is available.

b) How do you sample Z_l?

c) How do you integrate the model into algorithm 9.1?

Exercise 9.2

This is a continuation of the preceding exercise. Suppose $z_{l1}, \dots z_{lM_l}$ are claimes settled after l years of delay, $l = 0, 1, \dots, K$. A simple way to estimate θ_l is to utilize that

$$\log(Z_l) = \log(Z_0) + \theta_l.$$

Thus

$$\hat{\theta}_l = \frac{1}{M_l}\sum_{i=1}^{m_l} z_{li} - \sum_{i=1}^{m_l} z_{0i}.$$

The shape of the claim size distribution may be dertemined from

$$z'_{li} = z_{li}\exp(-\hat{\theta}_l),$$

which removes the effect of the delay. The set of all $\{z'_{li}\}$ have then approximately the same distribuiton and can be treated by one of the methods in chapter 7.

For expectation independence is *not* needed, which means that (1.4) holds *generally.* We may, for example, take

$$r = E(\mathcal{R}_k) \qquad \text{and} \qquad \lambda_k = J_k E(\mu_k) h$$

making r the expected return of a complicated portfolio of assets or λ_k the expected Poission parameter a under stochastic process model for claim intensitites. It is not necessary to apply Monte Carlo to get hold of *expectation* under such models. Recursions under these more general conditions can be developped even for the variance and third order moment, but much of the simplicity is now lost, and, as pointed out above, the normal and normal power approximations can no longer be used.

Present values

Another special case is the present value

$$\text{PV}_k = \frac{\mathcal{X}_1}{1+r} + \dots + \frac{\mathcal{X}_k}{(1+r)^k} \tag{1.35}$$

of a stream of liabiltities $\{\mathcal{X}_k\}$. Suppose we in the recursion for $\{\mathcal{V}_k\}$ take $v_0 = 0$, $\pi = 0$ and $o = 0$. Then

$$\mathcal{V}_k = (1+r)\mathcal{V}_{k-1} - \mathcal{X}_k, \qquad \mathcal{V}_0 = 0$$

from which it follows that

$$\text{PV}_k = -\frac{\mathcal{V}_k}{(1+r)^k}.$$

This tells us that the approximation formulas for the distribution of $\mathcal{V}_k$ with minor modfications apply to present values too. Indeed,

$$E(\text{PV}_k) = J\mu\xi_z h\frac{1-(1+r)^{-k}}{r}, \tag{1.36}$$

$$\text{sd}(\text{PV}_k) = \text{sd}(\mathcal{X})\sqrt{\frac{1-(1+r)^{-2k}}{(1+r)^2-1}}, \tag{1.37}$$

$$\text{skew}(\text{PV}_k) = \text{skew}(\mathcal{X})\frac{(1+r)^{3k}-1}{\{(1+r)^{2k}-1\}^{3/2}}\frac{\{(1+r)^2-1\}^{3/2}}{(1+r)^3-1}, \tag{1.38}$$

which can be plugged into the normal and Normal Power approximations, as before.

Seasonal components are plotted on their own in Figure 11.2 left and are far from smooth (as is evident from the plot on the right too). Consider, for example, November, December and January, the start of the winter in Nothern Europe. Is it plausible that there are *fewer* car accidents in December? The tiny standard deviations in Table 11.2 means that discrepancies are statistically significant, but that is not in itself very interesting. Conditions in December certainely deviated during *those years the data were collected*, but nothing more can be read into it. What it also suggests is that randomness in background conditions is considerable and that it could be better to allow monthly parameters to be stochastic rather han fixed. That is developed next.