1 Solvency and pricing

1.1 Introduction

Principal tasks in general insurance is solvency and pricing. Solvency is financial control of liabilities under near worst-case scenarios. Target is the (upper) percentiles q_ϵ of the portfolio liability $\mathcal{X}$, known as the **reserve**. Modelling was reviewed in the preceding chapters, and the issue now is computation. We may need the entire distribution of $\mathcal{X}$, for example when dealing with re-insurance (section 10.6). Monte Carlo is the obvious, *general* tool, but some problems can be handled by simpler Gaussian approximations, possibly with a correction for skewness added. Computational methods for solvency are discussed in the next two sections.

The second main topic is the pricing of risk. This has a market side. A company will gladly charge what people are willing to pay! Yet a core should be the pure premium $\pi = E(X)$ or $\Pi = E(\mathcal{X})$; i.e. the expected policy or portfolio payout during a certain period of time. Evaluations of those are important not only as a basis for pricing, but also as an aid to decision making. Not all risks are worth taking! Pricing or **rating** methods follow two main lines. The first one draws on claim histories of *individuals*. Those with good records are considered lower risks and rewarded (premium reduced), those with bad ones punished (premium raised). The traditional approach is through the theory of **credibility**, a classic presented in Section 10.5. Price differentials can also be administered according to the experience with *groups*. Credibility is a possible approach even now, but it is often more natural to use regression where risk is allowed to depend on explanatory variables such as age, sex, what kind of car you own, where your residence are and so on. Section 10.4 makes use of regression methods from earlier chapters.

1.2 Portfolio liabilities by simple approximation

Introduction

The portfolio loss $\mathcal{X}$ for independent risks becomes Gaussian when the number of policies $J \to \infty$. This is a consequence of the central limit theorem and leads to straightforward assessments of the reserve that avoid detailed probabilistic modelling. The method is useful due to its simplicity, but the underlying conditions are too restrictive for it to be the only one. Normal approximations underestimate risk for small portfolios and in branches with large claim severities. Some of that is rectified by taking the skewness of $\mathcal{X}$ into account, leading to the so-called **NP**-version. The purpose of this section is to review these simple approximation methods, show how they are put to practical use and indicate their accuracy and range of application.

Normal approximations

Let μ be claim intensity and ξ_z and σ_z mean and standard deviation of the individual losses. If they are the same for all policy holders, the mean and standard deviation of $\mathcal{X}$ over a period of length T become

$$E(\mathcal{X}) = a_0 J, \qquad \text{and} \qquad \text{sd}(\mathcal{X}) = a_1\sqrt{J};$$

where

$$a_0 = \mu T \xi_z, \qquad \text{and} \qquad a_1 = (\mu T)^{1/2}(\sigma_z^2 + \xi_z^2)^{1/2}; \tag{1.1}$$

see Section 6.3. This leads to to the true percentile q_ϵ being approximated by

$$q_\epsilon^{\text{No}} = a_0 J + a_1 \phi_\epsilon \sqrt{J} \tag{1.2}$$

where ϕ_ϵ is the (upper) ϵ percentile of the the standard normal distribution. Estimates of μ, ξ_z and σ_z are required, but *not* the entire claim size distribution. Detailed modelling can be avoided by using the sample mean and the sample standard deviation as estimates $\hat{\xi}_z$ and $\hat{\sigma}_z$, but they can also be found by fitting a parametric distribution.

The approximation (1.2) is nearly always valid for large portfolios even when μ, ξ_z and σ_z depend on j. This is due to the Lindeberg extension of the central limit theorem; see Appendix A.4. The coefficients a_0 and a_1 now become

$$a_0 = \frac{T}{J}\sum_{j=1}^{J}\mu_j \xi_{zj} \qquad \text{and} \qquad a_1 = \sqrt{\frac{T}{J}\sum_{j=1}^{J}\mu_j(\sigma_{zj}^2 + \xi_{zj}^2)}. \tag{1.3}$$

which reduce to (1.1) when all parameters are equal. With μ_j, ξ_{zj} and σ_{zj} available on file this method gives (when applicable) a quick appraisal of the reserve.

Still another version emerges when the underlying parameters are random. The most important special case is when claim frequencies $\mu_1, \dots, \mu_J$ are drawn (independently of each other) from a distribution with common mean and standard deviation ξ_μ and σ_μ. If the mean and standard deviation ξ_z and σ_z of the size of claims are fixed, the coefficients (1.1) now become

$$a_0 = \xi_\mu T \xi_z, \qquad \text{and} \qquad a_1 = T^{1/2}\{\xi_\mu(\sigma_z^2 + \xi_z^2) + \sigma_\mu^2 \xi_z^2\}^{1/2}, \tag{1.4}$$

see (**??**) and (**??**) in Section 6.3. The following example examines the numerical impact.

Example: Motor insurance

The Norwegian autmobile portfolio was introduced in Chapter 8. Its parameters are

$$\underbrace{\hat{\xi}_\mu = 5.6\%, \qquad \hat{\sigma}_\mu = 2.0\%}_{\textit{annual parameters}} \qquad \text{and} \qquad \underbrace{\hat{\xi}_z = 0.30, \qquad \hat{\sigma}_z = 0.35,}_{\textit{unit: 1000 euro}}$$

where the model for claim intensity was identified in Section 8.3. The loss parameters $\hat{\xi}_z$ and $\hat{\sigma}_z$ exclude personal injuries and were obtained from almost 7000 incidents; see also Section 10.4. This is sufficient information to evaluate the reserve under the normal approximation. With $J = 10000$ policies (and $T = 1$) the coefficients a_0 and a_1 are obtained from (1.1) and (1.4) and leads to the following assessments (in 1000 euro):

Fixed claim frequency			*Random claim frequency*	
1860,	1934	and	1860,	1935.
95% *reserve*	99% *reserve*		95% *reserve*	99% *reserve*

Note how little hetereogeneity among policy holders matters! The message was the same in Section 6.3. Even a quite substantial variation among individuals (as in the present example) is of no more than minor importance for the reserve.

The normal power approximation

Normal approximations are refined by adjusting for skewness in $\mathcal{X}$. This is in actuarial science known as the **normal power** (or **NP**) approximation, in reality the leading term in a series of corrections to the central limit theorem. Another name is the Cornish-Fisher expansion; see Feller (1971) for a probabilistic introduction and Hall (1992) for one in statistics. The underlying theory is beyond the scope of this book, but a brief sketch of the structure is indicated in Section 10.7. Only the pure Poisson model is considered below. The extension to the negative binomial and other models is treated in Daykin, Pentikäinen and Pesonen (1994), but as has been argued earlier, the practical impact is limited.

Let ζ_z be the skewness coefficient of the claim size distribution. The modified approximation then reads

$$q_\epsilon^{\mathrm{NP}} = q_\epsilon^{\mathrm{No}} + a_2(\phi_\epsilon^2 - 1)/6 \qquad \text{where} \qquad a_2 = \frac{\zeta_z\sigma_z^3 + 3\xi_z\sigma_z^2 + \xi_z^3}{\sigma_z^2 + \xi_z^2}. \tag{1.5}$$

The extra term is due to skewness and is in practice positive; see Section 10.7 for the justification. When (1.1) replaces the normal approximation q_ϵ^{No}, this yields

$$q_\epsilon^{\mathrm{NP}} = \underbrace{a_0 J + a_1\phi_\epsilon\sqrt{J}}_{\textit{the normal component}} + \underbrace{a_2(\phi_\epsilon^2 - 1)/6}_{\textit{NP correction}}$$

which is a series in falling powers of $\sqrt{J}$. The NP correction term is *independent* of portfolio size.

To use the approximation in practice skewness ζ_z must be estimated in addition to ξ_z and σ_z (μ as well). There is no new ideas in this. We may fit a parametric family to the historical data or use the sample skewness coefficient introduced in Section 9.2. The mathematics becomes more complicated to write down when the parameters vary over the portfolio, but the approximation is still valid; see Section 10.7.

Example: Danish fire claims

Consider a portfolio for which

$$\underbrace{\hat{\mu} = 1\%}_{\textit{annual}} \qquad \text{and} \qquad \underbrace{\hat{\xi}_z = 3.385, \quad \hat{\sigma}_z = 8.507, \quad \hat{\zeta}_z = 18.74.}_{\textit{Unit: Million Danish kroner}}$$

The parameters for claim size are those found for the Danish fire data in Chapter 9. With $J = 1000$ and $J = 100000$ policies the assessments of the reserve becomes those in Table 10.1. The NP correction has considerable impact on the small portfolio on the left, raising the 99% the reserve by as much as 60%. The principal reason is the losses being strongly skewed towards the right (with coefficient exceeding 18). When the number of policies is higher, the relative effect of the adjustment is smaller. With 100000 policies the difference between the two methods is of minor importance and their almost common assessment one to be trusted.

But what about the small portfolio? The huge impact of the NP correction on the left in Table 10.1 is ominous and should make us suspicous. Indeed, the more reliable Monte Carlo assessments in the next section match neither. The approximations of this section is likely to work best when the NP term isn't a dominating one.

Money unit: Million DKK

	Portfolio size: $J = 1000$		Portfolio: $J = 100000$	
	95% *reserve*	99% *reserve*	95% *reserve*	99% *reserve*
Normal	80	100	3860	4060
Normal power	120	160	3900	4120

Table 10.1 *Normal and normal power approximations to the reserve for the Danish fire claims.*

1.3 Portfolio liabilities by simulation

Introduction

Monte Carlo has several advantages over the methods of the preceeding section. It is more *general* (no restriction on use), more *versatile* (easier to adapt changing circumstances) and better suited for long time horizons (Chapter 11). But the method is slow computationally and doesn't it demand the entire claim size distribution whereas the normal approximation could do with only mean and variance? The last point is deceptive. If the portfolio size is so large that the Gaussian is a reasonable approximation, the claim size distribution (apart from mean and variance) doesn't matter anyhow.

Computational speed is unlikely to be a problem, at least not with fast compilers such as C or Fortran. To give you an idea suppose there are 1000 policies with average claim frequency $\mu T = 5\%$. Then a Fortran implementation of Algorithm 10.1 on a T60p processor produced one thousand portfolio simulations in 0.02 seconds when the claim size distribution was the empirical one. The Gamma distribution (laborious to sample) required twice as much, still only 0.04 seconds. Computer time is not very far from being proportional to the mean number of claims $J\mu T$.

A skeleton algorithm

Portfolio liabiltity is a central issue in general insurance, and its seems worthwile to sketch a general method that collects algorithms spread over several chapters. Suppose claim intensities $\mu_1, \ldots, \mu_J$ are stored on file along with J different claim size distributions and payments functions $H_1(z), \ldots, H_J(z)$. If Algorithm 2.10 are used for the Poisson sampling, the programming steps can be organized as follows:

Algorithm 10.1 Portfolio liabilities in the general case

Input: $\lambda_j = \mu_j T$ $(j = 1, \ldots, J)$, claim size models, $H_1(z), \ldots, H_J(z)$.
$\mathcal{X}^* \leftarrow 0$
For $j = 1, \ldots, J$ do
  Draw $U^* \sim$ uniform and $S^* \leftarrow -\log(U^*)$
  Repeat while $S^* < \lambda_j$
    Draw claim size Z^* %*Might depend on j*
    $\mathcal{X}^* \leftarrow \mathcal{X}^* + H_j(Z^*)$ %*Add loss*
    Draw $U^* \sim$ uniform and $S^* \leftarrow S^* - \log(U^*)$ %*Update for Poisson*
Endfor
Return $\mathcal{X}^*$

Distribution		Distribution	
Empirical distribution	Algorithm 4.1	Extended Pareto	Algorithm 9.1
Pareto mixing	Algorithm 9.2	Weibull	Exercise 2.5.1
Gamma	Algorithms 2.13, 2.14	Fréchet	Exercise 2.5.2
Log-normal	Algorithm 2.2	Logistic	Exercise 2.5.3
Pareto	Algorithm 2.8	Burr	Exercise 2.5.4

Table 10.2 *List of claim size algorithms*

Poisson sampling has been integrated into the code. The algorithm goes through the entire portfolio and add costs of settling incidents until the citerion on Line 4 is *not* satisfied. There are many different algorithms for Line 5. Table 10.2 lists examples from this book.

Often individual losses require most of the computer time. If so, there is little point in faster Poisson samplers such as the guide tables (Section 4.2) and the Atkinson method (Section 2.6) which may *not* bring worthwile improvements. Neither is speed enhanced very much when risks are identical and the algorithm built around the portfolio number of claims $\mathcal{N}$.

Danish fire data: The impact of the claim size model

The Danish fire data was examined in Section 9.6 and a number of models were tried. Some worked better than others, and Table 10.3 shows how the fit or lack of it is passed on to the reserve. Models considered were the empirical distribution function without or with Pareto mixing for the extremes, pure Pareto, Gamma and log-normal. All were fitted the historical fire claims as described in Chapter 9. The portfolio size were $J = 1000$ with annual claim rate $\mu = 1\%$, producing no more than 10 claims per year on average. Ten million simulations were used, making Monte Carlo uncertainty very small indeed.

The model scenario is the same as on the left in Table 10.1 and testfies to the difficulty of calculating the reserves for small portfolios. On its own the empirical distribution function underestimates risk, but it seems to work well when mixed with the Pareto distribution, and the results are not overly dependent on where the threshold b is placed. Another well-fitting model in Section 9.6 was the Gamma distribution on log-scale, and the reserve calculated under it does not deviate much from Pareto mixing. Other models in Section 9.6 were grossly in error, and produce strongly deviating results here. If you compare with the normal power method in Table 10.1 you will discover that it over-shoots at 95% and under-shoots at 99%.

[a] *EDF: The empirical distribuiton* [b] *Thresholds are* 5%, 10%, 25%, 50% [c] *Log-transformed claims*

	EDF[a]	EDF[a] with Pareto above b[b]				Other claim size models		
Reserve		b=10	b=5.6	b=3.0	b=1.8	Pareto	Gamma[c]	Log-normal
95%	72	100	104	105	100	71	94	49
99%	173	200	217	230	225	137	214	61
99.97%	330	590	870	1400	1750	900	1944	84

Table 10.3 *Calculated reserves for the Danish fire data. Money unit: Million Danish kroner (about* 8 *Danish kroner for one euro).*

Reserves at level 99.97% have been added. Luckily those figures are not in demand! The results are a mess of unstability, an example of the extreme difficulty of evaluations very far out into the tails of a distribution where they become sensitive to modelling details. Percentiles that close to one are rarely needed in insurance, but they are used by rating bureaus in finance.

1.4 Differentiated pricing through regression

Introduction

Very young male drivers or owners of fast cars are groups of clients notoriously more risky than others, and it may not be unfair to charge them more. The technological development which makes it easier to collect and store information with bearing on risk, can only further such practice. A picture of how insurance incidents and their cost are connected to circumstances, conditions and the people causing them must be built up from experience, and a principal tool is regression, typically on log-linear form. The purpose of this section is to indicate how Poisson, Gamma and log-normal regression from the preceding chapters are put to work.

Explanatory variables (observations, registrations, measurements) $x_1 \dots, x_v$ are then linked to claim intensity μ and mean loss per event ξ_z through

$$\log(\mu) = b_{\mu 0}x_0 + \dots + b_{\mu v}x_v \qquad \text{and} \qquad \log(\xi_z) = b_{\xi 0}x_0 + \dots + b_{\xi v}x_v,$$

where $b_{\mu 0}, b_{\mu 1} \dots$, and $b_{\xi 0}, b_{\xi 1}, \dots$ are coefficients. By default $x_0 = 1$, a convention introduced to make formulae neater. The explanatory variables do not have to be the same for μ and ξ_z, but the mathematics becomes simpler to write down if they are, and we can always 'zero' irrelevant ones away; i.e. take $b_{\xi i} = 0$ if (for example) x_i isn't included in the regression for ξ_z. In motor insurance (the example below) regression relationships are typically stronger for μ than for ξ_z. Inserting the defining equations for μ and ξ_z into the pure premium $\pi = \mu T \xi_z$ yields

$$\pi = Te^{\eta} \qquad \text{where} \qquad \eta = (b_{\mu 0} + b_{\xi 0})x_0 + \dots + (b_{\mu v} + b_{\xi v})x_v,$$

and estimates for the coefficients must be supplied.

Estimates of the pure premium

The pure premium of a policy holder with $x_1, \dots, x_v$ as explanatory variables is estimated as

$$\hat{\pi} = Te^{\hat{\eta}} \qquad \text{where} \qquad \hat{\eta} = (\hat{b}_{\mu 0} + \hat{b}_{\xi 0})x_0 + \dots + (\hat{b}_{\mu v} + \hat{b}_{\xi v})x_v.$$

Here $\hat{b}_{\mu i}$ and $\hat{b}_{\xi i}$ are obtained from historical data, usually through statistical software. Assessment of their standard deviations is provided too, and we must learn how they are passed on to $\hat{\pi}$ itself. Bootstrapping (Section 7.4) can be used (as always), but there is also a simpler Gaussian technique. Since the estimated regression coefficients are often approximately normal, their sum $\hat{\eta}$ is as well, and $\hat{\pi}$ becomes log-normal. This is a large-sample result which requires (in principle) much historical data, but a robust attitude is here in order. High accuracy in error estimates isn't that important.

There are two sets of estimated coefficients $(\hat{b}_{\mu 0}, \dots, \hat{b}_{\mu_v})$ and $(\hat{b}_{\xi 0}, \dots, \hat{b}_{\xi_v})$ coming from two different regression analyses. It is usually unproblematic to assume independence *between* sets so that $(\hat{b}_{\mu i}, \hat{b}_{\xi j})$ is uncorrelated for all (i,j). If $\sigma_{\mu ij} = \mathrm{cov}(\hat{b}_{\mu_i}\hat{b}_{\mu_j})$ and $\sigma_{\xi ij} = \mathrm{cov}(\hat{b}_{\xi_i}\hat{b}_{\xi_j})$ are the covariances *within* sets, then

$$E(\hat{\eta}) \doteq \eta \qquad \text{and} \qquad \mathrm{var}(\hat{\eta}) \doteq \sum_{i=0}^{v}\sum_{j=0}^{v} x_i x_j(\sigma_{\mu ij} + \sigma_{\xi ij}) = \tau^2,$$

where the relationship on the right follows from the general variance formula for sums (rule A.20 in Table A.2). These results are passed on to $\hat{\pi}$ through the usual formulae for the log-normal which yield

$$E(\hat{\pi}) \doteq \pi \exp(\tau^2/2) \qquad \text{and} \qquad \mathrm{sd}(\hat{\pi}) \doteq E(\hat{\pi})\sqrt{\exp(\tau^2) - 1}.$$

Note that $E(\hat{\pi}) > \pi$ so that $\hat{\pi}$ is biased upwards, but usually not by very much (see below). Bias and standard deviation is estimated by

$$\underbrace{\hat{\pi}(e^{\hat{\tau}^2/2} - 1)}_{\textit{bias}}, \qquad \underbrace{\hat{\pi}e^{\hat{\tau}^2/2}\sqrt{e^{\hat{\tau}^2} - 1}}_{\textit{standard deviation}} \qquad \text{where} \qquad \hat{\tau}^2 = \sum_{i=0}^{v}\sum_{j=0}^{v} x_i x_j(\hat{\sigma}_{\mu ij} + \hat{\sigma}_{\xi ij}).$$

Here $\hat{\sigma}_{\mu ij}$ and $\hat{\sigma}_{\xi ij}$ are estimates of their variances/covariances (provided by standard software). In the formula for $\hat{\tau}^2$ take $\hat{\sigma}_{\mu ij} = 0$ or $\hat{\sigma}_{\xi ij} = 0$ if variable i or j (or both) isn't included in the regression.

Designing regression models

Log-linear regression is a general tool that offers many possibiltities within a framework that adds contributions on logarithmic scale. On the natural scale such specifications are multiplicative; i.e.

$$\mu \; = \; \underbrace{\mu_0}_{\textit{baseline}} \cdot \underbrace{e^{(b_{\mu 1}+b_{\xi 1})x_1}}_{\textit{variable 1}} \cdots \underbrace{e^{(b_{\mu v}+b_{\xi v})x_v}}_{\textit{variable } v} \qquad \text{where} \qquad \mu_0 = e^{b_{\mu 0}+b_{\xi 0}}.$$

Here μ_0 is claim intensity when $x_1 = \ldots = x_v = 0$, and the explanatory variables drive intensities up and down compared to this baseline. As an example suppose x_1 is binary, (0 for males and 1 for females). Then

$$\underbrace{\mu_m = \mu_0 e^{(b_{\mu 2}+b_{\xi v})x_2}\cdots e^{(b_{\mu v}+b_{\xi v})x_v}}_{\textit{for males}} \qquad \text{and} \qquad \underbrace{\mu_f = \mu_0 e^{b_{\mu 1}+b_{\xi 1}} e^{(b_{\mu 2}+b_{\xi v})x_2}\cdots e^{(b_{\mu v}+b_{\xi v})x_v}}_{\textit{for females}},$$

and $\mu_f/\mu_m = e^{b_{\mu 1}+b_{\xi 1}}$, is fixed and *independent of all other covariates.*

The female drivers of Section 8.3 who produced less claims than men when young and more when old are not captured by this, but modifications are possible. One way is to design crossed categories. The problem with such procedures as a general approach is that the number of parameters grows rapidly. For example, suppose there are three variables consisting of 2, 6 and 6 categories. The total number of combinations is then $2 \times 6 \times 6 = 72$, and the cross-classification comprises 72 groups. This may not appear much when the historical material is almost 200000 policy years (as in the example below). On average there would then be around 2500 policy years for each group, enough for fairly accurate assessments of claim intensities through the elementary estimate (**??**). However, historical data are often very unequally divided among such groups which makes some of the estimates highly inaccurate, and often the number of groups is much higher

[a] *Estimated shape of the Gamma distribution:* $\hat{\alpha} = 1.1$

	Intercept	Age				
		≤ 26	> 26			
Freq.	-2.43 (.08)	0 (0)	-0.55 (.07)			
Size[a]	8.33 (.07)	0 (0)	-0.36 (.06)			
	Distance limit on policy (in 1000 km)					
	8	12	16	20	25-30	No limit
Freq.	0 (0)	.17 (.04)	0.28 (.04)	0.50 (.04)	0.62 (.05)	0.82 (.08)
Size[a]	0 (0)	.02 (.04)	0.03 (.04)	0.09 (.04)	0.11 (.05)	0.14 (.08)
	Geographical regions with traffic density from **high** to **low**					
	Region 1	Region 2	Region 3	Region 4	Region 5	Region 6
Freq.	0 (0)	-0.19 (.0.4)	-0.24 (.06)	-0.29 (.04)	-0.39 (.05)	-0.36 (.04)
Size[a]	0 (0)	-0.10 (.0.4)	-0.03 (.05)	-0.07 (.04)	-0.02 (.05)	0.06 (.04)

Table 10.4 *Estimated coefficients of claim intensity and claim size for automobile data (standard deviation in parenthesis). Methods: Poisson and Gamma regression.*

than 72. Simplifications through log-linear regression enables us to dampen random error; see also Exercise 10.4.3.

Example: The Norwegian automobile portfolio
A useful case for illustration is the Norwegian automobile portfolio of Chapter 8. There are around 100000 policies extending two years back with much customer turnover. Almost 7000 claims were registered as basis for claim size modelling. Explanatory variables used are

- age (2 categories that were ≤ 26 and > 26 years)
- driving limit (6 categories)
- geographical region (6 categories).

Driving limit is a proxy for how much people drive. Age is simplified drastically compared to what would be done in practice. The regression equation for μ now becomes

$$\log(\mu) \;=\; b_{\mu 0} \;+\; \underbrace{b_{\mu 1}x_1}_{age} \;+\; \underbrace{\textstyle\sum_{i=2}^{6} b_{\mu 1}(i)x_2(i)}_{distance\ limit} \;+\; \underbrace{\textstyle\sum_{i=2}^{6} b_{\mu 1}(i)x_3(i)}_{region},$$

with a similar relation for ξ_z. Coding is the same as in Section 8.4. Note that x_1 is 0 or 1 according to the whether the individual is below or above 26. Regression methods used were Poisson (claim frequency) and Gamma (claim size).

The estimated parameters in Table 10.4 vary smoothly with the categories. As expected, the more people drive and the heavier the traffic the larger is the risk. Claim frequency fluctuates stronger than claim size (coefficients larger in absolute value). Accidents of young people appear to be both more frequent and more severe. The results in Table 10.5 yield estimates of the pure premia for the 72 groups along with their standard deviation, as explained above. Those for the region with heaviest traffic (Oslo area) is shown in Table 10.5. Estimates are smooth and might

Age	Distance limit on policy (in 1000 km)					
	80	120	160	200	250-300	No limit
≤ 26 years	365 (6.3)	442 (6.8)	497 (7.5)	656 (8.3)	750 (9.0)	951 (9.8)
> 26 years	148 (2.9)	179 (3.0)	201 (3.7)	265 (3.7)	303 (4.1)	385 (4.3)

Table 10.5 *Estimated pure premium (in euro) for Region 1 of the Scandinavian autmobile portfolio (standard deviation in parenthesis)*

be used as basis for pricing. The log-normal bias introduced above varied from 0.2 to 0.5, much smaller than the standard deviations.

1.5 Differentiated pricing through credibility

Introduction

The preceding section differentiated premium according to observable attributes such as age, sex, geographical location and so on. Other factors with impact on risk could be personal ones that are not easily measured or observed. Drivers of automobiles may be able and concentrated or reckless and inexperienced. Such things influence driving and the accidents caused. The issue which is now raised is whether it is possible to assess risk individually from people's own track records. If so, charge unequally! Related examples are shops robbed repeatedly or buildings frequently damaged which might lead to higher premia.

Rating risks from experience has been done all along by fitting models to historical data, but the focus is now different, and the approach has much in common with the Bayesian ideas of Section 7.6. Policy holders are assigned pure premia $\pi = \pi^{\mathrm{pu}}$ that are random and vary from one person to another. They can be determined by **credibility** estimation. This is a method where prior knowledge of how π varies over the portfolio is combined with individual records. The idea can be applied to groups of policy holders too, and both viewpoints are introduced below. Credibility is a classical theme in actuarial science; see Bühlman and Gisler (2005).

Credibility: Approach and modelling

The basic assumption is that policy holders carry a list of attributes ω with impact on risk. What ω is immaterial; the important thing is that it exists and has been drawn randomly for each individual. Let X be the sum of claims from a certain future period (say a year) and introduce

$$\underbrace{\pi(\omega) = E(X|\omega)}_{\textit{conditional pure premium}} \qquad \text{and} \qquad \sigma(\omega) = \mathrm{sd}(X|\omega). \tag{1.6}$$

where the notation reflects that both quantities depend on the underlying ω. As basis for pricing we seek $\pi = \pi(\omega)$, the **conditional** pure premium of the policy holder. The concept can be used on group or portfolio level too. There is now a common ω that apply to all risks jointly, and the target is $\Pi = E(\mathcal{X}|\omega)$ where $\mathcal{X}$ is the sum of claims from many individuals.

Let $X_1, \ldots, X_K$ (policy level) or $\mathcal{X}_1, \ldots, \mathcal{X}_K$ (group level) be past claims dating K years back. The most accurate estimate of π and Π from such records are (Section 6.4) the conditional means

$$\underbrace{\hat{\pi} = E(X|x_1, \ldots, x_K)}_{\textit{policy level}} \qquad \text{and} \qquad \underbrace{\hat{\Pi} = E(\mathcal{X}|x_1, \ldots, x_K)}_{\textit{group level}} \tag{1.7}$$

where $x_1, \dots x_K$ are the values of $X_1, \dots X_K$ or $\mathcal{X}_1, \dots, \mathcal{X}_K$. Claims may be broken down on frequency and individual losses (this viewpoint will be introduced later), but for the moment stay with the estimates (1.7). The issue is essentially the same on either level, and the argument will be written out for single policies. As basic framework introduce the common factor model of Section 6.3. where $X_1 \dots, X_K, X$ are identically and independently distributed given ω. Surely this is plausible? It won't be true when underlying conditions change systematically during the K periods in question; consult some of the references in Section 10.8.

Complicated modelling can be avoided by leaning on the so-called **structural parameters**. There are three of them; i.e.

$$\zeta = E\{\pi(\omega)\}, \qquad \upsilon^2 = \text{var}\{\pi(\omega)\}, \qquad \tau^2 = E\{\sigma^2(\omega)\}, \tag{1.8}$$

where ζ is the *average* pure premium over the entire population and υ and τ both represent variation. The former is caused by diversity between individuals and the latter by physical processes leading to incidents. These parameters determine mean and standard deviation of X through

$$E(X) = \zeta \qquad \text{and} \qquad \text{sd}(X) = \sqrt{\tau^2 + \upsilon^2} \tag{1.9}$$

which are verified by the rules of double expectation and double variance. Indeed,

$$E(X) = E\{E(X|\omega)\} = E\{\pi(\omega)\} = \zeta,$$

and

$$\text{var}(X) = E\{\text{var}(X|\omega)\} + \text{var}\{E(X|\omega)\} = E\{\sigma^2(\omega)\} + \text{var}\{\pi(\omega)\} = \tau^2 + \upsilon^2;$$

see also (**??**) in Section 6.3.

Linear credibility
Let $\hat{\pi}_K$ be an estimate of π based on the claim record $X_1, \dots, X_K$. The simplest procedure would be to go linear. This means that the estimate is of the form

$$\hat{\pi}_K = b_0 + b_1 X_1 + \dots + b_K X_K,$$

where $b_0, b_1, \dots, b_K$ are carefully selected coefficients. The fact that $X_1, \dots, X_K$ are condionally independent with the same distribution forces $b_1 = \dots = b_K$. Write w/K for the common value, and the estimate becomes

$$\hat{\pi}_K = b_0 + w\bar{X}_K \qquad \text{where} \qquad \bar{X}_K = (X_1 + \dots + X_K)/K. \tag{1.10}$$

A natural way to proceed is to demand that b_0 and w minimize the mean squared error $E(\hat{\pi}_K - \pi)^2$. This sets up a mathematical problem with solution

$$\hat{\pi}_K = (1-w)\zeta + w\bar{X}_K, \qquad \text{where} \qquad w = \frac{\upsilon^2}{\upsilon^2 + \tau^2/K}; \tag{1.11}$$

see Section 10.7 where the argument is given. There is a close resemblance with the Bayes estimate of the normal mean in Section 7.6. The weight w defines a compromise between the average pure premium ζ of the population and the individual record of the policy holder. Note that $w = 0$ if $K = 0$; i.e. with no claim information available the best estimate is the population average. Other interpretations are given among the exercises.

It is also proved in Section 10.7 that

$$E(\hat{\pi}_K - \pi) = 0 \qquad \text{and that} \qquad \text{sd}(\hat{\pi}_K - \pi) = \frac{\upsilon}{\sqrt{1 + K\upsilon^2/\tau^2}}. \tag{1.12}$$

The linear credibility estimate is unbiased, and its standard deviation decreases with K.

Optimal credibility

The preceding estimate is the best *linear* method, but the Bayesian estimate (1.7) offers an improvement since it is optimal among *all* methods; see Section 6.4. Now the aggregate claims $x_1, \ldots, x_K$ are broken down on annual frequencies $n_1, \ldots, n_K$ and individual losses $z_1, \ldots, z_n$ where $n = n_1 + \ldots + n_K$. Suppose claim numbers and losses are stochastically independent. Then the Bayes estimate of $\pi = E(X) = E(N)E(Z)$ becomes

$$\hat{\pi} = E(X|n_1, \ldots, n_K, z_1, \ldots, z_n) = E(N|n_1, \ldots, n_K)E(Z|z_1, \ldots, z_n),$$

and the estimation problem has been decoupled into two separate ones. Both the claim intensity μ and the mean claim size $\xi_z = E(Z)$ may vary between individuals, but often more strongly for the former, and a possible simplification is to fix ξ_z, the same for everybody. Then $\xi_z = E(Z|z_1, \ldots, z_n)$, and the preceding estimate becomes

$$\hat{\pi} = \xi_z E(N|n_1, \ldots, n_K), \tag{1.13}$$

For credibility estimation of ξ_z consult Bühlman and Gisler (2005).

A model for past and future claim numbers $N_1, \ldots, N_K, N$ is needed. The natural one is to assume conditional independence given μ with each claim number being Poisson(μT). As model for μ the customary choice is

$$\mu = \xi_\mu G \qquad \text{and} \qquad G \sim \text{Gamma}(\alpha)$$

where G is a standard gamma variable with expectation one. It is verified in Section 10.7 that the estimate (1.13) now becomes

$$\hat{\pi}_K = \zeta \frac{\bar{n} + \alpha/K}{\xi_\mu T + \alpha/K} \qquad \text{where} \qquad \bar{n} = (n_1 + \ldots + n_K)/K, \tag{1.14}$$

and the population average ζ is adjusted up or down according to whether the average claim number $\bar{n}$ is larger or smaller than its expectation $\xi_\mu T$. The error is

$$E(\hat{\pi}_K - \pi) = 0 \qquad \text{and} \qquad \text{sd}(\hat{\pi}_K - \pi) = \frac{\zeta}{\sqrt{\alpha + \xi_\mu K T}} \tag{1.15}$$

which is also proved in Section 10.7.

Credibility on group level

The preceding estimates apply to groups of policies as well. Suppose we seek $\Pi(\omega) = E(\mathcal{X}|\omega)$ where $\mathcal{X}$ is the sum of claims from a group of policy holders. Now ω is common bacground uncertainty, and the linear credibility estimate (1.11) is applied to the claim record $\mathcal{X}_1, \dots, \mathcal{X}_K$ of the entire group. The structural parameters differ from what they were above. A reasonable assumption is that individual risks are independent given ω. Then

$$E(\mathcal{X}|\omega) = J\pi(\omega) \qquad \text{and} \qquad \text{sd}(\mathcal{X}|\omega) = \sqrt{J}\,\sigma(\omega),$$

and the structural parameters (1.8) become $J\zeta$, $J^2\upsilon^2$ and $J\tau^2$ instead of ζ, υ^2 and τ^2. It follows from (1.11) that the best linear estimate is

$$\hat{\Pi}_K = (1-w)J\zeta + w\bar{\mathcal{X}}_K, \qquad \text{where} \qquad w = \frac{\upsilon^2}{\upsilon^2 + \tau^2/(JK)}, \tag{1.16}$$

Here $\bar{\mathcal{X}}_K = (\mathcal{X}_1 + \ldots + \mathcal{X}_K)/K$ is the average claim on group level. Its weight is much larger than for individual policies and increases with the group size J.

Estimation error is from (1.12)

$$E(\hat{\Pi}_K - \Pi) = 0 \qquad \text{and} \qquad \text{sd}(\hat{\Pi}_K - \Pi) = \frac{J\upsilon}{\sqrt{(1 + KJ\upsilon^2/\tau^2}}, \tag{1.17}$$

so that the method is unbiased as before. Note that

$$\frac{\text{sd}(\hat{\Pi}_K - \Pi)}{\text{sd}(\hat{\Pi}_0 - \Pi)} = \frac{1}{\sqrt{1 + KJ\upsilon^2/\tau^2}},$$

which decreases with KJ. The gain from the claim record is much higher when J is large which is an important observation. It will be suggested below that the accuracy on individual level might be poor, but it could be different for groups.

Even the optimal credibility estimates (1.13) apply on group level. The historical claim numbers $n_1, \dots, n_K$ are now aggregates from J policies. Their distribution then changes to Poisson($J\mu T$), and the only thing we have to do is to replace T with JT in (1.14) and (1.15); see Exercise 10.5.5.

How accurate is credibility estimation?

Suppose ξ_z is fixed for all policy holders and μ is random. With $\xi_\mu = E(\mu)$, $\sigma_\mu = \text{sd}(\mu)$, $\xi_z = E(Z)$ and $\sigma_z = \text{sd}(Z)$ we have

$$\pi(\mu) = E(X|\mu) = \mu T\xi_z \qquad \text{and} \qquad \sigma^2(\mu) = \text{var}(X|\mu) = \mu T(\xi_z^2 + \sigma_z^2);$$

see Section 6.3. The structural parameters (1.8) become

$$\zeta = \xi_\mu T\xi_z \qquad \upsilon^2 = \sigma_\mu^2 T^2 \xi_z^2 \qquad \text{and} \qquad \tau^2 = \xi_\mu T(\xi_z^2 + \sigma_z^2),$$

and when these expressions are inserted into in (1.12) right, we obtain for the linear credibility estimate

$$\text{sd}(\hat{\pi}_K - \pi) = \frac{\sigma_\mu T\xi_z}{\sqrt{1 + K\theta_z T\sigma_\mu^2/\xi_\mu}} \qquad \text{where} \qquad \theta_z = \xi_z^2/(\xi_z^2 + \sigma_z^2).$$

Optimal			Linear, $\sigma_z = 0.1\xi_z$			Linear, $\sigma_z = \xi_z$		
$K = 0$	$K = 10$	$K = 20$	$K = 0$	$K = 10$	$K = 20$	$K = 0$	$K = 10$	$K = 20$
200.0	193.2	187.1	200.0	193.3	187.2	200.0	196.5	193.2

Table 10.6 *Standard deviations of credibility estimates under conditons in the text.*

Accurate estimation requires the standard deviation to go down fast as K is raised, and much hinges on the ratio σ_μ^2/ξ_μ. Unfortunately, this is always a rather small number.

A numerical illustration is reported in Table 10.6 . Standard deviations of the optimal and linear credibility estimate (1.15) and (1.11) are there compared when $\xi_\mu = 5.6\%$ and $\sigma_\mu = 2\%$ annually ($T = 1$) which are the parameters of the Norwegian automobile portfolio. Additional assumptions are $\xi_z = 10000$ and $\sigma_z = 0.1\xi_z$ or $\sigma_z = \xi_z$. Errors conveyed by the computations in Table 10.6 are huge. The mean annual claim $\zeta = 10000 \cdot 0.056 = 560$, and even 20 years of experience with the same client hasn't reduced uncertainty more than a trifle. Nor is the optimal method much of an improvement over the linear one. Errors go down when σ_μ is smaller than in this example, but now the historical record has even less impact. On group level the picture might not be the same at all; see Exercise 10.5.3.

Finding the parameters

It remains to determine the parameters underlying the credibility estimates. How this can be done for claim numbers was discussed was discussed in Section 8.3, and only linear credibility is treated here. Historical data for J policies that have been in the company $K_1, \ldots, K_J$ years are then of the form

$$\begin{array}{cccc|cc} 1 & x_{11} & \ldots & x_{1K_1} & \bar{x_1} & s_1 \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \\ J & x_{J1} & \ldots & x_{JK_J} & \bar{x_J} & s_J, \\ \textit{Policies} & & \textit{Annual claims} & & \textit{mean} & \textit{sd} \end{array}$$

where the the j'th row $x_{j1}, \ldots, x_{jK_j}$ are the annual claims from client j and $\bar{x}_j$ and s_j its mean and standard deviation. The following estimates are essentially due to Sundt (1983) with a forerunner in Bühlman and Straub (1970) and even in the biostatistical literature; see Sokal and Rohlf (1981), Section 9.2.

Let $\mathcal{K} = K_1 + \ldots + K_J$. Unbiased, moment estimates of the structural parameters are then

$$\hat{\zeta} = \frac{1}{\mathcal{K}} \sum_{j=1}^{J} K_j \bar{x}_j, \qquad \hat{\tau}^2 = \frac{1}{\mathcal{K} - J} \sum_{j=1}^{J} (K_j - 1) s_j^2 \tag{1.18}$$

and

$$\hat{\upsilon}^2 = \frac{\sum_{j=1}^{J} (K_j/\mathcal{K})(\bar{x}_j - \hat{\xi})^2 - \hat{\tau}^2 (J-1)/\mathcal{K}}{1 - \sum_{j=1}^{J} (K_j/\mathcal{K})^2}; \tag{1.19}$$

for verfication see Section 10.7. The expression for $\hat{\upsilon}^2$ may be negative. If it is, the pragmatic (and sensible) position is to take $\hat{\upsilon} = 0$. Variation in the individual pure premium over the portfolio is then too small to be detected.

1.6 Re-insurance

Introduction

Re-insurance was introduced in Section 3.2. Parts of **primary** risks placed with a **cedent** are now passed on to **re-insurers** who may in turn go to other re-insurers leading to a global network of risk sharers. Re-insurers may provide cover to incidents far away both geographically and in terms of intermediaries, but for the original clients at the bottom of the chain all this is irrelevant. For them re-insurance instruments used higher up are without importance as long as the companies involved are solvent. These arrangements are ways to spread risk and may enable small or medium-sized companies to take on heavier responsibilites than its own capital base allows.

Method doesn't change much from ordinary insurance. The primary risks rest with cedents, and the stochastic modelling is the same as before. Cash flows differ, but those are merely modifications handled through fixed functions $H(z)$ defining the payment rules and are easily taken care of by Monte Carlo (Section 3.3). The economic impact may be huge, the methodological not. This section outlines some of the most common contracts and indicate consequences for pricing and solvency.

Types of contracts

Re-insurance contracts may apply to single events or to sums of claims affecting the entire portfolio. These losses (denoted Z and $\mathcal{X}$) are then divided between re-insurer and cedent according to

$$\underbrace{Z^{\text{re}} = H(Z), \quad Z^{\text{ce}} = Z - H(Z)}_{\textit{single events}} \qquad \text{and} \qquad \underbrace{\mathcal{X}^{\text{re}} = H(\mathcal{X}), \quad \mathcal{X}^{\text{ce}} = \mathcal{X} - H(\mathcal{X}),}_{\textit{on portfolio level}} \tag{1.20}$$

where $0 \leq H(z) \leq 1$. Here Z^{ce} and $\mathcal{X}^{\text{ce}}$ are the *net* cendent responsibility after subtracting re-insurance.

One of the most common contracts is the $a \times b$ type considered in Chapter 3. When drawn up in terms of single events, re-insurer and cedent responsibilities are

$$Z^{\text{re}} = \begin{array}{ll} 0, & \text{if } Z < a \\ Z - a, & \text{if } a \leq Z < a + b \\ b - a, & \text{if } Z \geq a + b, \end{array} \qquad \text{and} \qquad Z^{\text{ce}} = \begin{array}{ll} Z, & \text{if } Z < a \\ a, & \text{if } a \leq Z < a + b \\ Z - b, & \text{if } Z \geq a + b, \end{array}$$

where $Z^{\text{re}} + Z^{\text{ce}}$=Z. The lower bound a is the **retention** limit of the cedent who must cover all claims below. Responsibility (i.e. Z^{ce}) appears unlimited, but in practice there is usually a maximum insured sum S that makes $Z \leq S$. If $b - a = S$, the scheme gives good cedent protection. If the upper bound b (the retention limit of the *re-insurer*) is infinite (rare in practice), the contract is known as **excess of loss**. This type of arrangement is also used with $\mathcal{X}$. Now $\mathcal{X}^{\text{re}}$ and $\mathcal{X}^{\text{ce}}$ are related to $\mathcal{X}$ in a manner similar to the previous relationships for $Z^{\text{re}}$and Z^{ce}, and if b is infinite, the treaty is known as **stopp loss**.

Another type of contract is the **proportional** one for which

$$\underbrace{Z^{\text{re}} = cZ, \quad Z^{\text{ce}} = (1 - Z)}_{\textit{single events}} \qquad \text{and} \qquad \underbrace{\mathcal{X}^{\text{re}} = c\mathcal{X}, \quad \mathcal{X}^{\text{ce}} = (1 - c)\mathcal{X}}_{\textit{on portfolio level}} \tag{1.21}$$

Risk is now shared by cedent and re-insurer in a fixed proportion. Suppose there are J separate re-insurance treaties, one for each of J contracts placed with the cedent. Such an arrangement is known as **quota share** if the constant of proportionality c is the same for all policies. Consider the opposite case where $c = c_j$ depends on the contract. Specifically, suppose that

$$c_j = \max(0, 1 - \frac{a}{S_j}) \qquad \text{so that} \qquad Z_j^{\text{re}} = \begin{array}{ll} 0 & \text{if } a \geq S_j \\ (1 - a/S_j)Z_j & \text{if } a < S_j, \end{array} \tag{1.22}$$

where S_j is the maximum insured sum of the j'th primary risk. This is known as **surplus** re-insurance. Note that a (the cedent retention limit) does not depend on j. As S_j increases from a, the re-insurer part grows.

Pricing re-insurance
Examples of pure re-insurance premia are

$$\underbrace{\pi^{\text{re}} = \mu T \xi^{\text{re}} \quad \text{for} \quad \xi^{\text{re}} = E\{H(Z)\}}_{\textit{single event contracts}} \qquad \text{and} \qquad \underbrace{\Pi^{\text{re}} = E\{H(\mathcal{X})\}}_{\textit{contracts on portfolio level}}$$

with Monte Carlo approximations

$$\underbrace{\pi^{\text{re}*} = \frac{\mu T}{m} \textstyle\sum_{i=1}^{m} H(Z_i^*)}_{\textit{single event contracts}} \qquad \text{and} \qquad \underbrace{\Pi^{\text{re}*} = \frac{1}{m} \textstyle\sum_{i=1}^{m} H(\mathcal{X}_i^*).}_{\textit{contracts on portfolio level}}$$

Simulation is usually the simplest way if you know the ropes and often takes *less* time to implement than to work out exact formulae (and the latter may not be possible at all). On portfolio level simulations $\mathcal{X}^*$ of the total portfolio loss (obtained from Algorithm 10.1) are inserted into the re-insurance contract $H(x)$

There is a useful formula for $a \times b$ contracts in terms of *single* events. If $f(z)$ and $F(z)$ are density and distribution function of Z, then the mean re-insurance claim is

$$\begin{aligned} \xi^{\text{re}} &= \int_a^{a+b} (z-a) f(z)\, dz + \int_{a+b}^{\infty} b f(z)\, dz \\ &= -(z-a)\{1 - F(z)\}\Big|_a^{a+b} + \int_a^{a+b} \{1 - F(z)\}\, dz + b\{1 - F(a+b)\} = \int_a^{a+b} \{1 - F(z)\}\, dz \end{aligned}$$

after integration by parts. Writing $F(z) = F_0(z/\beta)$ as in Section 9.2 yields

$$\pi^{\text{re}} = \mu T \int_a^{a+b} \{1 - F_0(z/\beta)\} dz, \tag{1.23}$$

which is is possible to evaluate under the Pareto distribution; i.e. when $1 - F_0(z) = (1+z)^{-\alpha}$. Then

$$\pi^{\text{re}} = \mu T \frac{\beta}{\alpha - 1} \left(\frac{1}{(1 + a/\beta)^{\alpha-1}} - \frac{1}{(1 + (a+b)/\beta)^{\alpha-1}} \right) \qquad \text{for} \qquad \alpha > 0, \tag{1.24}$$

with special treatment being needed for $\alpha = 1$ (Exercise 10.6.2).

The example

$$\mu T = 1\%,\ \ a = 50,\ \ b = 500\ \ \alpha = 2,\ \ \beta = 100 \qquad \text{gives} \qquad \pi^{\rm re} = 0.50,$$

which was used to test the accuracy of Monte Carlo. With $m = 100000$ simulations the answer was reproduced to two decimal places. Three decimals would take one hundred times more; i.e. $m = 10$ million.

The effect of inflation
Inflation drives claims upwards into the regions where re-insurance treaties apply, and contracts will be mis-priced if the re-insurance premium is not adjusted. The mathematical formulation rests on the rate of inflation I which changes the parameter of scale from $\beta = \beta_0$ to $\beta_I = (1+I)\beta_0$ (Section 9.2), but the rest of the model is as before. Fot $a \times b$ contracts in terms of single events (1.23) shows that that the pure premium $\pi_I^{\rm re}$ under inflation is related to the original one through

$$\frac{\pi_I^{\rm re}}{\pi_0^{\rm re}} = \frac{\int_a^b \{1 - F_0(z/\beta_I)\}dz}{\int_a^b \{1 - F_0(z/\beta_0)\}dz}.$$

How other types of contracts react to inflation is studied among the exercises.

Consider, in particular, the case of infinite b with Pareto distributed claims. Then

$$\frac{\pi_I^{\rm re}}{\pi_0^{\rm re}} = (1+I)\left(\frac{1 + a\beta_0^{-1}}{1 + a\beta_0^{-1}/(1+I)}\right)^{\alpha-1}$$

which is not negligable for values of α of some size; try some suitable values if $I = 5\%$, for example. The ratio is also an increasing function of α which means that the *lighter* the tail of the Pareto distribution, the *higher* the impact of inflation.

That appears to be a general phenomenon. A second example is

$$\underbrace{Z_0 \sim \text{Gamma}(\alpha)}_{\textit{orginal model}} \qquad \text{and} \qquad \underbrace{Z_I = (1+I)Z_0,}_{\textit{inflated model}}$$

and the pure premia $\pi_0^{\rm re}$ and $\pi_I^{\rm re}$ can be computed by Monte Carlo. Suppose the lower limit a is varied, the upper one b infinite and that $I = 5\%$. Then the relative change $(\pi_I^{\rm re} - \pi_0^{\rm re})/\pi_0^{\rm re}$ under variation of the shape parameter α of the Gamma distribution becomes

9%	23%	76%		17%	46%	169%
$\alpha = 1$	$\alpha = 10$	$\alpha = 100$	and	$\alpha = 1$	$\alpha = 10$	$\alpha = 100$
a median of Z_0				a upper 10% percentile of Z_0		

Note the huge increase in the effect of inflation as α moves from the heavy-tailed $\alpha = 1$ to the light-tailed, almost normal $\alpha = 100$.

The effect of re-insurance on the reserve

Number of simulations: One million

	Annual claim frequency: 1.05				Annual claim frequency: 5.25			
Upper limit (b)	0	2200	4200	10200	0	2200	4200	10200
Pure premium	0	82	92	100	0	410	460	500
Cedent reserve	2170	590	510	480	6300	3800	1800	1200

Table 10.7 *Re-insurance premium and net cedent reserve (1%) under the conditions in the text. Money unit: Million NOK (8 NOK for 1 euro).*

Re-insurance may lead to substantial reduction in capital requirements. The cedent company loses money on average, but it can get around on less own capital, and its value per share could be higher. A re-insurance strategy must balance extra cost against capital saved. An illustration is given in Table 10.7. Losses were those of the Norwegain pool of natural disasters in Chapter 7 for which a possible claim size distribution is

$$Z \sim \text{Pareto}(\alpha, \beta) \qquad \text{with} \qquad \alpha = 1.71 \text{ and } \beta = 140.$$

The re-insurance contract was a $a \times b$ arrangement per event with $a = 200$ and b varied. Maximum cedent responsibility is $S = 10200$ for each incident. Monte Carlo was used for computation.

Table 10.7 shows cedent net reserve against the pure re-insurance premium. When the claim frequency is 1.05 annually, the 1% reserve is down from 2170 to about one fourth in exchange for the premium paid. Five-doubling claim frequency yields smaller savings in per cent, but larger in value. How much does the cedent lose by taking out re-insurance? It depends on the deals available in the market. If the premium paid is $(1+\gamma^{\text{re}})\pi^{\text{re}}$ where π^{re} is pure premium and γ^{re} the loading, the average loss due to re-insurance is

$$\underbrace{(1+\gamma^{\text{re}})\pi^{\text{re}}}_{\textit{premium paid}} - \underbrace{\pi^{\text{re}}}_{\textit{claims saved}} = \underbrace{\gamma^{\text{re}}\pi^{\text{re}}}_{\textit{net loss}}.$$

In practice γ^{re} is determined by market conditions and may vary enormously in certain branches of insurance. Going from barely zero to 100% and even 200% in short time (a year or two) are not unheard of!

1.7 Mathematical arguments

Section 10.2

The normal power approximation: The NP approximation of Section 10.2 is a special case of the Cornish-Fisher expansion (Hall 1992) which sets up a series of approximations to the percentile q_ϵ of a random sum $\mathcal{X}$. The first two are

$$q_\epsilon \doteq \underbrace{E(\mathcal{X}) + \text{sd}(\mathcal{X})\phi_\epsilon}_{\textit{normal approximation}} + \underbrace{\text{sd}(\mathcal{X})\frac{1}{6}(\phi_\epsilon^2 - 1)\text{skew}(\mathcal{X})}_{\textit{skewness correction}}. \tag{1.25}$$

A fourth term on the right would involve the kurtosis, but that one isn't much in use in property insurance. The approximation become exact as the portfolio size $J \to \infty$. *Relative* error is proportional to $J^{-1/2}$ (skewness omitted) and to J^{-1} (skewness included), which means that skewness adjustments enhance accuracy a good deal.

Suppose $\mathcal{X}$ is the total portfolio liability based on identical Poisson risks with intensity μ and with ξ_z, σ_z and ζ_z as mean, standard deviation and skewness of the claim size distribution. Mean, variance and third order moment of $\mathcal{X}$ are then

$$E(\mathcal{X}) = J\mu T\xi_z, \qquad \text{var}(\mathcal{X}) = J\mu T(\sigma_z^2 + \xi_z^2), \qquad \nu_3(\mathcal{X}) = J\mu T(\zeta_z\sigma_z^3 + 3\sigma_z^2\xi_z + \xi_z^3),$$

where the third order moment is verified below (the other two were derived in Chapter 6, see Exercise 6.3.1). Skewness is $\nu_3(\mathcal{X})/\text{var}(\mathcal{X})^{3/2}$, and some straightforward manipulations yield

$$\text{skew}(\mathcal{X}) = \frac{1}{(J\mu T)^{1/2}} \, \frac{\zeta_z\sigma_z^3 + 3\sigma_z^2\xi_z + \xi_z^3}{(\sigma_z^2 + \xi_z^2)^{3/2}}.$$

The NP approximation (1.4) follows when the formulae for $\text{sd}(\mathcal{X})$ and $\text{skew}(\mathcal{X})$ are inserted into (1.25).

Skewness under hetereogeneous portfolios Suppose $\mathcal{X} = X_1 + \ldots + X_J$ where the parameters of policy j are μ_j, ξ_{zj}, σ_{zj} and ζ_{zj}, depending on j. The NP-approximation remains the same except for a new expression for $\text{skew}(\mathcal{X})$. It can be derived by utilizing $\nu_3(\mathcal{X}) = \nu_3(X_1) + \ldots + \nu_3(X_J)$ (rule A.13 of Appendix A) with a similar rule for variances. Hence

$$\text{skew}(\mathcal{X}) = \frac{\nu_3(\mathcal{X})}{\text{var}(\mathcal{X})^{3/2}} = \frac{\nu_3(X_1) + \ldots + \nu(X_J)}{\{\text{var}(X_1) + \ldots + \text{var}(X_J)\}^{3/2}}$$

into which we must insert

$$\text{var}(X_j) = \mu_j T(\sigma_{zj}^2 + \xi_{zj}^2) \qquad \text{and} \qquad \nu_3(X_j) = \mu_j T(\zeta_{zj}\sigma_{zj}^3 + 3\sigma_{zj}^2\xi_{zj} + \xi_{zj}^3).$$

for $j = 1, \ldots, J$.

The third order moment of $\mathcal{X}$ Let $\lambda = J\mu T$ be the Poisson parameter for the total number of claims $\mathcal{N}$. The third order moment of $\mu_3(\mathcal{X})$ is then the expectation of

$$\{\mathcal{X} - \lambda\xi_z\}^3 = \{(\mathcal{X} - \mathcal{N}\xi_z) + (\mathcal{N} - \lambda)\xi_z\}^3 = B_1 + 3B_2 + 3B_3 + B_4$$

where

$$B_1 = (\mathcal{X} - \mathcal{N}\xi_z)^3, \qquad B_2 = (\mathcal{X} - \mathcal{N}\xi_z)^2(\mathcal{N} - \lambda)\xi_z,$$
$$B_3 = (\mathcal{X} - \mathcal{N}\xi_z)(\mathcal{N} - \lambda)^2\xi_z^2, \qquad B_4 = (\mathcal{N} - \lambda)\xi_z^3.$$

Expectations of all these terms follow by computing the *conditional* expectation given $\mathcal{N}$ and applying the rule of double expectation. This is simple since $\mathcal{X}$ is a sum of identically and independently distributed random variables. Start with B_1. It follows from rule A.13 in Appendix A that the *conditional* third order moment of $\mathcal{X}$ is $\mathcal{N}$ times the third order moment of Z. Hence

$$E(B_1|\mathcal{N}) = \mathcal{N}(EZ_1 - \xi_z)^3 = \mathcal{N}\zeta_z\sigma_z^3 \qquad \text{which yields} \qquad E(B_1) = \lambda\zeta_z\sigma_z^3.$$

Similarly, from the sum of variance formula

$$E(B_2|\mathcal{N}) = \mathcal{N}\sigma_z^2(N - \lambda)\xi_z \qquad \text{and} \qquad E(B_2) = E\{\mathcal{N}(N - \lambda)\}\sigma_z^2\xi_z = \lambda\sigma_z^2\xi_z.$$

It has here been utilized that $E\{\mathcal{N}(N-\lambda)\} = \text{var}(\mathcal{N}) = \lambda$. For the two remaining terms

$$E(B_3|\mathcal{N}) = 0 \qquad \text{so that} \qquad E(B_3) = 0$$

and

$$E(B_4) = E(\mathcal{N}-\lambda)^3\xi_z^3 = \mu_3(\mathcal{N})\xi_z^3 = \lambda\xi_z^3$$

since $\nu_3(\mathcal{N}) = \lambda$; see Section 8.3. Binding all these expectations together yields

$$E(\mathcal{X}-\lambda\xi_z)^3 = E(B_1) + 3E(B_2) + 3E(B_3) + E(B_4) = \lambda(\zeta_z\sigma_z^3 + 3\sigma_z^2\xi_z + \xi_z^3)$$

which is $\nu_3(\mathcal{X})$.

Section 10.5
Statistical properties of $\bar{X}$. The first part of this section derives the linear credibility estimate and verifies its statistical properties. Three auxiliary results are

$$E(\bar{X}) = \zeta, \qquad \text{var}(\bar{X}) = \upsilon^2 + \tau^2/K \qquad \text{cov}\{\bar{X}, \pi(\omega\} = \upsilon^2. \tag{1.26}$$

The expectation follows from $E(\bar{X}) = E(X_1) = \zeta$. To derive the variance note that

$$E(\bar{X}|\omega) = E(X_1|\omega) = \pi(\omega) \qquad \text{and} \qquad \text{var}(\bar{X}|\omega) = \text{var}(X_1|\omega)/K = \sigma^2(\omega)/K,$$

and the rule of double variance yields

$$\text{var}(\bar{X}) = \text{var}\{E(\bar{X}|\omega)\} + E\{\text{var}(\bar{X}|\omega)\} = \text{var}\{\pi(\omega)\} + E\{\sigma^2(\omega)/K\} = \upsilon^2 + \frac{\tau^2}{K},$$

as asserted. Finally for the covariance

$$E\{(\bar{X}-\eta)(\pi(\omega)-\eta)|\omega\} = E\{\bar{X}-\eta\}\{\pi(\omega)-\eta\} = \{\pi(\omega)-\eta\}^2,$$

and by the rule of double expectation

$$E\{(\bar{X}-\eta)(\pi(\omega)-\eta)\} = E\{\pi(\omega)-\eta\}^2 = \upsilon^2,$$

and the term on the left is $\text{cov}\{\bar{X}, \pi(\omega)\}$. In the following we shall write $\pi = \pi(\omega)$.

Linear credibility Let $\hat{\pi}_K$ be the estimate in (1.10). Then

$$\hat{\pi}_K - \pi = b_0 + b\bar{X} - \pi = b_0 - (1-b\zeta) + b(\bar{X}-\zeta) - (\pi-\zeta)$$

after a little reorganization. Hence

$$\{\hat{\pi}_K - \pi\}^2 = \{b_0 - (1-b\zeta)\}^2 + b^2(\bar{X}-\zeta)^2 + (\pi-\zeta)^2$$
$$+2\{b_0-(1-b\zeta)\}(\bar{X}-\zeta) + 2\{b_0-(1-b\zeta)\}(\pi-\zeta) - 2b(\bar{X}-\zeta)(\pi-\zeta),$$

and $Q = E\{\hat{\pi}_K - \pi\}^2$ is calculated by taking expectation on both sides. Since $E(\bar{X}) = \zeta$ and $E(\pi) = \zeta$, this yields

$$Q = (b_0 - (1-b\zeta)^2 + b^2\text{var}(\bar{X}) + \text{var}(\pi) + 0 + 0 - 2b\text{cov}\{\bar{X}, \pi(\omega)\}$$

and after inserting (1.26) (middle and right) and $\upsilon^2 = \text{var}(\pi)$, we obtain

$$Q = (b_0 - (1-b)\zeta)^2 + b^2(\upsilon + \tau^2/K) + \upsilon^2 - 2b\upsilon^2$$

This is minimized by taking

$$b_0 = 1 - b\zeta \qquad \text{and} \qquad b = w = \frac{\upsilon^2}{\upsilon + \tau^2/K},$$

the solution of b_0 being obvious and that for b being found by differentiation afterwards. This yields the credibility estimate $\hat{\pi}_K$ defined in (1.11).

The statistical properties Unbiasedness is a consequence of

$$E(\hat{\pi}_K) = E\{(1-w)\zeta + w\bar{X}\} = (1-w)\zeta + wE(\bar{X}) = (1-w)\zeta + w\zeta = \zeta = E(\pi).$$

The variance of the error is calculated by inserting $b_0 = 1 - w\xi$ and $b = w$ in the expression for Q. This yields

$$Q = \left(\frac{\upsilon^2}{\upsilon^2 + \tau^2/K}\right)^2 (\upsilon^2 + \tau^2/K) + \upsilon^2 - 2\frac{\upsilon^2}{\upsilon^2 + \tau^2/K}\upsilon^2 = \frac{\upsilon^2\tau^2/K}{\upsilon^2 + \tau^2/K},$$

so that

$$E(\hat{\pi}_K - \pi)^2 = Q = \frac{\upsilon^2}{1 + K\upsilon^2/\tau^2}$$

as asserted in (1.12).

Optimal credibility We must determine the distribution of μ given $N_1, \ldots, N_K$. The prior density function assumed for μ is

$$f(\mu) = C\mu^{\alpha-1}e^{-\mu\alpha/\xi_\mu}$$

where C is a constant whereas the the claim numbers are conditionally independent and Poisson given μ. Their joint density function is

$$f(n_1, \ldots, n_K|\mu) = \prod_{k=1}^{K}\left(\frac{(\mu T)^{n_k}}{n_k!}e^{-\mu T}\right) = C\mu^{n_1+\ldots+n_K}e^{-\mu KT}$$

where C (another constant) is an expression not depending on μ. Multiplying the pair of density functions together yields the posterior density function $p(\mu|n_1, \ldots, n_K)$ up to a constant. In other words,

$$p(\mu|n_1, \ldots, n_K) = C\left(\mu^{\alpha-1}e^{-\mu\alpha/\xi_\mu}\right)\cdot\left(\mu^{n_1+\ldots+n_K}e^{-\mu KT}\right) = Ce^{\alpha+K\bar{n}-1}e^{-\mu(\alpha/\xi+KT)}$$

where $\bar{n} = (n_1 + \ldots + n_K)/K$. This is another Gamma density function with expectation

$$E(\mu|n_1, \ldots, n_K) = \frac{\alpha + K\bar{n}}{\alpha/\xi_\mu + KT} = \xi_\mu\frac{\bar{n} + \alpha/K}{\xi_\mu T + \alpha/K}.$$

Multiply with $T\xi_z$, and you get

$$\hat{\pi}_K = \zeta \frac{\bar{n} + \alpha/K}{\xi_\mu T + \alpha/K}.$$

which is the credibility estimate (1.14).

Optimal credibility: Error

Note that

$$\hat{\pi}_K - \pi = \zeta \frac{\bar{n} + \alpha/K}{\xi_\mu T + \alpha/K} - \mu T \xi_z = \zeta \left(\frac{\bar{n} + \alpha/K}{\xi_\mu T + \alpha/K} - \frac{\mu}{\xi_\mu} \right).$$

Since $E(\bar{n}|\mu) = \mu T$ and $\text{var}(\bar{n}|\mu) = \mu T/K$, this implies that

$$E(\hat{\pi}_K - \pi|\mu) = \zeta \left(\frac{\mu T + \alpha/K}{\xi_\mu T + \alpha/K} - \frac{\mu}{\xi_\mu} \right) \qquad \text{and} \qquad \text{var}(\hat{\pi}_K - \pi|\mu) = \zeta^2 \frac{\mu T/K}{(\xi_\mu T + \alpha/K)^2},$$

and by the rule of double variance

$$\text{var}(\hat{\pi}_K - \pi) = \zeta^2 \left(\frac{T}{\xi_\mu T + \alpha/K} - \frac{1}{\xi_\mu} \right)^2 \sigma_\mu^2 + \zeta^2 \frac{\xi_\mu T/K}{(\xi_\mu T + \alpha/K)^2}.$$

Under the model assumed $\sigma_\mu = \xi_\mu/\sqrt{\alpha}$, and when this is inserted, the preceding expression reduces to

$$\text{var}(\hat{\pi}_K - \pi) = \frac{\zeta^2}{\alpha + K\xi_\mu T}.$$

which is (1.15).

The estimates of ξ, τ and υ.

We shall examine the estimates (1.18) and (1.19). The principal part of the argument is to verify unbiasedness. Firstly, since $E(\bar{x}_j) = \zeta$ and $K_1 + \ldots + K_J = \mathcal{K}$ we have

$$E(\hat{\zeta}) = \sum_{j=1}^{J} \frac{K_j}{\mathcal{K}} E(\bar{x}_j) = \sum_{j=1}^{J} \frac{K_j}{\mathcal{K}} \zeta = \zeta.$$

For τ we must utilize that s_j^2 is the ordinary empirical variance. Thus

$$E(s_j^2|\omega) = \sigma^2(\omega),$$

and by the rule of double expectation

$$E(s_j^2) = E\{E(s_j^2|\omega)\} = E\{\sigma^2(\omega)\} = \tau^2.$$

Now (1.18) right yields

$$E(\hat{\tau}^2) = \sum_{j=1}^{J} \frac{K_j - 1}{\mathcal{K} - J} E(s_j^2) = \sum_{j=1}^{J} \frac{K_j - 1}{\mathcal{K} - J} \tau^2 = \tau^2.$$

Finally, note that

$$\hat{\zeta} - \zeta = \sum_{j=1}^{J} \frac{K_j}{\mathcal{K}} (\bar{x_j} - \zeta),$$

so that

$$Q_\upsilon = \sum_{j=1}^{J} \frac{K_j}{\mathcal{K}} (\bar{x_j} - \hat{\zeta})^2 = \sum_{j=1}^{J} \frac{K_j}{\mathcal{K}} (\bar{x_j} - \zeta)^2 - (\hat{\zeta} - \zeta)^2,$$

and from (1.26) middle

$$E(\bar{x}_j - \zeta)^2 = \upsilon^2 + \frac{\tau^2}{K_j}.$$

Since

$$E(\hat{\zeta} - \zeta)^2 = \text{var}(\hat{\zeta}) = \sum_{j=1}^{J} \left(\frac{K_j}{\mathcal{K}} \right)^2 \text{var}(x_j) = \sum_{j=1}^{J} \left(\frac{K_j}{\mathcal{K}} \right)^2 (\upsilon^2 + \frac{\tau^2}{K_j}),$$

it now follows that

$$E(Q_\upsilon) = \sum_{j=1}^{J} \frac{K_j}{\mathcal{K}} (\upsilon^2 + \frac{\tau^2}{K_j}) - \sum_{j=1}^{J} \left(\frac{K_j}{\mathcal{K}} \right)^2 (\upsilon^2 + \frac{\tau^2}{K_j})$$

or since $K_1 \ldots + K_J = \mathcal{K}$

$$E(Q_\upsilon) = \upsilon^2 + \frac{J}{\mathcal{K}} \tau^2 - \upsilon^2 \sum_{j=1}^{J} \left(\frac{K_j}{\mathcal{K}} \right)^2 - \frac{\tau^2}{\mathcal{K}}.$$

Thus

$$E(Q_\upsilon) = 1 - \sum_{j=1}^{J} \left(\frac{K_j}{\mathcal{K}} \right)^2 \upsilon^2 + \frac{J-1}{M} \tau^2,$$

and the the estimate $\hat{\upsilon}^2$ is determined by solving the equation

$$Q_\upsilon = \sum_{j=1}^{J} \frac{K_j}{\mathcal{K}} (\bar{x}_j - \hat{\eta})^2 = 1 - \sum_{j=1}^{J} \left(\frac{K_j}{\mathcal{K}} \right)^2 \hat{\upsilon}^2 + \frac{J-1}{\mathcal{K}} \hat{\tau}^2.$$

This yields the estimate (1.19) for $\hat{\upsilon}$, and the argument has also shown that $\hat{\upsilon}$ is unbiased.

1.8 Bibliographic Notes

The most versatile method for computing reserves is Monte Carlo which is except for the normal approximation also the fastest one to implement if you know how to program a computer. Normal power is arguably a method belonging to the pre-computer age (it only extends the area of application of the ordinary normal a little). The Panjer recursion (Section 3.8) and other recursive methods are reviewed in Chapter 4 of Dickson (2005). Those might have been important historically, but they are now superseded by the much more flexible and general Monte Carlo.

Credibility (Section 10.5) when applied to individuals is an ambitous attempt to attach risk to policy holders according to their own track record, and errors must be huge as we saw in Section 10.5. An obvious place to study the method further is Bühlman and Gisler (2005), but you are also likely to appreciate the elegantly written review article Norberg (2004) which draws methodological lines to certain parts of statistics. The simple solution for a Gamma prior with Poisson counts in Section 10.5 goes back to Jewel (1974) with later extensions in Kaas, Danneburg and Goovaerts (1997) and Ohlsson and Johansson (2006). Credibility is still (2008) an area for basic research, in particular with respect to time effects and the use of explanatory variables. Frees (2003), Huang Song and Liang (2003), Purcaru and Denuit (2003), Luo, Young and Frees (2004), Pitselis (2004), Frees and Wang (2006), Yeo and Valdez (2006) Lo, Wing and Zhu (2006) and Pitselis (2008) all present extensions in one of those directions.

The presentation of reinsurance in Section 10.6 owes something to Daykin, Pentikäinen and Pesonen (1994). Academic research is in this area largely confined to how such contracts should be designed, a topic that wasn't covered above. A lot effort has been put into strategies that work over time (references in Section 11.9). One-period contributions are Gajek and Zagrodny (2004), Kaluscka (2004) and (2005), Kravich and Sherris (2006), Lee and Yu (2007) and Cai and Tan (2007) with many others in the years prior to 2004.

Bühlman, H. and Straub, E. (1970). Glaubwüdigkeit für Schadebsätze. *Mitteleiungen der Vereinigung Scweizerischer Versicherungsmatematiker*, 70, 111-133.

Bühlman, H. and Gisler, A. (2005). *A first Course in Credibility Theory and its Applications.* Springer Verlag. Berlin Heidelberg.

Cai, J. and Tan, K.S. (2007). Optimal Retention for Stop-Loss Reinsurance under the VaR and CTE Risk Measures. *Astin Bulletin*, 37, 93-112.

Daykin, C.D., Pentikäinen, T. and Pesonen, M. (1994). *Practical Risk theory for Actuaries.* Chapman & Hall/CRC, London.

Dickson, D.C.M. (2005). *Insurance Risk and Ruin.* Cambridge University Press, Cambridge.

Feller, W. (1971). *An Introduction to Probability Theory and its Applications.* Volume II. John Wiley & Sons, New York.

Frees, E. (2003). Multivariate Credibility for Aggregate Loss Models. *North American Actuarial Journal*, 1, 13-37.

Frees, E. and Wang, P. (2006). Copula Credibility for Aggregate Loss Models. *Insurance: Mathematics and Economics*, 38, 360-373.

Gajek, L. and Zagrodny, D. (2004). Optimal Reinsurance under General Risk measures. *Insurance: Mathematics and Economics*, 34, 227-240.

Hall, P. *The Bootstrap and Edgeworth Expansison.* Springer Verlag, New York.

Huang, X, Song,L. and Liang, Y. (2003). Semiparametric Credibility Rate making Using a Piecewise Linear Predictor. *Insurance: Mathematics and Economics*, 33. 585-593.
Jewel, W. (1974). Credible Means are Exact Bayesian for Exponential Families. *Astin Bulletin*, 8, 77-90.
Kaas, R, Danneburg, D. and Goovaerts, M.J. (1997). Exact Credibility for Weighted Observations. *Astin Bulletin*, 27,287-295.
Kaluszka, M. (2004). Mean-Variance Optimal Reinsurance Arrangements. *Scandinavian Actuarial Journal*, 28-41.
Kaluszka, M. (2005). Truncated Stopp Loss as Optimal Reinsurance Agreement in One-period Models. *Astin Bulletin*, 35, 337-349.
Kravitch, Y. and Sherris, M. (2006), Enhancing Reinsurer Value through Optimization. *Insurance: Mathematics and Economics*, 38, 495-517.
Lee, J-P. and Yu, M-T. (2007). Valuation of Catastrophy Reinsurance with Catastrophy Bonds. *Insurance: Mathematics and Economics*, 41, 264-278.
Lo, C.h., Wing, K. and Zhu, Z.Y. (2006). Generalized Estimating Equations for Variance and Covariance Parameters in Regression Credibility Models *Insurance: Mathematics and Economics*, 39, 99-113.
Luo, Y, Young, V.R and Frees, E.W. (2004). Credibility Rate Making Using Collateral Information. *Scandinavian Actuarial Journal*, 448-461.
Nordberg, R. (2004). Credibility Theory. In *Encyclopedia of Actuarial Science*, Teugels, J, and Sundt, B. (eds), John Wiley & Sons, Chichester, 398-407.
Ohlsson, E. and Johansson, B. (2006). Exact Credibility and Tweedie Models, *Astin Bulletin*, 36, 121-133.
Pitselis, G. (2004). A Seemingly Unrelated Regression Model in a Credibility Framework. *Insurance: Mathematics and Economics*, 34, 37-54.
Pitselis, G. (2008). Robust Regression Credibility. the Influence Function Approach. *Insurance: Mathematics and Economics*, 42, 288-300.
Purcaru, O. and Denuit, M. (2003). Dependence in Dynamic Claim Credibility Models. *Astin Bulletin*, 33, 23-40.
Sokal, R.S. and Rholf, F.J., (1981), second ed. *Biometry*. W.H. Freeman and Company, San Francisco.
Sundt, B. (1983). Parameter Estimation in some Credibility Models. *Scandinavian Actuarial Journal*, 239-255
Yeo, K.L. and Valdez, M.A. (2006). Claim Dependence with Common Effects in Credibility Models. *Insurance: Mathematics and Economics*, 38, 609-623.

1.9 Exercises

Section 10.2

Exercise 10.2.1 Consider a portfolio of identical risks where the standard deviation of the claim size model is $\sigma_z = \theta\xi_z$ where θ is a parameter. **a)** Show that the normal approximation for the reserve can be written

$$q_\epsilon^{\mathrm{No}} = E(\mathcal{X})(1+\gamma) \qquad \text{where} \qquad \gamma = \sqrt{\frac{1+\theta^2}{\mu T J}}\,\phi_\epsilon.$$

Let Z be Gamma-distributed so that $\theta = 1/\sqrt{\alpha}$ where α is the shape parameter. **b)** Compute γ when $\mu = 5\%$, $T = 1$, $\alpha = 1$, $\epsilon = 1\%$ (so that $\phi_\epsilon = 2.33$) and $J = 100$ and $J = 10000$. For which of the two values of J is the approximation most reliable?

Exercise 10.2.2 This is an extension of the preceding exercise. **a)** Show that the normal power approximation can be expressed as

$$q_\epsilon^{\rm NP} = E(\mathcal{X})(1+\gamma) \qquad \text{where} \qquad \gamma = \sqrt{\frac{1+\theta^2}{\mu TJ}}\,\phi_\epsilon + \frac{(\zeta_z\theta^3+3\theta^2+1)(\phi_\epsilon^2-1)}{6(1+\theta^2)\mu TJ}$$

For a Gamma distribution $\theta = 1/\sqrt{\alpha}$ and $\zeta_z = 2/\sqrt{\alpha}$. **b)** Insert those into the preceding expression for γ. **c)** Investigate the impact of the NP-term on γ numerically under the same conditions as in Exercise 10.2.1b).

Section 10.3

Exercise 10.3.1 For a portfolio of identical risks suppose $Z = \xi_z G$ where $G \sim$ Gamma(α). **a)**. Compute the normal and normal power approximation to the reserve at level $\epsilon = 1\%$ when $\alpha = 1$ and $J\mu = 10$, 100 and 1000. **b)** Repeat the computations in a) by means of simulations using $m = 10000$. **c)** Compare the results in a) and b) and comment on how the discrepancies depend on $J\mu$.

Exercise 10.3.2 Suppose you want to plan a simulation experiment for the reserve so that Monte Carlo error is less than a certain fraction γ of the final result. One strategy is to run B batches of m_1 simulations. For each batch b sort the simulations in descending order and compute $q_{\epsilon b}^* = \mathcal{X}_{(m_1\epsilon)}^*$ as the reserve. That gives you B assessments $q_{\epsilon 1}^*, \dots, q_{\epsilon B}^*$ from which you may compute their mean $\bar{q}_\epsilon^*$ and standard deviation s_ϵ^*. **a)** Carry out $B = 10$ rounds of such experiments when $J\mu = 10$ and Z is exponentially distributed with mean one. Use $\epsilon = 1\%$ and $m_1 = 10000$. **b)** Compute mean and standard deviation $\bar{q}_\epsilon^*$ and s_ϵ^* of $q_{\epsilon 1}^*, \dots, q_{\epsilon B}^*$. For large m there is the approximation $\mathrm{sd}(q_\epsilon^*) \doteq a_\epsilon/\sqrt{m}$ where a_ϵ doesn't depend on m; see Section 2.2. **c)** Estimate a_ϵ as the value $a_\epsilon^* = s_\epsilon^*\sqrt{m_1}$ and argue that

$$m = m_1\left(\frac{s_\epsilon^*}{\gamma\bar{q}_\epsilon^*}\right)^2$$

is approximately the number of simulations you need. **d)** Compute it for the values you found in b) when $\gamma = 1\%$.

Section 10.4

Exercise 10.4.1 A quick way to explore statistical significance is the **Wald** test. Let $\hat\theta$ be an estimate of a parameter θ and $\hat\sigma_\theta$ its estimated standard deviation. Then pronounce the underlying θ different from zero if $|\hat\theta/\hat\sigma_\theta| > 2$. The significance level is close to 5% under the normal approximation which is a fair assumption in many applications of regression methodology. **a)** Apply this test to the second age category in Table 10.4. Can we from this information be sure that age has real impact on claim frequency and size? **b)** Examine the two other explanatory variables in Table 10.4 (distance limit and geographical region) in the same way. Which of the categories deviate significantly from the first one?

Exercise 10.4.2 The estimate of the pure premium for a customer with a given set of explanatory variables is $\hat\pi = Te^{\hat\eta}$ where mathematical expressions for $\hat\eta$ and its estimated standard deviation $\hat\tau$ were given in Section 10.4. Let $\underline{\pi} = \hat\pi e^{-2\hat\tau\phi_\epsilon}$ and $\bar\pi = \hat\pi e^{2\hat\tau\phi_\epsilon}$ where ϕ_ϵ is the $1-\epsilon$ percentile of the standard normal distribution. **a)** Argue using the normal approximation that $\underline{\pi} < \pi < \bar\pi$ is $1-2\epsilon$ confidence interval for π. **b)** Compute 95% condifence intervals for the pure premia in Table 10.5 utilizing that $\phi_\epsilon \doteq 2$ when $\epsilon = 2.5\%$.

Exercise 10.4.3 Consider a portfolio where regression models for claim intensity and claim size have been fitted. Then $\log(\mu_j) = b_{\mu 1}x_{j0} + \ldots + b_{\mu v}x_{jv}$ and $\log(\xi_{zj}) = b_{\xi 1}x_{j0} + \ldots + b_{\xi v}x_{jv}$ are known relationships for policy holder j. If $\sigma_{zj} = \sigma_z$ is the common standard deviation for all j, use the central limit theorem to compute the approximate reserve for the portfolio at level ϵ.

Section 10.5

Exercise 10.5.1 Consider the linear credibility estimate $\hat{\pi}_K = (1-w)\zeta + w\bar{X}_K$ where $w = \upsilon^2/(\upsilon^2+\tau^2/K)$ and $\zeta = E\{\pi(\omega)\}$, $\upsilon^2 = \text{var}\{\pi(\omega)\}$ and $\tau^2 = E\{\sigma^2(\omega)\}$ are the three structural parameters. Explain and interprete why the credibility set-up yields a weight w which is increasing in K and υ and decreasing in τ.

Exercise 10.5.2 The accuracy of the group credibility estimate $\hat{\Pi}_K$ may be examined by calculating relative error. **a)** Use the standard deviation formula (1.17) to verify that

$$\frac{\text{sd}(\hat{\Pi}_K - \Pi)}{E(\Pi)} = \frac{\upsilon/\zeta}{\sqrt{(1+KJ\upsilon^2/\tau^2}},$$

b) What happens to the ratio as $J \to \infty$? Explain what this tells us about the accuracy of credibility estimation on group level.

Exercise 10.5.3 Exercise 10.5.2 enables us to re-analyse the accuracy of credibility estimation. The following conditions are those of Table 10.7 with μ being a common, random factor influencing the entire portfolio. Then, $\zeta = \xi_\mu\xi_z$, $\upsilon^2 = \sigma_\mu^2\xi_z^2$ and $\tau^2 = \xi_\mu(\xi_z^2 + \sigma_z^2)$ if $T = 1$; see Section 10.5. Suppose $\mu = 5.6\%$, $\sigma_\mu = 2\%$, and $\sigma_z = 0.1\xi_z$ as in Table 10.7. **a)** Show that the standard deviation/mean ratio in Exrecise 10.5.2 doesn't depend on ξ_z. **b)** Compute it for $K = 0$, 10 and 20 both when $J = 1$ (single polices) and when $J = 10000$. **c)** What conclusions do you draw from these computations?

Exercise 10.5.4 The optimal credibility estimate $\hat{\pi}_K = \zeta(\bar{n} + \alpha/K)/(\xi_\mu T + \alpha/K)$ is an adjustment of the average, pure premium ζ of the population. Clearly $\hat{\pi}_K > \zeta$ if $\bar{n} > \xi_\mu T$ and $\hat{\pi}_K \leq \zeta$ in the opposite case. **a)** What is the intuition behind this? **b)** Show that the adjustment increases with K and decreases with α. Why must the credibility set-up lead to these results?

Exercise 10.5.5 Let $\Pi = E(\mathcal{X}|\mu)$ be the average claim against a portfolio when claim intensity μ is a common random factor with prior distribution $\mu = \xi G$ with $G \sim \text{Gamma}(\alpha)$. Suppose $\bar{n}$ is the average number of claims against the portfolio over K years. **a)** Explain why $\hat{\Pi}_K = J\zeta(\bar{n} + \alpha/K)/(J\xi_\mu T + \alpha/K)$ is the optimal credibility estimate of Π. **b)** Use the standard deviation formula (1.15) right to deduce that

$$\frac{\text{sd}(\hat{\Pi}_K - \Pi)}{E(\Pi)} = \frac{1}{\sqrt{\alpha + J\xi_\mu KT}}.$$

c) What is the limit as $J \to \infty$? Comment on the potential of optimal credibility estimation on group level.

Section 10.6

Exercise 10.6.1 Consider J independent portfolios where the total claim $\mathcal{X}_j$ against portfolio j is approximately normally distributed with mean ξ_j and standard deviation σ_j, $j = 1, \ldots, J$. A re-insurer has taken partial responsibility for all the portfolios through proportional re-insurance. This means that the re-insurer share for portfolio j is $\mathcal{X}_j^{\text{re}} = c_j\mathcal{X}_j$ where $0 < c_j < 1$ is a fixed coefficient. **a)** Explain why the total re-insurance obligation $\mathcal{Y}^{\text{re}} = \mathcal{X}_1^{\text{re}} + \ldots + \mathcal{X}_J^{\text{re}}$ is approximately normally distributed and identify the mean and standard deviation. **b)** The normal approximation is arguably better than for the portfolios individually. Why is that? **c)** Compute the 99% reserve for the re-insurer when $J = 20$ and all $\xi_j = 1$, $\sigma_j = 0.2$ and all $c_j = 1/3$.

Exercise 10.6.2 The situation is the same as in the preceding exercise except that the contracts are of the $a \times b$ type so that $\mathcal{X}_j^{\text{re}} = 0$ if $\mathcal{X}_j < a_j$, $\mathcal{X}_j^{\text{re}} = \mathcal{X}_j - a_j$ if $a_j \leq \mathcal{X}_j < a_j + b_j$ and $\mathcal{X}_j^{\text{re}} = b_j$ if $\mathcal{X}_j \geq a_j + b_j$. Suppose $J = 20$ and all $\xi_j = 1$ and $\sigma_j = 0.2$. **a)** Use Monte Carlo to compute the re-insurer 99% reserve when when all $a_j = 0.9$ and $b_j = 1.2$ **b)** Repeat a) when $a_j = 1.1$ and $b_j = 1.4$. **c)** Why is more money required in a)?

Exercise 10.6.3 Suppose claims against a portfolio is log-normally distributed of the form $Z = e^{-\tau^2/2+\tau\varepsilon}$ where $\varepsilon \sim N(0,1)$. Upper responsibility per claim is $b = 4$. The portfolio is re-insured through a $a \times b$ contract where b coincides with the maximum claim against the cedent. The average number of claims annually is 20. **a)** Use Monte Carlo to determine the pure premium of the re-insurance when $\tau = 1$ and $a = 1$, $a = 2$ and $a = 3$. **b)** Compute the 99% reserve of the cedent under the same conditions as in *a*), again using Monte Carlo. Suppose τ increases from $\tau = 1$ and that the other conditions are the same as before. **c)** Is the cedent reserve and re-insurance premium now going to go up or down? It is possible to answer this through intuition!