Nonlife Actuarial Models

Chapter 3

Aggregate-Loss Models

Learning Objectives

- Individual risk model
- Collective risk model
- De Pril recursion
- Compound process for collective risk
- Approximation methods
- Stop-loss reinsurance

3.1 Individual Risk and Collective Risk Models

(1) Individual risk model:

- The number of policies in the block is n. We assume the loss of each policy, denoted by X_i, for $i = 1, \cdots, n$, to be *independently and identically distributed* as X. The aggregate loss of the block of policies, denoted by S, is then given by

$$S = X_1 + \cdots + X_n. \tag{3.1}$$

- X follows a mixed distribution with a probability mass at point zero.

(2) Collective risk model:

- Let N be the number of losses in the block of policies, and X_i be the amount of the ith loss, for $i = 1, \cdots, N$.

- The aggregate loss S is given by

$$S = X_1 + \cdots + X_N. \tag{3.2}$$

- N is a random variable, and N and X are assumed to be independent.

- S has a compound distribution.

Remarks:

- There are advantages in modeling the claim frequency and claim severity separately, and then combine them to obtain the aggregate-loss distribution.

- For example, expansion of insurance business may have impacts on the claim frequency but not the claim severity.

- Cost control (or general cost increase) and innovation in technology may affect the claim severity with no effects on the claim frequency.

3.2 Individual Risk Model

- The basic equation of the individual risk model is

$$S = X_1 + \cdots + X_n. \tag{3.1}$$

- As n is fixed, the mean and variance of S are given by

$$\mathrm{E}(S) = n\,\mathrm{E}(X) \qquad \text{and} \qquad \mathrm{Var}(S) = n\,\mathrm{Var}(X). \tag{3.3}$$

- X is a mixed-type distribution with prob mass at point 0.

- Let the probability of a loss be θ and the probability of no loss be $1 - \theta$.

- When there is a loss, the loss amount is Y, which is a positive continuous random variable with mean μ_Y and variance σ_Y^2.

- Thus, $X = Y$ with probability θ, and $X = 0$ with probability $1 - \theta$. We can write X as

$$X = IY, \tag{3.4}$$

where I is a Bernoulli random variable distributed independently of Y, so that

$$I = \begin{cases} 0, & \text{with probability } 1 - \theta, \\ 1, & \text{with probability } \theta. \end{cases} \tag{3.5}$$

- Thus, the mean of X is

$$\mathrm{E}(X) = \mathrm{E}(I)\mathrm{E}(Y) = \theta\mu_Y, \tag{3.6}$$

- Using equation (A.118) in Appendix A.11, we have

$$
\begin{aligned}
\mathrm{Var}(X) &= \mathrm{Var}(IY) \\
&= [\mathrm{E}(Y)]^2\,\mathrm{Var}(I) + \mathrm{E}(I^2)\mathrm{Var}(Y) \\
&= \mu_Y^2\,\theta(1-\theta) + \theta\,\sigma_Y^2. \qquad (3.7)
\end{aligned}
$$

- Equations (3.6) and (3.7) can be plugged into equation (3.3) to obtain the mean and variance of S.

- **Example 3.1:** Assume there is a chance of 0.2 that there is a claim. When a claim occurs the loss is exponentially distributed with parameter $\lambda = 0.5$. Find the mean and variance of the claim distribution. Suppose there are 500 independent policies with this loss distribution, compute the mean and variance of their aggregate loss.

- **Solution:** The mean and variance of the loss in a loss event is

$$\mu_Y = \frac{1}{\lambda} = \frac{1}{0.5} = 2,$$

and

$$\sigma_Y^2 = \frac{1}{\lambda^2} = \frac{1}{0.5^2} = 4.$$

Thus, the mean and variance of the loss incurred by a random policy are

$$\mathrm{E}(X) = (0.2)(2) = 0.4,$$

and

$$\mathrm{Var}(X) = (2)^2(0.2)(1 - 0.2) + (0.2)(4) = 1.44.$$

The mean and variance of the aggregate loss are

$$\mathrm{E}(S) = (500)(0.4) = 200,$$

and

$$\mathrm{Var}(S) = (500)(1.44) = 720.$$

3.2.1 Exact distribution of S using convolution:

- We first consider the distribution of $X_1 + X_2$, where X_1 and X_2 are both continuous.

- The pdf of $X_1 + X_2$ is given by the 2-fold convolution

$$f^{*2}(x) = f_{X_1+X_2}(x) = \int_0^x f_1(x-y) f_2(y)\, dy = \int_0^x f_2(x-y) f_1(y)\, dy. \tag{3.8}$$

- The pdf of $X_1 + \cdots + X_n$ can be calculated recursively. Suppose the pdf of $X_1 + \cdots + X_{n-1}$ is given by the $(n-1)$-fold convolution $f^{*(n-1)}(x)$, then the pdf of $X_1 + \cdots + X_n$ is the n-fold convolution given by

$$
\begin{aligned}
f^{*n}(x) &= f_{X_1+\cdots+X_n}(x) = \int_0^x f^{*(n-1)}(x-y) f_n(y)\,dy \\
&= \int_0^x f_n(x-y) f^{*(n-1)}(y)\,dy.
\end{aligned} \tag{3.9}
$$

- Now we consider the case where X_i are mixed-type random variables.
- We assume that the pf-pdf of X_i is given by

$$
f_{X_i}(x) = \begin{cases} 1-\theta_i, & \text{for } x = 0, \\ \theta_i\, f_{Y_i}(x), & \text{for } x > 0, \end{cases} \tag{3.10}
$$

in which $f_{Y_i}(\cdot)$ are well defined pdf of positive continuous random variables.

- The df of $X_1 + X_2$ is given by the 2-fold convolution in the Stieltjesintegral form, i.e.,

$$\begin{aligned} F^{*2}(x) &= F_{X_1+X_2}(x) = \int_0^x F_{X_1}(x-y)\, dF_{X_2}(y) \\ &= \int_0^x F_{X_2}(x-y)\, dF_{X_1}(y). \end{aligned} \tag{3.11}$$

- The df of $X_1 + \cdots + X_n$ can be calculated recursively.

- **Example 3.2:** For the block of insurance policies defined in Example 3.1, approximate the loss distribution by a suitable discrete distribution. Compute the df $F_S(s)$ of the aggregate loss of the portfolio for s from 110 through 300 in steps of 10, based on the discretized distribution.

- **Solution:** We approximate the exponential loss distribution by a discrete distribution taking values $0, 1, \cdots, 10$. As the df of $\mathcal{E}(\lambda)$ is $F_X(x) = 1 - \exp(-\lambda x)$, we approximate the pf by

$$f_X(x) = (0.2)\left\{\exp\left[-\lambda(x - 0.5)\right] - \exp\left[-\lambda(x + 0.5)\right]\right\}, \qquad \text{for } x = 1, \cdots, 9,$$

with

$$f_X(0) = 0.8 + (0.2)\left\{1 - \exp\left[-0.5\lambda\right]\right\},$$

and

$$f_X(10) = (0.2)\exp\left[-9.5\lambda\right].$$

The discretized approximate pf of the loss is

Table 3.1: Discretized probabilities

x	$f_X(x)$	x	$f_X(x)$
0	0.8442	6	0.0050
1	0.0613	7	0.0031
2	0.0372	8	0.0019
3	0.0225	9	0.0011
4	0.0137	10	0.0017
5	0.0083		

- Using the convolution method, the df of the aggregate loss S for selected values of s is given in Table 3.2.

Table 3.2: The df of S by convolution

s	$F_S(s)$	s	$F_S(s)$
110	0.0001	210	0.7074
120	0.0008	220	0.8181
130	0.0035	230	0.8968
140	0.0121	240	0.9465
150	0.0345	250	0.9746
160	0.0810	260	0.9890
170	0.1613	270	0.9956
180	0.2772	280	0.9984
190	0.4194	290	0.9994
200	0.5697	300	0.9998

3.2.3 Approximation of the Individual Risk Model:

- When n is large, by virtue of the Central Limit Theorem, S is approximately normal.

- The (exact) mean and variance of S are given in equations (3.3), (3.6) and (3.7), which can be used to compute the approximate distribution of S.

- Thus,

$$\begin{aligned}
\Pr(S \leq s) &= \Pr\left(\frac{S - \mathrm{E}(S)}{\sqrt{\mathrm{Var}(S)}} \leq \frac{s - \mathrm{E}(S)}{\sqrt{\mathrm{Var}(S)}}\right) \\
&\simeq \Pr\left(Z \leq \frac{s - \mathrm{E}(S)}{\sqrt{\mathrm{Var}(S)}}\right) \\
&= \Phi\left(\frac{s - \mathrm{E}(S)}{\sqrt{\mathrm{Var}(S)}}\right). \qquad (3.21)
\end{aligned}$$

3.3 Collective Risk Model

- The aggregate loss S is given by

$$S = X_1 + \cdots + X_N. \tag{3.2}$$

- N is the primary distribution, and X is the secondary distribution.

- S satisfies properties for compound distributions.

- The mgf $M_S(t)$ of the aggregate loss S is given by

$$M_S(t) = M_N\left[\log M_X(t)\right]. \tag{3.24}$$

- If the claim-severity takes nonnegative discrete values, S is also nonnegative and discrete, and its pgf is

$$P_S(t) = P_N\left[P_X(t)\right]. \tag{3.25}$$

- The mean and variance of S are

$$\mathrm{E}(S) = \mathrm{E}(N)\mathrm{E}(X), \tag{3.26}$$

and

$$\mathrm{Var}(S) = \mathrm{E}(N)\mathrm{Var}(X) + \mathrm{Var}(N)\left[\mathrm{E}(X)\right]^2. \tag{3.27}$$

These results hold whether X is continuous or discrete.

- If S_i has a compound Poisson distribution with claim-severity distribution X_i, which may be continuous or discrete, for $i = 1, \cdots, n$, then $S = S_1 + \cdots + S_n$ has also a compound Poisson distribution.

- When X is continuous, the pdf of S is

$$\begin{aligned} f_S(s) &= \sum_{n=1}^{\infty} f_{X_1+\cdots+X_n \,|\, n}(s)\, f_N(n) \\ &= \sum_{n=1}^{\infty} f^{*n}(s)\, f_N(n), \qquad (3.28) \end{aligned}$$

where $f^{*n}(\cdot)$ is the n-fold convolution. Thus, the exact pdf of S is a weighted sum of convolutions, and the computation is highly complex.

- There are some special cases for which the compound distribution can be analytically derived, such as

- **Theorem 3.2:** For the compound distribution specified in equation (3.2), assume $X_1, \cdots, X_N$ are iid $\mathcal{E}(\lambda)$, and $N \sim \mathcal{GM}(\theta)$. Then the compound distribution S is a mixed distribution with a probability mass of θ at 0 and a continuous component of $\mathcal{E}(\lambda\theta)$ weighted by $1 - \theta$.

- **Proof:** The mgf of $X \sim \mathcal{E}(\lambda)$ is

$$M_X(t) = \frac{\lambda}{\lambda - t},$$

and the mgf of $N \sim \mathcal{GM}(\theta)$ is

$$M_N(t) = \frac{\theta}{1 - (1 - \theta)e^t}.$$

Thus, we conclude that the mgf of S is

$$
\begin{aligned}
M_S(t) &= M_N\left[\log M_X(t)\right] \\
&= \frac{\theta}{1-(1-\theta)\left(\dfrac{\lambda}{\lambda-t}\right)} \\
&= \frac{\theta(\lambda-t)}{\lambda\theta-t} \\
&= \frac{\theta(\lambda\theta-t)+(1-\theta)\lambda\theta}{\lambda\theta-t} \\
&= \theta+(1-\theta)\left(\frac{\lambda\theta}{\lambda\theta-t}\right).
\end{aligned}
$$

3.3.2 Panjer recursion

- Theorem 1.5 provides the efficient Panjer recursion method to compute the exact distribution of a compound process which satisfies the conditions that (a) the primary distribution belongs to the $(a, b, 0)$ class, and (b) the secondary distribution is discrete and nonnegative.

- Thus, if a continuous claim-severity distribution can be suitably discretized, and the primary distribution belongs to the $(a, b, 0)$ class, we can use the Panjer approximation to compute the distribution of the aggregate loss.

3.3.3 Approximations of the Collective Risk Model

- If the mean number of claims is large, we may expect the normal approximation to work.

- Thus, using the mean and variance formulas of S in equations (3.26) and (3.27), we may approximate the df of S by

$$\Pr(S \leq s) = \Pr\left(\frac{S - \mathrm{E}(S)}{\sqrt{\mathrm{Var}(S)}} \leq \frac{s - \mathrm{E}(S)}{\sqrt{\mathrm{Var}(S)}}\right) \simeq \Phi\left(\frac{s - \mathrm{E}(S)}{\sqrt{\mathrm{Var}(S)}}\right). \tag{3.29}$$

- **Example 3.6:** Assume the aggregate loss S in a collective risk model has a primary distribution of $\mathcal{PN}(100)$ and a secondary distribution of $\mathcal{E}(0.5)$. Approximate the distribution of S using the normal distribution. Compute the df of the aggregate loss for $s = 180$ and 230.

- **Solution:** From equation (3.26) we have

$$\mathrm{E}(S) = \mathrm{E}(N)\mathrm{E}(X) = (100)\left(\frac{1}{0.5}\right) = 200,$$

and

$$\begin{aligned} \mathrm{Var}(S) &= \mathrm{E}(N)\mathrm{Var}(X) + \mathrm{Var}(N)\left[\mathrm{E}(X)\right]^2 \\ &= (100)\left(\frac{1}{0.5^2}\right) + (100)\left(\frac{1}{0.5}\right)^2 \\ &= 800. \end{aligned}$$

Thus, we approximate the distribution of S by $\mathcal{N}(200, 800)$. Using the normal approximation the required probabilities are

$$\Pr(S \leq 180) \simeq \Pr\left(Z \leq \frac{180.5 - 200}{\sqrt{800}}\right) = \Phi(-0.6894) = 0.2453,$$

and

$$\Pr(S \leq 230) \simeq \Pr\left(Z \leq \frac{230.5 - 200}{\sqrt{800}}\right) = \Phi(1.0783) = 0.8596.$$

3.4 Coverage Modifications and Stop-Loss Reinsurance

- We study the effects of coverage modifications on aggregate loss through their effects on the claim frequency and severity.

- We first consider the effects of a deductible of amount d. For the individual risk model, the number of policies n remains unchanged, while the policy loss X_i becomes the loss amount of the claim after the deductible, which we shall denote by $\tilde{X}_i$. Thus, the pf-pdf of $\tilde{X}_i$ is

$$f_{\tilde{X}_i}(x) = \begin{cases} 1 - \theta_i + \theta_i F_{Y_i}(d), & \text{for } x = 0, \\ \theta_i f_{Y_i}(x + d), & \text{for } x > 0. \end{cases} \tag{3.36}$$

- For the collective risk model the primary distribution of the compound distribution is now modified. Also, the secondary distribution is that of the claim after the deductible, i.e., $\tilde{X}$ with pdf given by

$$f_{\tilde{X}}(x) = \frac{f_X(x+d)}{1 - F_X(d)}, \quad \text{for } x > 0. \tag{3.37}$$

- Second, we consider the effects of a policy limit u. For the individual risk model, the number of policies again remains unchanged, while the claim-severity distribution is now capped at u. If we denote the modified claim-severity distribution by $\tilde{X}_i$, then the pf-pdf of $\tilde{X}_i$ is given by

$$f_{\tilde{X}_i}(x) = \begin{cases} 1 - \theta_i, & \text{for } x = 0, \\ \theta_i f_{Y_i}(x), & \text{for } 0 < x < u, \\ \theta_i \left[1 - F_{Y_i}(u)\right], & \text{for } x = u, \\ 0, & \text{otherwise.} \end{cases} \tag{3.38}$$

- For the collective risk model, the primary distribution is not affected, while the secondary distribution $\tilde{X}$ has a pf-pdf given by

$$f_{\tilde{X}}(x) = \begin{cases} f_X(x), & \text{for } 0 < x < u, \\ 1 - F_X(u), & \text{for } x = u, \\ 0, & \text{otherwise.} \end{cases} \tag{3.39}$$

- Insurance companies may purchase reinsurance coverage for a portfolio of policies they own. The coverage may protect the insurer from aggregate loss S exceeding an amount d, called **stop-loss reinsurance**. From the reinsurer's point of view this is a policy with a deductible of amount d. Thus, the loss to the reinsurer is $(S-d)_+$.

- $$\mathrm{E}\left[(S-d)_+\right] = \int_d^\infty \left[1 - F_S(s)\right] ds. \tag{3.40}$$

  This can be computed as

  $$\mathrm{E}\left[(S-d)_+\right] = \int_d^\infty (s-d) f_S(s)\, ds \tag{3.41}$$

when S is continuous, or

$$\mathrm{E}\left[(S-d)_{+}\right]=\sum_{s>d}(s-d) f_{S}(s) \tag{3.42}$$

when S is discrete.

Table 3.7: Methods for computing the aggregate-loss distribution

Model	Exact methods	Approximate methods
Individual risk	1) Convolution: with discretized claim-severity distribution 2) De Pril recursion: with specific set-up of policy stratification	1) Normal approximation 2) Compound Poisson distribution and Panjer recursion
Collective risk	1) Convolution: with discretized claim-severity distribution and assumed primary distribution 2) Panjer recursion: Primary distribution follows $(a, b, 0)$ class, secondary distribution discretized 3) Some limited analytic results	1) Normal approximation