Chapter 8

Risk process

- Individual model
- Collective model
- Comparison of the two models

Generally, the risk process is the stochastic process (S_t) representing the cumulative claim amounts, for the company, of a portfolio of (same kind of) contracts up to time t.

In this chapter, we will only need the r.v. S corresponding to the time interval $[0; 1]$. In the next chapter, we will use the generalization (S_t).

Individual model

- Definition and hypotheses
- Probability distribution of the risk process
  - Cumulative distribution function
  - Moment generating function
  - Moments
- Particular case : degenerated claim amounts
- Example

Definition and hypotheses

We consider here the portfolio as the sum of n independent risks (= contracts) $Y_1, \dots, Y_n$

For each risk (the j-th), the claim amount is a r.v. and we define

$$p_j = \Pr[0 \text{ claim}]$$
$$q_j = 1 - p_j = \Pr[1 \text{ claim}]$$

$$I_j \sim \begin{pmatrix} 0 & 1 \\ p_j & q_j \end{pmatrix}$$

Y_j = claim amount, conditionally to $[I_j = 1]$

Then,

$$S = I_1 Y_1 + \cdots + I_n Y_n$$

with the hypothesis of independence of

$$I_1, \dots, I_n, Y_1, \dots, Y_n$$

Probability distribution of the risk process

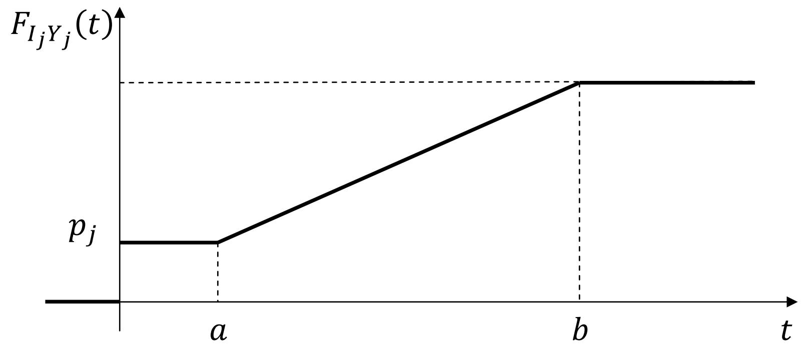
Cumulative distribution function

$$F_S(t) = \left(F_{I_1Y_1} * \ldots * F_{I_nY_n}\right)(t)$$

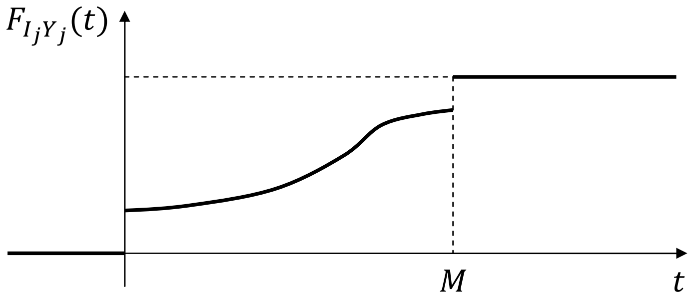
For $t \geq 0$,

$$\begin{aligned} F_{I_jY_j}(t) &= \Pr\left[I_jY_j \leq t\right] \\ &= \Pr\left(\left[I_jY_j \leq t\right] \middle| \left[I_j = 0\right]\right) \cdot \Pr\left[I_j = 0\right] \\ &\quad + \Pr\left(\left[I_jY_j \leq t\right] \middle| \left[I_j = 1\right]\right) \cdot \Pr\left[I_j = 1\right] \\ &= p_j + q_j F_{Y_j}(t) \end{aligned}$$

Example : if $Y_j \sim \mathcal{U}(a; b)$, then



Example of shape of the c.d.f. of I_jY_j



where

- the height of the jump at 0 is the probability that no claim occur
- M is the maximum intervention of the company

Moment generating function

$$\begin{aligned} m_S(t) &= \prod_{j=1}^{n} m_{I_j Y_j}(t) \\ &= \prod_{j=1}^{n} E\left(e^{tI_j Y_j}\right) \\ &= \prod_{j=1}^{n} E\left(E\left(e^{tI_j Y_j} \middle| I_j\right)\right) \end{aligned}$$

But, for an ω such that

- $I_j(\omega) = 0,$ then $E\left(e^{tI_j Y_j} \middle| I_j\right) = 1$
- $I_j(\omega) = 1,$ then $E\left(e^{tI_j Y_j} \middle| I_j\right) = m_{Y_j}(t)$

$$m_S(t) = \prod_{j=1}^{n} \left(p_j + q_j m_{Y_j}(t)\right)$$

Moments

$$E(S) = \sum_{j=1}^{n} E(I_j Y_j) = \sum_{j=1}^{n} q_j E(Y_j)$$

$$\begin{aligned} &var(S) \\ &= \sum_{j=1}^{n} var(I_j Y_j) \\ &= \sum_{j=1}^{n} \left\{E\left(var\left(I_j Y_j \middle| I_j\right)\right) + var\left(E\left(I_j Y_j \middle| I_j\right)\right)\right\} \\ &= \sum_{j=1}^{n} \left\{E\left(I_j^2 \cdot var\left(Y_j \middle| I_j\right)\right) + var\left(I_j \cdot E\left(Y_j \middle| I_j\right)\right)\right\} \\ &= \sum_{j=1}^{n} \left\{E\left(I_j^2\right) var\left(Y_j\right) + var\left(I_j\right) E^2\left(Y_j\right)\right\} \\ &= \sum_{j=1}^{n} \left\{q_j var\left(Y_j\right) + p_j q_j E^2\left(Y_j\right)\right\} \end{aligned}$$

$$var(S) = \sum_{j=1}^{n} q_j \left\{var\left(Y_j\right) + p_j E^2\left(Y_j\right)\right\}$$

Particular case : degenerated claim amounts

If $Y_j \equiv y_j$ for any j,

• c.d.f. : $F_{Y_j}(t) = \begin{cases} 0 & \text{if} \quad t < y_j \\ 1 & \text{if} \quad t \geq y_j \end{cases}$

Note : even in this particular case, convolutions are not easy to calculate : if $n = 2$ with $y_1 < y_2$,

$$S \sim \begin{pmatrix} 0 & y_1 & y_2 & y_1 + y_2 \\ p_1 p_2 & q_1 p_2 & p_1 q_2 & q_1 q_2 \end{pmatrix}$$

(Do it for $n = 5$)

• m.g.f. : $m_S(t) = \prod_{j=1}^{n} \left(p_j + q_j e^{t y_j}\right)$

• moments

$$E(S) = \sum_{j=1}^{n} q_j y_j \qquad var(S) = \sum_{j=1}^{n} p_j q_j y_j^2$$

Example

A portfolio of 14 risks, with degenerated claim amounts, sorted in 3 categories (3 different values for q_j)

N°	cat.	cl amount	q_j	p_j	$q_j y_j$	$p_j q_j y_j^2$
1	1	100	0,05	0,95	5	475
2	1	200	0,05	0,95	10	1900
3	1	200	0,05	0,95	10	1900
4	1	200	0,05	0,95	10	1900
5	1	300	0,05	0,95	15	4275
6	1	300	0,05	0,95	15	4275
7	2	300	0,10	0,90	30	8100
8	2	400	0,10	0,90	40	14400
9	2	400	0,10	0,90	40	14400
10	2	400	0,10	0,90	40	14400
11	3	200	0,15	0,85	30	5100
12	3	300	0,15	0,85	45	11475
13	3	300	0,15	0,85	45	11475
14	3	400	0,15	0,85	60	20400
			1,30		395	114475

It is not easy to obtain c.d.f. or m.g.f (possible values of S : $\{0, 100, 200, \ldots, 4000\}$)

$$E(S) = 395 \qquad var(S) = 114\,475$$

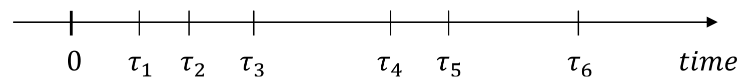
Collective model

- Definition and hypotheses
- Probability distribution of the risk process
  - Cumulative distribution function
  - Moment generating function
  - Moments
- Panjer's recursion formula
- Other distributions ?
  - In the original counting process
  - Thanks to Panjer's formula

Definition and hypotheses

The claims are no more generated individually : they are generated by the portfolio during the time interval

Occurrences of claim is a counting process



N (N_t in the general case) = number of claims during 1 year

X_j = claim amount for the claim occurring at τ_j

Then,

$$S = X_1 + \cdots + X_N$$

where

- $X_1, X_2, \ldots$ are i.i.d. r.v.
- N is independent of the X_k's

Such a process is a compound counting process

- N : claim frequency
- X : severity of the claim

Various conditions can affect these factors in different ways :

- seat belt : affects more X than N
- daytime headlights : affects more N than X

So, in practice,

- study the two components separately
- put them together in the model

Additional hypothesis : the counting process is a Poisson process with parameter λ

The risk process is then a compound Poisson process

Probability distribution of the risk process

Cumulative distribution function

$$F_S(t) = \sum_{k=0}^{\infty} F_X^{*k}(t) \cdot \Pr[N = k]$$
$$= \sum_{k=0}^{\infty} F_X^{*k}(t) \cdot e^{-\lambda} \frac{\lambda^k}{k!}$$

Moment generating function

$$m_S(t) = m_N(\ln m_X(t))$$
$$= e^{\lambda(e^{\ln m_X(t)}-1)}$$
$$= e^{\lambda(m_X(t)-1)}$$

Moments

$$E(S) = E(N) \cdot E(X) = \lambda E(X)$$

$$var(S) = E(N) \cdot var(X) + var(N) \cdot E^2(X)$$
$$= \lambda E(X^2)$$

Panjer's recursion formula

The probability distribution of S is again not easy to calculate, because of convolutions

Solution for the case where the possible values of the X_k's are positive integers : Panjer's formula

$$\Pr[S=0] = \exp\{-\lambda(1-\Pr[X=0])\}$$

and, for $s \in \mathbb{N}_0$,

$$\Pr[S=s] = \frac{\lambda}{s}\sum_{j=1}^{s} j \cdot \Pr[X=j] \cdot \Pr[S=s-j]$$

Proof for $s=0$

$$\begin{aligned}
\Pr[S=0] &= \sum_{k=0}^{\infty} \Pr([S=0] \mid [N=k]) \Pr[N=k] \\
&= \sum_{k=0}^{\infty} \Pr([X_1+\cdots+X_k=0] \mid [N=k]) \\
&\quad \cdot \Pr[N=k] \\
&= \sum_{k=0}^{\infty} \Pr[X_1+\cdots+X_k=0] \Pr[N=k] \\
&= \textstyle\sum_{k=0}^{\infty} (\Pr[X=0])^k \, e^{-\lambda} \frac{\lambda^k}{k!} \\
&= \exp\{-\lambda(1-\Pr[X=0])\}
\end{aligned}$$

Proof for $s \geq 1$

$$\Pr[S = s]$$
$$= \sum_{k=0}^{\infty} \Pr([S = s] | [N = k]) \Pr[N = k]$$
$$= \sum_{k=1}^{\infty} \Pr([S = s] | [N = k]) \Pr[N = k]$$
$$= \sum_{k=1}^{\infty} \Pr([X_1 + \cdots + X_k = s] | [N = k])$$
$$\cdot \Pr[N = k]$$
$$= \sum_{k=1}^{\infty} \Pr[X_1 + \cdots + X_k = s] \Pr[N = k]$$

But, for $i \in \{1, \ldots, k\}$

$$E(X_i | [X_1 + \cdots + X_k = s])$$
$$= s/k$$
$$= \sum_{j=1}^{s} j \cdot \Pr\big([X_1 = j] \big| [X_1 + \cdots + X_k = s]\big)$$
$$= \sum_{j=1}^{s} j \cdot \frac{\Pr([X_1 = j] \cap [X_2 + \cdots + X_k = s - j])}{\Pr[X_1 + \cdots + X_k = s]}$$

and

$$\Pr[X_1 + \cdots + X_k = s]$$
$$= \frac{k}{s} \sum_{j=1}^{s} j \cdot \Pr[X = j] \cdot \Pr[X_1 + \cdots + X_{k-1} = s - j]$$

Then,

$$\Pr[S = s]$$
$$= \sum_{k=1}^{\infty} \frac{k}{s} \sum_{j=1}^{s} j \cdot \Pr[X = j]$$
$$\cdot \Pr[X_1 + \cdots + X_{k-1} = s - j] \cdot \Pr[N = k]$$
$$= \frac{1}{s} \sum_{j=1}^{s} j \cdot \Pr[X = j]$$
$$\cdot \sum_{k=1}^{\infty} \Pr[X_1 + \cdots + X_{k-1} = s - j]\, e^{-\lambda} \frac{\lambda^k}{(k-1)!}$$

$$= \frac{\lambda}{s} \sum_{j=1}^{s} j \cdot \Pr[X = j]$$

$$\cdot \sum_{k=1}^{\infty} \Pr[X_1 + \cdots + X_{k-1} = s - j]$$

$$\cdot \Pr[N = k - 1]$$

$$= \frac{\lambda}{s} \sum_{j=1}^{s} j \cdot \Pr[X = j]$$

$$\sum_{k=0}^{\infty} \Pr[X_1 + \cdots + X_k = s - j] \cdot \Pr[N = k]$$

$$= \frac{\lambda}{s} \sum_{j=1}^{s} j \cdot \Pr[X = j] \cdot \Pr[S = s - j]$$

Other probability distributions ?

In the original counting process

The Poisson distribution is characterized by the fact that

$$E(N) = var(N) \quad (= \lambda)$$

Sometimes, because of the heterogeneity of risks, we have $var(N) > E(N)$

It is then possible to use mixed Poisson r.v.

- Λ is a positive r.v.
- Conditionally to $[\Lambda = \lambda]$, $N \sim \mathcal{P}(\lambda)$

• Probability distribution : for $k \in \mathbb{N}$,

$$\Pr[N = k] = E(\Pr([N = k] | \Lambda))$$

$$= \int_0^{+\infty} \Pr([N = k] | [\Lambda = \lambda]) \; dF_\Lambda(\lambda)$$

$$= \int_0^{+\infty} e^{-\lambda} \frac{\lambda^k}{k!} \; dF_\Lambda(\lambda)$$

- Moment generating function

$$\begin{aligned} m_N(t) &= E(e^{tN}) \\ &= E\big(E(e^{tN}|\Lambda)\big) \\ &= E\big(e^{\Lambda(e^t-1)}\big) \\ &= m_\Lambda(e^t-1) \end{aligned}$$

- Moments

$$\begin{aligned} E(N) &= E\big(E(N|\Lambda)\big) \\ &= E(\Lambda) \end{aligned}$$

$$\begin{aligned} var(N) &= E\big(var(N|\Lambda)\big) + var\big(E(N|\Lambda)\big) \\ &= E(\Lambda) + \mathrm{var}(\Lambda) \end{aligned}$$

Note 1 : $var(N) > E(\Lambda) = E(N)$

Note 2

- If Λ degenerated ($\Lambda \equiv \lambda$), then $N \sim \mathcal{P}(\lambda)$
- If $\Lambda \sim$ gamma, then $N \sim$ negative binomial
- If $\Lambda \sim$ exponential, then $N \sim$ geometric

Thanks to Panjer's formula

The Poisson distribution is such that

$$\Pr[N=k] = e^{-\lambda}\frac{\lambda^k}{k!} = \frac{\lambda}{k}\Pr[N=k-1]$$

In fact, Panjer's formula may be generalized for any distribution such that

$$\Pr[N=k] = \left(a + \frac{b}{k}\right)\cdot \Pr[N=k-1]$$

$$\Pr[S=0] = \begin{cases} \Pr[N=0] & if \quad \Pr[X=0]=0 \\ m_N(\ln(\Pr[X=0])) & if \quad \Pr[X=0]>0 \end{cases}$$

and, for $s \in \mathbb{N}_0$,

$$\Pr[S=s] = \frac{1}{1-a\cdot\Pr[X=0]} \cdot \sum_{j=1}^{s}\left(a+\frac{bj}{s}\right)\cdot\Pr[X=j]\cdot\Pr[S=s-j]$$

- If $a = 0$ and $b = \lambda (> 0)$, then $N \sim \mathcal{P}(\lambda)$

- If $a = 1 - p$ and $b = (1 - p)(r - 1)$ with $0 < p < 1$ and $r \in \mathbb{N}_0$, then $N \sim \mathcal{NB}(r; p)$

- If $a = \frac{-p}{1-p}$ and $b = \frac{p(n+1)}{1-p}$ with $0 < p < 1$ and $n \in \mathbb{N}_0$, then $N \sim \mathcal{B}(n; p)$

Comparison of the two models

- Quality of modeling
- Comparison of the two probability laws
    - Principle
    - Expected values
    - Variances
- Example

Quality of modeling

- Individual model is concerned by situations where only 0 or 1 claim occur during the time interval
    - No problem for life insurance
    - For non-life insurance, the model is wrong, except if the claim probability is such that $\Pr[\geq 2 \text{ claims for one risk}]$ is negligible

- Collective model is adaptable to much more situations

- The collective model is easier to handel, thanks to Panjer's formula

Comparison of the two probability laws

Principle

We identify the two first moments of

- X for the collective model
- the weighted average of the Y_j's for the individual model

$$E\left(X^k\right) = E\left(\sum_{j=1}^{n} \frac{q_j}{\sum q_i} Y_j^k\right) \qquad (k = 1, 2)$$

Expected values

$$E\left(S^{coll}\right) = \lambda E(X)$$
$$= \lambda E\left(\sum_j \frac{q_j}{\sum q_i} Y_j\right)$$
$$= \frac{\lambda}{\sum q_i} \sum_j q_j\, E(Y_j)$$
$$= \frac{\lambda}{\sum q_i} E\left(S^{ind}\right)$$

The two models are linked by the relation

$$\lambda = \sum_{j=1}^{n} q_j$$

Variances

$$var\left(S^{coll}\right) = \lambda E(X^2)$$
$$= \lambda E\left(\sum_j \frac{q_j}{\sum q_i} Y_j^2\right)$$
$$= \sum_j q_j E\left(Y_j^2\right)$$
$$= \sum_j q_j \left\{var\left(Y_j\right) + E^2(Y_j)\right\}$$
$$> \sum_j q_j \left\{var\left(Y_j\right) + p_j\, E^2(Y_j)\right\}$$
$$= var\left(S^{ind}\right)$$

The collective model is more “careful” because of its greater dispersion

Example

With the same data (portfolio with 14 risks)

N°	cat.	cl amount	q_j	p_j	$q_j\, y_j$	$p_j\, q_j\, y_j^2$
1	1	100	0,05	0,95	5	475
2	1	200	0,05	0,95	10	1900
3	1	200	0,05	0,95	10	1900
4	1	200	0,05	0,95	10	1900
5	1	300	0,05	0,95	15	4275
6	1	300	0,05	0,95	15	4275
7	2	300	0,10	0,90	30	8100
8	2	400	0,10	0,90	40	14400
9	2	400	0,10	0,90	40	14400
10	2	400	0,10	0,90	40	14400
11	3	200	0,15	0,85	30	5100
12	3	300	0,15	0,85	45	11475
13	3	300	0,15	0,85	45	11475
14	3	400	0,15	0,85	60	20400
			1,30		395	114475

We have $\lambda = \sum q_i = 1{,}3$

$$X \sim \begin{pmatrix} 100 & 200 & 300 & 400 \\ 0{,}0385 & 0{,}2308 & 0{,}3846 & 0{,}3462 \end{pmatrix}$$

e.g.

$$\Pr[X = 200] = \frac{3 \cdot 0{,}05 + 0{,}15}{1{,}3} = 0{,}2308$$

By Panjer's formula,

s	Pr[S = s]
0	0,27253
100	0,01363
200	0,08210
300	0,14036
400	0,14182
500	0,04780
600	0,07430
700	0,07111
800	0,04689
900	0,02694
1000	0,02699
1100	0,01962
1200	0,01183
1300	0,00780
1400	0,00611
1500	0,00387
1600	0,00230
1700	0,00151
1800	0,00101
1900	0,00060
2000	0,00035
2100	0,00022
2200	0,00013
2300	0,00008
2400	0,00004
2500	0,00003
2600	0,00001
2700	0,00001
2800	0,00000
2900	0,00000
3000	0,00000

Distribution du coût cumulé des sinistres