第2讲　聚合风险模型
——复合Poisson模型

No.2 2019.9.11

本讲主要内容

- 聚合模型简介
  - 模型
  - 背景
  - 个体与聚合
  - 一般性质
  - 例子
- Poisson聚合模型
  - 分布递推计算
  - 复合Poisson随机变量的独立和与分解

聚合模型（Collective ）—数学表示

$$S = \begin{cases} X_1 + X_2 + \cdots + X_N, & N > 0, \\ 0, & N = 0, \end{cases}$$

$$p_n = \Pr(N = n), \quad n = 0, 1, 2, \ldots,$$

$$F_X(x) = \Pr(X_i \le x), x \ge 0, \quad F_X(0) = 0$$

独立性条件

聚合模型—实际背景

- 开放的“组合”（portfolio）系统。
- 在保险中，并不会明确对应某个固定的保单组，而是在一定时期内的总损失。例如，某业务类一年的总损失，显然这个业务类一年内既有新的保单随时签发，又有保单会到期，这时可以认为系统是开放的。

聚合模型—理论背景

- 将损失的**发生次数**（频率）与每次损失发生后的**损失量**区别对待
  - 索赔数变量（claim number）
  - 索赔额变量（claim amount 或severity）
- 背景条件：风险规模（profile）或者风险敞口是隐形的但是确定的，例如：虽然不断有新业务发生和到期，但是沉淀的规模相对稳定

聚合模型-实际背景：信用风险

- 信用事件（违约default）的发生（PD）
  - 两点分布
  - 事件的频率
- 信用事件发生后违约的损失程度（LGD）
  - 连续的
  - 贝塔分布
  - 指数分布
  - 伽马分布

1. 索赔数变量(频率–frequency)分析

- 计数随机变量
  - 取值非负整数
  - 描述了事件的发生数
  - 数据和模型简单易得
- 常见的计数分布（ Counting distribution）
  - Poisson（单参数）
  - 负二项（两参数）
  - 几何分布
  - 二项（有限）

Poisson

$$p_k = \frac{\lambda^k}{k!} e^{-\lambda}, \; k = 0, 1, 2, \ldots$$

$$E[X] = Var[X] = \lambda$$

$$M_X(t) = e^{\lambda(e^t - 1)}, P_X(t) = e^{\lambda(t-1)}$$

负二项（两参数）

$$NB(r,\beta),\ NB(r,p);\ p=\frac{1}{1+\beta}$$

$$p_k=\binom{k+r-1}{k}\left(\frac{1}{1+\beta}\right)^r\left(\frac{\beta}{1+\beta}\right)^k, k=0,1,\ldots$$

$$E[X]=r\beta<Var[X]=r\beta(1+\beta)$$

$$M_X(t)=[1-\beta(e^t-1)]^{-r} \quad P_X(t)=E\left[t^X\right]=[1-\beta(t-1)]^{-r}$$

负二项（两参数）（续）

$$NB(r,p)$$

$$p_k = \begin{pmatrix} k+r-1 \\ k \end{pmatrix} p^r (1-p)^k, k = 0, 1, \ldots$$

$$E[X] = \frac{1-p}{p} r < Var[X] = \frac{1-p}{p^2} r,$$

$$M_X(t) = \left(\frac{p}{1-(1-p)e^t}\right)^r, \ P_X(t) = \left(\frac{p}{1-(1-p)t}\right)^r$$

二项（两参数）

$$B(n,q),$$

$$p_k = \binom{n}{k} q^k (1-q)^{n-k}, \quad k = 0, 1, \ldots, n$$

$$E[X] = nq > Var[X] = nq(1-q)$$

$$M_X(t) = [1+q(e^t-1)]^n \qquad P_X(t) = [1+q(t-1)]^n$$

从个体模型到聚合模型

对于 n 个独立同分布的 $X_1, X_2, \cdots, X_n$ 随机变量，可以将 $X_1, X_2, \cdots, X_n$ 中非零的变量按照观测的顺序加和为：

$$S = \tilde{X}_1 + \tilde{X}_2 + \cdots + \tilde{X}_N$$

则：

$$N \sim B(n, q)\text{，}\quad q = \Pr(X_i > 0)\text{，}\quad \tilde{X}_j = X_i \mid X_i > 0$$

$$\Pr(\tilde{X}_j = 0) = 0\text{，}\quad f_{\tilde{X}_j}(x) = \frac{f_X(x)}{\Pr(X > 0)}, x > 0$$

聚合模型与个体模型

- 从“个体”模型变换为“聚合”模型：

$$N \sim B(n,q) \qquad f_{\tilde{X}}(x) = \frac{f_X(x)}{\Pr(X>0)}, x>0$$

- 从“聚合”模型变换为“个体”模型

$$F_{\tilde{X}}(x) = p + qF_X(x)$$

$$n\Pr\left(\widetilde{X}_i > 0\right) = EN, \quad \left[\Pr\left(\widetilde{X}_i = 0\right)\right]^n = \Pr(N=0)$$

聚合模型的一般性质

- 分布表达式1

$$F_S(x) = \Pr(S \le x)$$

$$= \sum_{n=0}^{\infty} \Pr(S \le x \mid N = n) \Pr(N = n)$$

$$= \sum_{n=0}^{\infty} \Pr\{\sum_{1}^{n} X_i \le x\} \Pr(N = n)$$

$$= \sum_{n=0}^{\infty} \Pr(N = n) F^{*(n)}(x)$$

- 分布表达式2

$$
\begin{aligned}
F_S(x) &= \Pr(N=0) + \Pr(N>0)\cdot\Pr(0<S\le x \mid S>0) \\
&= \Pr(N=0) + \Pr(N>0)\cdot\Pr\left(\widetilde{S}\le x\right) \\
&= \Pr(N=0) + \Pr(N>0)\cdot F_{\widetilde{S}}(x) \\
&= p_0 + (1-p_0)\cdot F_{\widetilde{S}}(x)
\end{aligned}
$$

$$\widetilde{S} = S \mid S>0$$

矩的性质：

- 期望

$$E[S]=E[E(S\mid N)]$$
$$=E[N\cdot E(X)]=E[N]\cdot E[X]$$

- 方差

$$Var[S]=E[N]\cdot Var(X)+E^2[X]\cdot Var[N]$$

概率生成函数

- 矩母函数

$$M_S(t) = E[e^{tS}] = E[E[e^{tS} \mid N]] = E[(M_X(t))^N]$$

$$= P_N(M_X(t)) = E[e^{N\log(M_X(t))}] = M_N(\log(M_X(t)))$$

- Laplace变换

$$L_S(t) = L_N(\log(M_X(t)))$$

例1-8 复合几何分布

- 几何分布+指数=混合指数
  - 矩母函数表示

$$F_S(x)=\frac{1}{1+\beta}+\frac{\beta}{1+\beta}(1-e^{-\frac{\lambda}{1+\beta}x})$$

Poisson聚合模型

- 模型
- Panjer递推（定理1.1）
- 复合Poisson随机变量的独立和仍然为复合Poisson模型（定理1.2）
- 复合Poisson随机变量的分解（定理1.3）
- 例及精算考试题目

模型

- 分布

$$N \sim Poisson(\lambda)$$

- 记号

$$S \sim CP(\lambda, f_X(x))$$

基本性质：

$$E[S]=\lambda E(X)$$

$$Var[S]=\lambda E(X^2)$$

$$M_s(t)=e^{\lambda(M_X(t)-1)}$$

定理1.1-Panjer递推

$$S \sim CP\left(\lambda, f_X(x)\right),\ x = 1, 2, \ldots$$

则有以下结论：

$$f_S(0) = e^{-\lambda}$$

$$f_S(x) = \frac{\lambda}{x}\sum_{y=1}^{x} y f_X(y) f_S(x-y),\ x = 1, 2, \ldots$$

$$f_S(1) = \lambda f_X(1) f_S(0)$$

$$f_S(2) = \frac{\lambda}{2}[f_X(1) f_S(1) + 2 f_X(2) f_S(0)]$$

证明

$$M_S'(t)=\lambda M_X'(t)e^{\lambda(M_X(t)-1)}=\lambda M_X'(t)M_S(t)=\lambda\sum_{k=0}^{\infty}ke^{kt}f_X(k)\cdot\sum_{k=0}^{\infty}e^{kt}f_S(k)$$

$$M_S'(t)=\sum_{k=0}^{\infty}ke^{kt}f_S(k)$$

$$xf_S(x)=\lambda\sum_{y=1}^{x}yf_X(y)\cdot f_S(x-y)$$

定理1.2：独立的复合Poisson之和仍然为复合Poisson

- 已知：

$$S_i \sim CP(\lambda_i, f_i(x))$$

- 结论：

$$S = S_1 + S_2 + \cdots + S_m \sim CP\left(\lambda, f_X(x)\right)$$

$$\lambda = \sum_{i=1}^{m} \lambda_i,\ f_X(x) = \sum_{i=1}^{m} \frac{\lambda_i}{\lambda} f_i(x)$$

证明

$$M_S(t)=\prod_{i=1}^{m}M_{S_i}(t)=\prod_{i=1}^{m}\exp\left[\lambda_i\left(M_{X_i}(t)-1\right)\right]$$

$$=\exp\left\{\sum_{i=1}^{m}\lambda_i\left(M_{X_i}(t)-1\right)\right\}$$

$$=\exp\left\{\lambda\left[\sum_{i=1}^{m}\frac{\lambda_i}{\lambda}M_{X_i}(t)-1\right]\right\}=\exp\left\{\lambda\left[M_X(t)-1\right]\right\},$$

其中,

$$\lambda=\sum_{i=1}^{m}\lambda_i,\ M_X(t)=\sum_{i=1}^{m}\frac{\lambda_i}{\lambda}M_{X_i}(t)$$

例1-11

$$\sum_{1}^{m} N_i x_i \sim CP\left(\lambda, F_X(x)\right)$$

$$\lambda = \sum_{1}^{m} \lambda_i \qquad f_X(x) = \begin{cases} \dfrac{\lambda_i}{\lambda}, x = x_i \\ 0, \text{否则} \end{cases}$$

定理1.3-复合Poisson的分解

已知： $S \sim CP(\lambda, f_X(x))$

则X支集存在划分 $A_1, A_2, \ldots, A_m$ 记 $p_i = \Pr(X \in A_i)$ 有

$$S = S_1 + S_2 + \cdots + S_m$$

这里， $S_i = \sum_{j=1}^{N_i} X_j^{(i)} \sim CP(\lambda_i, f_i(x))$ 其中， $\lambda_i = \lambda p_i$

$$N_i = \begin{cases} \sum_{j=1}^{N} 1_{\{X_j \in A_i\}}, & N > 0 \\ 0, & N = 0 \end{cases}, i = 1, 2, \ldots, m$$

$$f_i(x) = \begin{cases} \frac{f(x)}{p_i}, & x \in A_i \\ 0, & \text{else} \end{cases}, i = 1, 2, \ldots, m$$

定理1.3的应用

- 保单限额的处理。
  - 赔付有最大限额
  - 总损失分为两个部分
- 免赔责任的处理。
  - 总损失分为两个部分

例1-12

已知：

$S \sim CP(0.8, f_X(x))$

$f_X(1) = 0.25, f_X(2) = 0.375, f_X(3) = 0.375$

计算S的概率函数。

方法一：直接计算

计算结果如下表

x	$p^{*0}(x)$	$p(x)$	$p^{*2}(x)$	$p^{*3}(x)$	$p^{*4}(x)$	$p^{*5}(x)$	$p^{*6}(x)$	$f_S(x)$
0	1							$e^{-0.8}$
1		0.25						$0.25\,e^{-0.8}$
2		0.375	0.0625					$0.32\,e^{-0.8}$
3		0.375	0.1875	0.015625				$0.36\,e^{-0.8}$
4			0.328125	0.070313	0.003906			$0.11\,e^{-0.8}$
5			0.28125	0.175781	0.023438	0.000977		$0.11\,e^{-0.8}$
6			0.140625	0.263672	0.076172	0.007324	0.000024	$0.07\,e^{-0.8}$

方法二

$$S = 1N_1 + 2N_2 + 3N_3$$

x	$P\{N_1 = x\}$	$P\{2N_2 = x\}$	$P\{3N_3 = x\}$	$P\{N_1 + 2N_2 = x\}$	$f_S(x) = P\{N_1 + 2N_2 + 3N_3 = x\}$
0	0.818731	0.740818	0.740818	0.606531	0.449329
1	0.163746	---	---	0.121306	0.089866
2	0.016375	0.222245	---	0.194090	0.143785
3	0.001092	---	0.222245	0.037201	0.162358
4	0.000055	0.033337	---	0.030974	0.049906
5	0.000002	---	---	0.005703	0.047960
6	0.000000	0.003334	0033337	0.003288	0.030923

方法三：Panjer递推

$\lambda = 0.8$， $f_S(0) = e^{-0.8}$，

$$f_S(1) = \frac{0.8}{1} p(1) f_S(0) = 0.8 \times 0.25 \times f_S(0) = 0.2e^{-0.8}$$

$$f_S(2) = \frac{0.8}{2} \{0.25 f_S(1) + 2 \times 0.375 f_S(0)\} = 0.032e^{-0.8}$$

$$f_S(3) = \frac{0.8}{3} \{0.25 f_S(2) + 2 \times 0.375 f_S(1) + 3 \times 0.375 f_S(0)\} = 0.361333e^{-0.8}$$

关于复合Poisson模型的进一步思考

- 独立性条件：
  - 索赔数变量与索赔额变量的独立
  - 索赔额变量之间的独立
- 在信用风险管理中的应用
  - 基本模型
  - 风险暴露
  - 相关性

作业和思考

- 作业：
  - 第一章习题1(P.40):9-12/16-19
- 思考：
  - 个体模型与聚合模型的主要差异
  - 现有两个随机变量：
    - $\boldsymbol{S_1 = X_1 + \cdots + X_n, X_1, \dots, X_n}$为$\boldsymbol{i.i.d.\sim B(1,p)}$
    - $\boldsymbol{S_2 \sim B(n,p)}$
    - 通过随机模拟(学习已知分布模拟样本的方法)的outcome说明两者的样本差异