

第三章 总理赔额模型

本章主要内容

- 个体风险模型: $S = X_1 + X_2 + \cdots + X_n$
- 集体风险模型: $S = X_1 + X_2 + \cdots + X_N$
- 复合泊松分布

第一节 个体风险模型

设在某一定时期内保险人承保 n 份保单，每份保单的赔付变量分别为 X_i ， $i=1,2,\cdots,n$ ，则总理赔额为

$$S = X_1 + X_2 + \cdots + X_n \quad (*)$$

假定

- (1) 每张保单是否发生理赔以及理赔额大小是相互独立的，互不影响的，即 X_1, X_2, \cdots, X_n 相互独立的随机变量；
- (2) 每张保单至多发生一次理赔。

用随机变量 I 表示理赔发生情况,

$$I = \begin{cases} 0, & \text{不发生理赔} \\ 1, & \text{发生理赔} \end{cases}$$

设 $q_j = P(I_j = 1)$ 表示第 j 份保单发生理赔的概率, 则第 j 张保单的实际赔付额 X_j 可以表示为

$$X_j = I_j B_j = \begin{cases} 0, & 1 - q_j \\ B_j, & q_j \end{cases}$$

其中 B_j 表示在第 j 次理赔发生条件下的理赔额。 $\{B_j\}$ 与 $\{I_j\}$ 分别是独立的随机变量序列, 并且 $\{B_j\}$ 与 $\{I_j\}$ 之间相互独立。总理赔额 S 可以表示为

$$S = \sum_{i=1}^n X_i = \sum_{i=1}^n I_i B_i$$

例 3.1: 人寿保险

$$X_j = \begin{cases} 0, & 1 - q_j \\ b_j, & q_j \end{cases}$$
$$I_j = \begin{cases} 0, & \text{不发生死亡} \\ 1, & \text{死亡发生} \end{cases}$$

例 3.2: 令 $n_{ij}, i=1,2,\dots,r$ 表示获赔 i 个单位的赔款, 发生理赔的概率为 $q_j, j=1,2,\dots,m$ 的保单个数, 则总理赔额为

$$S = \sum_{i=1}^r \sum_{j=1}^m \sum_{k=1}^{n_{ij}} X_{ij}^{(k)}$$

其中 $X_{ij}^{(k)} = iI_j, k=1,\dots,n_{ij}$ 表示获赔 i 个单位的赔款, 且发生理赔的概率为 q_j 的保单个数。

一、个体保单赔付额的期望和方差

设 $P(I_j=1)=q_j$ ， $f_{B_j}(x)$ 为 B_j 的分布， $u_j = E(B_j)$ 。则 X_j 的分布是一个混合分布，它的分布函数为

$$f_{X_j}(x) = \begin{cases} 1-q_j, & x=0 \\ q_j f_{B_j}(x), & x>0 \end{cases}$$

$$E(X_j) = u_j q_j, \quad E(S) = \sum_{j=1}^n u_j q_j$$

命题 1 设 $\sigma_j = \text{Var}(B_j)$ ，则

$$\text{Var}(X_j) = u_j^2 q_j (1-q_j) + \sigma_j^2 q_j$$

$$\text{Var}(S) = \sum_{j=1}^n \left[u_j^2 q_j (1-q_j) + \sigma_j^2 q_j \right]$$

证明：由方差分解公式

$$\text{Var}(X) = E(\text{Var}(X | I_j)) + \text{Var}(E(X | I_j))$$

其中 $\text{Var}(X_j | I_j = 1) = \text{Var}(B_j | I_j = 1) = \sigma_j^2$, $\text{Var}(X_j | I_j = 0) = 0$, 因此

$$\text{Var}(X_j | I_j) = \sigma_j^2 I_j$$

$$E(\text{Var}(X_j | I_j)) = \sigma_j^2 q_j$$

利用 $E(X_j | I_j) = u_j I_j$, 有

$$\text{Var}(E(X | I)) = \text{Var}(u_j I) = u_j^2 \text{Var}(I_j) = u_j^2 q_j (1 - q_j)$$

所以

$$\begin{aligned} \text{Var}(X) &= E(\text{Var}(X_j | I)) + \text{Var}(E(X_j | I)) \\ &= u_j^2 q_j (1 - q_j) + \sigma_j^2 q_j \end{aligned}$$

$$\text{Var}(S) = \sum_{j=1}^n \left[u_j^2 q_j (1 - q_j) + \sigma_j^2 q_j \right]$$

例 3.3 某公司为员工购买意外死亡寿险。假设对所有人明年的死亡概率为 0.01，且 30% 的死亡是由于意外事故发生的。75 名雇员分属两个保单组，其中的 50 人如果是正常死亡，保险人将赔付 5 万元；如果是意外死亡，保险人将赔付 10 万元。另外 25 人的赔付额分别为 7.5 万元和 15 万元。求总赔付额的期望和方差。

解：对这 75 人来说，死亡概率 $q_j = 0.01$ 。根据题意，50 名员工的赔付额为

$$B_j = \begin{cases} 50,000, & 0.7 \\ 100,000, & 0.3 \end{cases}, \quad j = 1, \dots, 50,$$

$$\mu_j = E(B_j) = 0.7 \times 50,000 + 0.3 \times 100,000 = 65,000,$$

$$\sigma_j^2 = \text{Var}(B_j) = 0.7 \times 50,000^2 + 0.3 \times 100,000^2 - 65,000^2 = 525,000,000$$

另外 25 名员工的赔付额

$$B_j = \begin{cases} 75,000, & 0.7 \\ 150,000, & 0.3 \end{cases}, \quad j = 51, 52, \dots, 75$$

$$\mu_j = E(B_j) = 0.7 \times 75,000 + 0.3 \times 150,000 = 97,500,$$

$$\sigma_j^2 = \text{Var}(B_j) = 0.7 \times 75,000^2 + 0.3 \times 150,000^2 - 97,500^2 = 118,125,000,0$$

由命题 1 有

$$E(S) = 50 \times 0.01 \times 65,000 + 25 \times 0.01 \times 97,500 = 56875$$

$$\begin{aligned} Var(S) &= 50 \times (0.01 \times 0.99 \times 65,000^2 + 525,000,000 \times 0.01) \\ &\quad + 25 \times (0.01 \times 0.99 \times 97,500^2 + 1,181,250,000 \times 0.01) \\ &= 5,001,984,375 \end{aligned}$$

二、独立分布和的卷积

对于 X 与 Y 相互独立的离散非负随机变量，设它们的分布列分别为 p_X 和 p_Y ，则 $S = X + Y$ 的分布为

$$\begin{aligned} P(S=s) &= P(X+Y=s) \\ &= \sum_{y=0}^s P(X+Y=s | Y=y) P(Y=y) \\ &= \sum_{y=0}^s P(X=s-y | Y=y) P(Y=y) \\ &= \sum_{y=0}^s p_X(s-y) p_Y(y) \end{aligned} \quad (1)$$

对于两个相互独立的连续型非负随机变量 X 和 Y ，设其分布密度分别为 $f_X(x)$ 和 $f_Y(y)$ ，它们的联合密度为 $f_{(X,Y)}(x,y)$ ，则独立性知

$f_{(X,Y)}(x,y) = f_X(x)f_Y(y)$, $S = X + Y$ 的分布函数为

$$\begin{aligned} F_S(s) &= P(X + Y \leq s) \\ &= \iint_{x+y \leq s} f_{(X,Y)}(x,y) dx dy \\ &= \iint_{x+y \leq s} f_X(x) f_Y(y) dx dy \\ &= \int_0^s f_X(x) \left[\int_0^{s-x} f_Y(y) dy \right] dx \\ &= \int_0^s f_X(x) F_Y(s-x) dx \end{aligned}$$

所以, S 的分布密度为

$$f_S(s) = \int_0^s f_X(x) f_Y(s-x) dx \quad (2)$$

利用求和的可交换性, S 的分布密度也可以写为

$$f_S(s) = \int_0^s f_X(s-y) f_Y(y) dy$$

设 $\{X_i, i=1, \dots, n\}$ 为相互独立的随机变量, X_i 的分布记为 F_i , $S_k = X_1 + X_2 + \dots + X_k$ 的分布函数为记 $F^{(k)}$, 则由卷积的定义

$$\begin{aligned}
 F^{(1)} &= F_1 \\
 F^{(2)} &= F_2 * F^{(1)} = F_2 * F_1 \\
 F^{(3)} &= F_3 * F^{(2)} \\
 &\vdots \\
 F_S &= F^{(n)} = F_n * F^{(n-1)}
 \end{aligned} \tag{3}$$

例 3.4: 设随机变量 X_1, X_2, X_3 相互独立, 它们的分布列分别为

$$X_1 \sim \begin{pmatrix} 0 & 1 & 2 \\ 0.5 & 0.3 & 0.2 \end{pmatrix}, \quad X_2 \sim \begin{pmatrix} 0 & 1 & 2 & 3 \\ 0.4 & 0.3 & 0.2 & 0.1 \end{pmatrix}$$

$$X_3 \sim \begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 0.5 & 0 & 0.3 & 0.1 & 0.1 \end{pmatrix}$$

用卷积方法求 $S = X_1 + X_2 + X_3$ 的分布。

解: 设 $f_1(x)$ 、 $f_2(x)$ 、 $f_3(x)$ 分别为 X_1, X_2, X_3 的分布列, $f^{(1)}(x)$ 、 $f^{(2)}(x)$ 和 $f^{(3)}(x)$ 为 S_1, S_2, S_3 的分布列,

利用两项和卷积公式, 先计算 $X_1 + X_2$ 的分布。

$$f^{(2)}(0) = P(X_1 = 0, X_2 = 0) = 0.4 \times 0.5 = 0.20$$

$$f^{(2)}(1) = P(X_1 = 0, X_2 = 1) + P(X_1 = 1, X_2 = 0) = 0.5 \times 0.3 + 0.3 \times 0.4 = 0.27$$

$$\begin{aligned}f^{(2)}(2) &= P(X_1 = 0, X_2 = 2) + P(X_1 = 1, X_2 = 1) + P(X_1 = 2, X_2 = 0) \\&= 0.5 \times 0.2 + 0.3 \times 0.3 + 0.2 \times 0.4 = 0.27\end{aligned}$$

请同学们计算

$$f^{(2)}(3) = ?$$

$$f^{(2)}(4) = ?$$

由于 $S = (X_1 + X_2) + X_3$ ，所以 S 的分布等于 $X_1 + X_2$ 的分布 $f^{(2)}(x)$ 与 X_3 分布 $f_3(x)$ 的卷积，例如

$$\begin{aligned} f^{(3)}(2) &= P(X_1 + X_2 = 0, X_3 = 2) + P(X_1 + X_2 = 1, X_3 = 1) \\ &\quad + P(X_1 + X_2 = 2, X_3 = 0) \\ &= 0.20 \times 0.3 + 0.27 \times 0 + 0.27 \times 0.5 = 0.195 \end{aligned}$$

请计算

$$f^{(3)}(3) = ?$$

表 3.2.1

x	$f_1(x)$	$f_2(x)$	$f_3(x)$	$f^{(1)}(x)$	$f^{(2)}(x)$	$f^{(3)}(x)$	$F_1(x)$	$F^{(2)}(x)$	$F^{(3)}(x)$
0	0.5	0.4	0.5	0.5	0.2	0.1	0.5	0.2	0.1
1	0.3	0.3	0	0.3	0.27	0.135	0.8	0.47	0.235
2	0.2	0.2	0.3	0.2	0.27	0.195	1	0.74	0.43
3		0.1	0.1		0.17	0.186		0.91	0.616
4			0.1		0.07	0.163		0.98	0.779
5					0.02	0.115		1	0.894
6						0.065		1	0.959
7						0.03		1	0.989
8						0.009		1	0.998
9						0.002		1	1

例 3.5: 设 X_1, X_2 独立, 且与 X 的分布函数相同

$$f_X(x) = \frac{2}{100^2} (100 - x), 0 < x \leq 100$$

$S = X_1 + X_2$, 计算 $f_S(120)$ 。

解：由卷积公式知

$$f_S(120) = \int_0^{120} f_X(x) f_X(120-x) dx$$

积分区域为 $0 < x \leq 100$ 和 $0 < 120 - x \leq 100$ ，即 $20 \leq x \leq 100$ ，故上式可化为

$$\begin{aligned} f_S(120) &= \int_{20}^{100} f_{X_1}(x) f_{X_2}(120-x) dx \\ &= \int_{20}^{100} \frac{2}{100^2} (100-x) \frac{2}{100^2} (x-20) dx \\ &= \frac{4}{100^4} \left(-\frac{1}{3} x^3 + 60x^2 - 2000x \right) \Big|_{20}^{100} \\ &= 0.0034133 \end{aligned}$$

一定要注意积分区域。

在个体风险模型中，卷积法适用于计算个体保单理赔额固定的总理赔额的分布。假设个体保单的赔付额可以表示为 $X_j = I_j b_j$ ，其中 I_j 是 0—1 变量， b_j 为理赔发生时的理赔额。

$$f_{X_j}(x) = \begin{cases} p_j = 1 - q_j, & x = 0 \\ q_j, & x = b_j \end{cases}$$

记 $S_j = X_1 + X_2 + \cdots + X_j = S_{j-1} + X_j$ ，则由公式 (1) 可以计算 S 的分布如下：

$$\begin{aligned}
f_{S_j}(s) &= f_{S_{j-1}} * f_{X_j}(s) = \sum_{x=0}^s f_{S_{j-1}}(s-x) f_{X_j}(x) \\
&= \begin{cases} f_{S_{j-1}}(s) f_{X_j}(0), & s < b_j \\ f_{S_{j-1}}(s) f_{X_j}(0) + f_{S_{j-1}}(s-b_j) f_{X_j}(b_j), & s \geq b_j \end{cases} \quad (4) \\
&= \begin{cases} p_j f_{S_{j-1}}(s), & s < b_j \\ p_j f_{S_{j-1}}(s) + q_j f_{S_{j-1}}(s-b_j), & s \geq b_j \end{cases}
\end{aligned}$$

例 3.6: 某公司为 14 名员工购买了一年期团体定期寿险。保险公司的精算师选择了如下死亡表来计算该团体的死亡率。每个员工按他的工资水平（近似到 1000）进行投保。具体资料如下：

表 3.2.3

员工	年龄	性别	赔付	死亡率 q
1	20	男	15000	0.00149
2	23	男	16000	0.00142
3	27	男	20000	0.00128
4	30	男	28000	0.00122
5	31	男	31000	0.00123
6	46	男	18000	0.00353
7	47	男	26000	0.00394
8	49	男	24000	0.00484
9	64	男	60000	0.02182
10	17	女	14000	0.0005
11	22	女	17000	0.0005

12	26	女	19000	0.00054
13	37	女	30000	0.00103
14	55	女	55000	0.00479

请用卷积法来计算总赔付额的分布。

解：由公式（4）得

$$f_{s_1}(0) = 0.99851,$$

$$f_{s_1}(15) = 0.00149,$$

$$f_{s_2}(0) = p_2 f_{s_1}(0) = 0.99709212$$

$$f_{s_2}(15) = p_2 f_{s_1}(15) = 0.00148788,$$

$$f_{s_2}(16) = p_2 f_{s_1}(16) + q_2 f_{s_1}(0) = 0.00141788$$

$$f_{s_2}(31) = p_2 f_{s_1}(31) + q_2 f_{s_1}(15) = 0.0000021158$$

具体计算的结果见表 3.2.4

表 3.2.4

x	$F_s(x)$	x	$F_s(x)$	x	$F_s(x)$	x	$F_s(x)$
0	0.95273905	20	0.96157969	40	0.97335098	60	0.99933062
1	0.95273905	21	0.96157969	41	0.97335892	61	0.99933187
2	0.95273905	22	0.96157969	42	0.97338128	62	0.99933191
3	0.95273905	23	0.96157969	43	0.97338740	63	0.99933193
4	0.95273905	24	0.96621337	44	0.97340884	64	0.99933198
5	0.95273905	25	0.96621337	45	0.97341351	65	0.99933202
6	0.95273905	26	0.96998201	46	0.97342561	66	0.99933206
7	0.95273905	27	0.96998201	47	0.97342840	67	0.99933209
8	0.95273905	28	0.97114577	48	0.97343397	68	0.99933217
9	0.95273905	29	0.97114648	49	0.97343866	69	0.99933450
10	0.95273905	30	0.97212949	50	0.97345889	70	0.99934141
11	0.95273905	31	0.97330507	51	0.97346040	71	0.99934796
12	0.95273905	32	0.97330747	52	0.97346606	72	0.99935031
13	0.95273905	33	0.97331344	53	0.97346608	73	0.99936659
14	0.95321566	34	0.97331962	54	0.97347547	74	0.99937973
15	0.95463736	35	0.97332386	55	0.97806678	75	0.99941735

16	0.95599217	36	0.97332585	56	0.97807068	76	0.99944759
17	0.95646878	37	0.97332829	57	0.97807536	77	0.99945823
18	0.95984386	38	0.97333493	58	0.97807660	78	0.99953355
19	0.96035862	39	0.97334251	59	0.97807808	79	0.99956734

请有兴趣的同学用 R 软件来实现这个例子。

```
#加载 sfsmisc、startupmsg、SweaveListingUtils、distr 包
```

```
library(sfsmisc)
```

```
library(startupmsg)
```

```
library(SweaveListingUtils)
```

```
library(distr)
```

```
#每个被保险人的保险金额
```

```
b = c(15, 16, 20, 28, 31, 18, 26, 24, 60, 14, 17, 19, 30, 55)
```

```
#每个被保险人的死亡概率
```

```
q = c(0.00149, 0.00142, 0.00128, 0.00122, 0.00123, 0.00353, 0.00394, 0.00484,  
0.02182, 0.0005, 0.0005, 0.00054, 0.00103, 0.00479)
```

```
#用卷积法计算支付给该团体保单的保险金额的分布 s
```

```
s = 0
```



```
for(i in 1:length(b)){  
  x = DiscreteDistribution(supp = c(0, b[i]), prob = c(1 - q[i], q[i]))  
  s = s + x  
}  
dens = d(s)(seq(0, 79, by = 1))  
prob = p(s)(seq(0, 79, by = 1))  
cbind(seq(0, 79, by = 1), dens, prob)
```

三 矩母函数法

设 \mathbf{X} 为表示随机变量, 定义 $M_{\mathbf{X}}(t) = E[e^{t\mathbf{X}}]$ 为 \mathbf{X} 的矩母函数, $P_{\mathbf{X}}(z) = E(z^{\mathbf{X}})$ 为 \mathbf{X} 的母函数。

由于 $X_1, X_2 \cdots X_n$ 是相互独立的随机变量, 因此 \mathbf{S} 的矩母函数等于

$$\begin{aligned} M_{\mathbf{S}}(t) &= E(e^{t\mathbf{S}}) = E(e^{t \sum_{i=1}^n X_i}) = E(e^{tX_1} e^{tX_2} \cdots e^{tX_n}) \\ &= \prod_{i=1}^n E(e^{tX_i}) = \prod_{i=1}^n M_{X_i}(t) \end{aligned} \quad (5)$$

$$P_{\mathbf{S}}(z) = E(z^{\mathbf{S}}) = \prod_{i=1}^n P_{X_i}(z) \quad (6)$$

例 3.7: 设 X_1, \dots, X_n 独立同分布, 且 X_i 服从 gamma 分布,

$$f(x) = \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)}$$

设 $S = X_1 + X_2 + \dots + X_n$, 求 S 的分布。

解: 经计算得到

$$M_X(t) = (1 - t / \beta)^{-\alpha}$$

由公式 (5)

$$M_S(t) = \prod_{i=1}^n M_{X_i}(t) = (M_X(t))^n = (1 - t / \beta)^{-n\alpha}$$

S 服从 gamma 分布。

在个体风险模型中

当 $X_j = I_j b_j$ 时, 其中 I_j 是 0—1 变量, 则由公式 (6) 可以计算得到 S 的母函数为

$$P_S(t) = \prod_{j=1}^n (1 - q_j + q_j^{b_j} t) \quad (7)$$

当 $X_j = I_j B_j$ 时, 类似可以计算 S 的矩母函数和母函数分别为

$$M_S(t) = \prod_{j=1}^n (1 - q_j + q_j M_{B_j}(t))$$

$$P_S(z) = \prod_{j=1}^n (1 - q_j + q_j P_{B_j}(z))$$

当 $X_{ij} = i I_j$ 时, 设 n_{ij} 表示获赔 i 个单位的赔款, 发生理赔的概率为 q_j

的保单个数, 总理赔额为 $S = \sum_{i=1}^r \sum_{j=1}^m \sum_{k=1}^{n_{ij}} X_{ij}^{(k)}$,

$$f_{x_{ij}}(x) = \begin{cases} 1 - q_j, & x = 0 \\ q_j, & x = i \end{cases}$$

$$P_S(z) = \prod_{i=1}^r \prod_{j=1}^m (1 - q_j + q_j z^i)^{n_{ij}}$$

有兴趣的同学阅读以下部分 #####

两边取对数得到

$$\log P_S(z) = \sum_{i=1}^r \sum_{j=1}^m n_{ij} \log(1 - q_j + q_j z^i)$$

$$P_S'(z) = P_S(z) \left[\sum_{i=1}^r \sum_{j=1}^m i q_j n_{ij} z^{i-1} (1 - q_j + q_j z^i)^{-1} \right] \quad (8)$$

令 $z=1$ ，我们可以得到总理赔额的期望为

$$E(S) = P'_S(1) = \sum_{i=1}^r \sum_{j=1}^m i q_j n_{ij}$$

将 (8) 改为

$$\begin{aligned} z P'_S(z) &= P_S(z) \left[\sum_{i=1}^r \sum_{j=1}^m i n_{ij} \left(\frac{q_j}{1-q_j} z^i \right) \left(1 + \frac{q_j z^i}{1-q_j} \right)^{-1} \right] \\ &= P_S(z) \left[\sum_{i=1}^r \sum_{j=1}^m i n_{ij} \sum_{k=1}^{\infty} (-1)^{k-1} \left(\frac{q_j}{1-q_j} \right)^k z^{ik} \right] \end{aligned} \quad (9)$$

其中第二个等式成立要求 $|z| < \min_{i,j} \{q_j^{-1}(1-q_j)\}^{1/i}$ 。若令

$$h(i, k) = i - \sum_{j=1}^m n_{ij} \frac{q_j}{1-q_j}^k \quad (10)$$

则 (9) 可以写为

$$zP'_S(z) = P_S(z) \left[\sum_{i=1}^r \sum_{k=1}^{\infty} h(i,k) z^{ik} \right] \quad (11)$$

将公式(11)用级数展开, 其中左边 z^x 的系数是 $xf_S(x)$, 而 $f_S(x)$ 是 $P_S(z)$ 中 z^x 的系数。公式(11)右边 z^x 的系数为

$$\sum_{i \leq x} h(i, k) f_S(x - ik) \quad (12)$$

记 $[z]$ 表示实数 z 的整数部分, 则(12)可以写为

$$\sum_{i=1}^x \sum_{k=1}^{[x/i]} h(i, k) f_S(x - ik)$$

当 $i > x$ 时, $h(i, k) = 0$ 。由于公式(11)左右两边系数相等, 因此

$$f_S(x) = \frac{1}{x} \sum_{i=1}^{x \wedge r} \sum_{k=1}^{[x/i]} h(i, k) f_S(x - ik) \quad (13)$$

由于

$$f_s(0) = P_s(0) = \prod_{i=1}^r \prod_{j=1}^m (1 - q_j)^{n_{ij}}$$

利用（13）计算得到

$$f_s(1) = h(1,1)f_s(0)$$

$$f_s(2) = \frac{1}{2} \{h(1,1)f_s(1) + [h(1,2) + h(2,1)]f_s(0)\}, \dots$$

依次计算就可以得到 S 的分布 $\{f_s(x), x=1, 2, \dots\}$ 。

从（10）可以看出， $h(i, k)$ 是 $\left[q_j / (1 - q_j) \right]^k$, $j=1, \dots, m$ 的加权和。当 q_j 越接近于 0 时， $\left[q_j / (1 - q_j) \right]^k$ 也越小。从而，当 k 的值增加时， $h(i, k)$ 的值将会下降。因此，在一定的精度保证下，公式（13）中 $f_s(x)$ 可用下面的求和项近似

$$f_s^{(K)}(x) = \frac{1}{x} \sum_{i=1}^{x \wedge r} \sum_{k=1}^{K \wedge [x/i]} h(i, k) f_s^{(K)}(x - ik) \quad (14)$$

事实上，De Pril 证明若 $q_j < \frac{1}{2}, j=1, \dots, m$ ，则

$$\sum_{x=0}^M |f_s(x) - f_s^{(K)}(x)| < e^{\delta(K)} - 1 \quad (15)$$

其中

$$\delta(K) = \frac{1}{K+1} \sum_{i=1}^r \sum_{j=1}^m n_{ij} \frac{1-q_j}{1-2q_j} \left(\frac{q_j}{1-q_j} \right)^{K+1}, \quad (16)$$

$M = \sum_{i=1}^r \sum_{j=1}^m in_{ij}$ 是总理赔额的最大取值。

例 3.6 续： 请用公式 (14) 计算 S 的分布。

由 (16) 计算得 $\delta(1)=5.947\times 10^{-4}$, $\delta(2)=3.9\times 10^{-6}$, $\delta(3)=6.369\times 10^{-8}$, $\delta(4)=1.131\times 10^{-9}$ 。当 $K=4$ 时, $f_s^{(4)}(x)$ 和 $f_s(x)$ 的误差小于小数点后八位数。表 3.3.1, 3.3.2, 3.3.3 分别给出了 $n_{ij}, h(i, k), f_s^{(4)}(x), F_s^{(4)}(x)$ 的值。

表 3.3.1

[illegible]

9	3.53	0	0	0	0	1	0	0	0	0	0	0	0	0	0
10	3.94	0	0	0	0	0	0	0	0	1	0	0	0	0	0
11	4.79	0	0	0	0	0	0	0	0	0	0	0	0	1	0
12	4.84	0	0	0	0	0	0	0	1	0	0	0	0	0	0
13	21.82	0	0	0	0	0	0	0	0	0	0	0	0	0	1

表 3.3.2

i	$h(i, k)$			
	k=1	k=2	k=3	k=4
14	7.0035018E-03	-3.5035026E-06	1.7526300E-09	-8.8000000E-13
15	2.2383351E-02	-3.3400961E-05	4.9841700E-08	-7.4370000E-11
16	2.2752308E-02	-3.2354221E-05	4.6008330E-08	-6.5420000E-11
17	8.5042522E-03	-4.2542531E-06	2.1281907E-9	-1.0646277E-12
18	6.3765090E-2	-2.2588816E-4	8.0020991E-7	-2.8347477E-9

19	1.0265543E-02	-5.5463885E-06	2.9966700E-09	-1.6200000E-12
20	2.5632810E-02	-3.2852047E-05	4.2104510E-08	-5.3960000E-11
24	1.1672495E-01	-5.6769640E-04	2.7610139E-06	-1.3428300E-08
26	1.0284521E-01	-4.0681297E-04	1.6091833E-06	-6.3652600E-09
28	3.4201726E-02	-4.1777074E-05	5.1030290E-08	-6.2330000E-11
30	3.0931860E-02	-3.1892665E-05	3.2883310E-08	-3.3900000E-11
31	3.8176958E-02	-4.7015487E-05	5.7900270E-08	-7.1310000E-11
55	2.6471800E-01	-1.2741022E-03	6.1323232E-06	-2.9515210E-08
60	1.3384040E+00	-2.9855420E-02	6.6597688E-04	-1.4855768E-05

表 3.3.3

x	$f_s^{(K)}(x)$	$F_s^{(K)}(x)$		$f_s^{(K)}(x)$	$F_s^{(K)}(x)$
0	0.95273905	0.952739	38	0.000006643644	0.973335
14	0.000476607829	0.953216	39	0.000007574225	0.973343

15	0.001421699516	0.954637	40	0.000008474451	0.973351
16	0.001354813286	0.955992	41	0.000007941654	0.973359
17	0.000476607829	0.956469	42	0.000022356100	0.973381
18	0.003375082888	0.959844	43	0.000006125396	0.973387
19	0.000514757056	0.960359	44	0.000021435448	0.973409
20	0.001221068952	0.96158	45	0.000004672158	0.973414
24	0.004633684032	0.966213	46	0.000012100769	0.973426
26	0.003768640299	0.969982	47	0.000002791514	0.973428
28	0.001163761430	0.971146	48	0.000005562190	0.973434
29	0.000000711205	0.971146	49	0.000004696495	0.973439
30	0.000983010770	0.972129	50	0.000020226469	0.973459
31	0.001175572312	0.973305	51	0.000001510358	0.97346
32	0.000002399591	0.973307	52	0.000005666162	0.973466
33	0.000005971631	0.973313	53	0.000000013706	0.973466
34	0.000006178405	0.97332	54	0.000009392317	0.973475
35	0.000004242488	0.973324	55	0.004591308358	0.978067

#通过循环计算矩阵 **h** 中各元素的值

```
for(i in 1:14){  
  for(k in 1:4){  
    x[i,k]=0  
    for(j in 1:13){  
      x[i,k]=x[i,k]+n[j,i]*(q[j]/(1-q[j]))^k}  
      h[i,k]=b[i]*(-1)^(k-1)*x[i,k]}}
```

#计算赔偿额 **fsk** 的近似密度函数

K=4;

```
fs0=1;  
for(i in 1:14){  
  for(j in 1:13){  
    fs0=fs0*(1-q[j])^(n[j,i])}}  
s = NULL  
fsk = NULL  
for(x in 1:100){  
  s[x] = 0
```

```

for(i in 1:14){
  if(min(K, floor(x/b[i]))>=1){
    for(k in 1:min(K, floor(x/b[i]))){
      if(x==b[i]*k){s[x]=s[x]+h[i,k]*fs0}
      else{s[x] = s[x] + h[i,k]*fsk[x-b[i]*k]}}}}
fsk[x]=1/x*s[x]}

```

#计算累积分布函数 Fs

```
Fs = NULL
```

```
Fs[1]=fs0
```

```
Fs[2]=fs0+fsk[1]
```

```
for(x in 3:100){
```

```
  Fs[x]=Fs[x-1]+fsk[x]}
```

#显示结果

```
h
```

```
fsk
```

```
Fs
```


四、近似算法

对于数目较大的保单组合来说，更实用的方法是求出近似分布。

定理 (独立同分布的中心极限定理) 设随机变量 X_1, X_2, \dots, X_n 相互独立，服从同一分布，且具有数学期望和方差， $E(X_k) = \mu$ ，

$Var(X_k) = \sigma^2 \neq 0 (k=1, 2, \dots)$ 。记 $S = \sum_{i=1}^n X_i$ ，则随机变量

$$\frac{S - E(S)}{\sqrt{Var(S)}} = \frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}\sigma}$$

的分布函数 $F_n(x)$ 对任意 x 满足

$$\lim_{n \rightarrow \infty} F_n(x) = \lim_{n \rightarrow \infty} P\left(\frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}\sigma} \leq x\right) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$$

对于独立但不同分布的随机变量和，在一定条件下，中心极限定理同样成立。

定理 (李雅普诺夫定理) 设随机变量 X_1, X_2, \dots, X_n 相互独立，它们具有数学期望和方差：

$$E(X_k) = \mu_k, \quad \text{Var}(X_k) = \sigma_k^2 \neq 0, \quad k = 1, 2, \dots,$$

记

$$B_n^2 = \sum_{k=1}^n \sigma_k^2$$

若存在正数 δ ，使得当 $n \rightarrow \infty$ 时，

$$\frac{1}{B_n^{2+\delta}} \sum_{k=1}^n E\{|X_k - \mu_k|^{2+\delta}\} \rightarrow 0$$

则随机变量

$$Z_n = \frac{\sum_{i=1}^n X_i - E(\sum_{i=1}^n X_i)}{\sqrt{\text{Var}(\sum_{i=1}^n X_i)}} = \frac{\sum_{i=1}^n X_i - \sum_{i=1}^n \mu_i}{B_n}$$

的分布函数对任意 x 有

$$\lim_{n \rightarrow \infty} F_n(x) = \lim_{n \rightarrow \infty} P\left(\frac{\sum_{i=1}^n X_i - \sum_{i=1}^n \mu_i}{B_n} \leq x\right) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$$

例 3.8 某保险公司出售了 300 张火灾险保单。已知保单分为两类，具体信息如下

类别	保单数	理赔发生概率	期望理赔额
1	200	0.05	1,000
2	100	0.01	2,000

假设：

- (1) 在理赔发生的条件下，理赔额服从指数分布。
- (2) 每张保单至多只发生一次理赔。
- (3) 理赔的发生与理赔额的大小独立。

求：

- (1) 计算这 300 张保单的总理赔额的方差。
- (2) 假设保险人收取的保费等于 $(1+\theta)E(S)$ ， θ 称为安全附加系数， $\theta = 0.1$ ，用正态近似法计算总理赔额超过保费收入的概率

$$P(S > (1 + \theta)E(S))。$$

解：设类别 1 的个体理赔额为 $X^{(1)}$ ，总理赔额为 $S_1 = X_1^{(1)} + \cdots + X_{n_1}^{(1)}$ 。

设类别 2 的个体理赔额为 $X^{(2)}$ ，总理赔额为 $S_2 = X_1^{(2)} + \cdots + X_{n_2}^{(2)}$ 。设总理赔额为 S ，则

$$Var(S) = Var(S_1) + Var(S_2)。$$

对于类别 1：

$$E(X^{(1)}) = 0.05 \times 1000 = 50，$$

$$\begin{aligned} Var(X^{(1)}) &= u_1^2 q_1 (1 - q_1) + \sigma_1^2 q_1 \\ &= 1000^2 \times 0.05 \times 0.95 + 0.05 \times 1000^2 \\ &= 97500 \end{aligned}$$

$$E(S_1) = 200 \times 50 = 10000，$$

$$\text{Var}(S_1) = 200 \times 97500 = 1.95 \times 10^7$$

类似可计算（练习）

$$E(X^{(2)}) = 0.01 \times 2000 = 20 ,$$

$$\begin{aligned}\text{Var}(X^{(2)}) &= u_2^2 q_2 (1 - q_2) + \sigma_2^2 q_2 \\ &= 2000^2 \times 0.01 \times 0.99 + 0.01 \times 2000^2 \\ &= 79600\end{aligned}$$

$$E(S_2) = 100 \times 20 = 2000 ,$$

$$\text{Var}(S_2) = 100 \times 79600 = 0.796 \times 10^7$$

$$\begin{aligned}\text{Var}(S) &= \text{Var}(S_1) + \text{Var}(S_2) \\ &= 1.95 \times 10^7 + 0.796 \times 10^7 \\ &= 2.746 \times 10^7\end{aligned}$$

(2) 用正态分布近似 S 的分布，其均值和标准差分别为

$$E(S) = 12000$$

$$Std(S) = \sqrt{27460000} = 5240.23$$

故

$$\begin{aligned} P(S > (1 + \theta)E(S)) &= P\left(\frac{S - E(S)}{Std(S)} > \frac{0.1E(S)}{Std(S)}\right) \\ &= 1 - \Phi\left(\frac{0.1 \times 12000}{5240.23}\right) \\ &= 1 - \Phi(0.229) = 0.409 \end{aligned}$$

例 3.6 续：假设保险人收取的保费的安全附加系数为 45%，请用正态近似计算总理赔额超过保费收入的概率， $P(S > (1 + \theta)E(S))$ 。

解：

$$E(S) = \sum_{j=1}^{14} b_j q_j = 2054.41$$

$$\text{Var}(S) = \sum_{j=1}^{14} b_j^2 q_j (1 - q_j) = 1.02534 \times 10^8$$

保费等于 $1.45 \times 2054.41 = 2978.89$

$$P(S > 2.92889) = P(Z > \frac{2.92889 - 2.05441}{\sqrt{102.534}}) = 1 - \Phi(0.0913) = 0.46$$

注：在例 3.6 中

$$P(S > 2.92889) \approx 1 - P(S < 3) = 1 - 0.95273905 = 0.0473$$

这说明两者的偏差很大，请同学们分析原因？

