现代精算风险理论课程作业二

刘辰昂* 2013年6月18日

^{*}农业资源与环境1001,学号3100100050,Email:liuchenang@gmail.com

1.1

假设S是复合分布,试确定下面分布中的未知参数,

 $(1)N \sim NB(2,0.25), X \sim P(\lambda)$, 并且已经有 $f_S(0) = 0.067$.

 $(2)N \sim NB(4,\theta), X \sim G(\beta)$,并且已经有 $f_S(0) = 0.4, f_S(1) = 0.04$

解: (1)由题设可知,本题满足Panjer递推的条件,且这里p(0) > 0,所以这里由Panjer递推知

$$f(0) = m_N(\log p(0))$$

由于 $X \sim P(\lambda)$,故 $p(0) = e^{-\lambda}$,代入上式得

$$f(0) = m_N(-\lambda)$$

$$= (\frac{p}{1 - (1 - p)e^{-\lambda}})^2$$

$$= (\frac{0.25}{1 - 0.75e^{-\lambda}})^2$$

$$= 0.067$$

解得 $\lambda = 3.09$

(2)与第一小题类似,此时满足Panjer的条件,所以有

$$f(0) = m_N(\log \beta)$$

$$= (\frac{\theta}{1 - (1 - \theta)\beta})^r$$

$$= (\frac{\theta}{1 - (1 - \theta)\beta})^4$$

$$= 0.4$$

H

$$f(1) = \frac{1}{1 - (1 - \theta)\beta} 4(1 - \theta)(\beta(1 - \beta)) \cdot 0.4$$
$$= \frac{1.6(1 - \beta)(\beta(1 - \theta))}{1 - (1 - \theta)\beta}$$
$$= 0.04$$

联立两式整理得

$$\beta = \frac{1 - 0.4^{-0.25}\theta}{1 - \theta}$$
$$(1 - 0.4^{-0.25}\theta)(0.4^{-0.25} - 1) = \frac{0.04 \cdot 0.4^{-0.25}}{1.6}$$

由此解得

$$\theta = 0.6982$$
 $\beta = 0.4046$

1.2

假设有n份独立的保单,每份保单没有索赔的概率是 $1-\theta$,当索赔发生时,索赔大小服从 $\Gamma(\alpha,\beta)$.

- (1)计算总索赔额的矩母函数.
- (2)总索赔是不是可写成一个混合分布的形式S = IX + (1 I)Y? (注:本小题意义不大)

解: (1)不妨先求每份保单索赔额的矩母函数,事实上,每份保单的索赔额都写成如下形式的混合分布,即

$$X = (1 - I)X_1 + IX_2$$

这里 $P(I=0)=\theta, P(I=1)=1-\theta$,且 $f_{X_1}(x)=f(x)$,这里f(x)为 $\Gamma(\alpha,\beta)$ 的密度函数, $X_2=0$ 由题设

$$M_X(t) = E[e^{tX}]$$

$$= E[e^{X_1}]Pr[I=0] + E[e^{X_2}]Pr[I=1]$$

$$= \theta(\frac{\beta}{\beta - t})^{\alpha} + (1 - \theta)$$

故总索赔额8的矩母函数为

$$M_S(t) = \left[\theta \left(\frac{\beta}{\beta - t}\right)^{\alpha} + (1 - \theta)\right]^n$$

(2)可以。由题设可知发生索赔的概率为 $p = 1 - (1 - \theta)^n$,故可构造随机变量I使得 $Pr[I = 1] = 1 - (1 - \theta)^n$, $Pr[I = 0] = (1 - \theta)^n$,随机变量Y服从恒等于0的退化分布,X为满足如下形式的复合二项分布

$$X = \sum_{i=1}^{N} X_i$$

这里 $N \sim B(n,\theta), X_i \sim \Gamma(\alpha,\beta)$, 此时总索赔S即服从混合分布

$$S = IX + (1 - I)Y$$

1.3

假设S是复合分布,其中 $N \sim G(0.5)$, $X \sim P(3)$.请采用Panjer递推公式计算 $f_S(s), s = 0, 1$.

解: 由题设可知,此时S满足Panjer递推的条件,且p(0) > 0,故有

$$f(0) = m_N(\log p(0))$$

又由于此时

$$m_N(t) = \frac{\theta}{1 - (1 - \theta)e^t}$$
$$= \frac{0.5}{1 - 0.5e^t}$$
$$= \frac{\theta}{2 - e^t}$$

且 $p(0) = e^{-3}$,故有

$$f(0) = \frac{1}{2 - e^{-3}}$$

再考虑f(1),由于 $N \sim G(0.5)$,故 $a = 1 - \theta = 0.5, b = 0$,由递推式得

$$f(1) \neq \frac{1}{1 - 0.5e^{-3}} \cdot 0.5 \cdot p(1) \cdot f(0)$$
$$= \frac{3e^{-3}}{(2 - e^{-3})^2}$$

故综上所述

$$f(0) = \frac{1}{2 - e^{-3}}$$
$$f(1) = \frac{3e^{-3}}{(2 - e^{-3})^2}$$

1.4

保险人有五个独立的保险合同。每个保险合同有一个索赔的概率为0.2,分别以0.3的概率索赔1,0.7的概率索赔2.采用*Panjer* 递推,计算总索赔小

于5的概率。

解:设总理赔额为S,则由题设可将S转化为一复合分布,即

$$S = \sum_{i=1}^{N} X_i$$

这里 $N \sim B(5,0.2), p(1) = 0.3, p(2) = 0.7$,故p(0) = 0,且a = -0.25, b = 1.5,由Panjer递推式知

$$f(0) = Pr[N = 0] = 0.8^{5}$$

$$f(1) = \frac{1}{1 - (-0.25) \cdot p(0)} (-0.25 + 1.5) \cdot p(1) \cdot f(0)$$

$$= 0.12288$$

为便于计算,编写用于计算Panjer递推的R语言函数Panjer(),程序代码及说明见附录。输入代码:

source("Panjer.R");

library(pracma); #to define matrix

P<-zeros(1,4);

for(i in 1:4){

PP<-Panjer(N=3,c(5,0.2),

c(0,0.3,0.7),s=i);

P[i]<-PP\$Pi;

}

输出为

[,1] [,2] [,3] [,4

[1,] 0.12288 0.305152 0.0873984 0.1100806

即f(1) = 0.12288, f(2) = 0.305152, f(3) = 0.0873984, f(4) = 0.1100806 若输入如下代码则可直接输出索赔小于5即小于等于4的概率

PP<-Panjer(N=3,c(5,0.2),c(0,0.3,0.7),s=4); PP\$P;

输出为

[1] 0.953191

即

$$Pr[S < 5] = 0.9532$$

1.5

设 X_1, \cdots, X_n 是一列独立同分布的风险, D_1, \cdots, D_n 是一列独立同分布的随机变量,服从B(1,p),并且与 $\{X_k\}, 1 \le k \le n$ 独立.先考虑短期个体风险模型

$$S_n = \sum_{i=1}^n D_i X_i$$

和聚合风险模型

$$S_N = \sum_{i=1}^N X_i$$

其中: $N \sim B(n,p)$ 与 $\{X_k\}$, $1 \le k \le n$ 独立。证明

$$S_n \stackrel{d}{=} S_N$$

解: 事实上,只需证明 $m_{S_N}(t)=m_{S_n}(t)$ 即可,不妨令 $T_i=D_iX_i,1\leq k\leq n$,显然由题设可知, T_i 独立同分布,令其矩母函数为 $m_T(t)$,对任意的i,则由题设

$$m_{T}(t) = E[e^{tD_{i}X_{i}}]$$

$$= E[e^{tD_{i}X_{i}}|D_{i} = 1]Pr[D_{i} = 1] + E[e^{tD_{i}X_{i}}|D_{i} = 0]Pr[D_{i} = 0]$$

$$= E[e^{tX}]Pr[D_{i} = 1] + Pr[D_{i} = 0]$$

$$= p \cdot m_{X}(t) + (1 - p)$$

这里X(t)为 X_i 矩母函数,又由 T_i 独立同分布可知

$$m_{S_n}(t) = [m_T(t)]^n$$

= $[p \cdot m_X(t) + (1-p)]^n$

再考虑 $m_{S_N}(t)$,由复合分布的矩母函数性质知

$$m_{S_N}(t) = m_N(\log(m_X(t)))$$
$$= (1 - p + p \cdot m_X(t))^n$$

故 $m_{S_N}(t) = m_{S_n}(t)$, 命题得证。

1.6

假设某一险种A有一组有50份保险单,每份保单索赔的可能性为0.2, 其索赔大小服从参数为0.1 的指数分布。另一险种B共有70份保单,每份保 单索赔的可能性为0.3,索赔大小服从Pareto(3,30)。

- (1)计算两个险种总索赔额的均值和方差。
- (2)是不是可以采用聚合风险模型来近似这两个险种的总索赔。并计算总索赔额的期望和方差。

解:(1)不妨设险种A的总索赔额为 S_A ,险种B的总索赔额为 S_B ,由题设可知 S_A , S_B 均为复合分布,分别可以写成 1

$$S_A = \sum_{i=1}^{N} X_i$$

$$S_B = \sum_{i=1}^{M} Y_i$$

先考虑 S_A 均值与方差,这里 $N\sim B(50,0.2), X_i\sim exp(0.1)$,故由全期望公式得

$$E[S_A] = \sum_{n=0}^{\infty} E[X_1 + X_2 + \dots + X_N | N = n] Pr[N = n]$$

$$= \sum_{n=0}^{\infty} E[X_1 + X_2 + \dots + X_n] Pr[N = n]$$

$$= \sum_{n=0}^{\infty} n E[X_i] P[N = n]$$

$$= E[N] E[X_i]$$

$$= 10 \cdot 10$$

$$= 100$$

¹这里不发生理赔的概率已经接近于0,故认为至少发生一次理赔

再由条件方差公式

$$Var[S_A] = E[Var[S|N]] + Var[E[S|N]]$$

$$= E[N \cdot Var[X_i]] + Var[N \cdot E[X_i]]$$

$$= E[N]Var[X_i] + E[X_i]^2 Var[N]$$

$$= 10 \cdot 100 + 100 \cdot 8$$

$$= 1800$$

再考虑 S_B ,这里 $M \sim B(70,0.3), Y_i \sim Pareto(3,30)$,故由全期望公式得

$$E[S_B] = \sum_{n=0}^{\infty} E[Y_1 + Y_2 + \dots + Y_M | M = n] Pr[M = n]$$

$$= E[M]E[Y_i]$$

$$= 21 \cdot 45$$

$$= 945$$

再由条件方差公式

$$Var[S_B] = E[Var[S|N]] + Var[E[S|N]]$$

$$= E[M \cdot Var[Y_i]] + Var[M \cdot E[Y_i]]$$

$$= E[M]Var[Y_i] + E[Y_i]^2 Var[M]$$

$$= 21 \cdot 675 + 45^2 \cdot 14.7$$

$$= 43942.5$$

(2)可以,解答见第一小题。

1.7

- (1)采用NP近似分别求出 χ_{18}^2 分布的 $\alpha = 0.05, 0.1, 0.5, 0.9, 0.95$ 分位数值,并与 χ^2 精确分布值作比较。
- (2)采用平移伽马分布重新计算 χ_{18}^2 分布 $\alpha = 0.05, 0.1, 0.5, 0.9, 0.95$ 分位数值。
- (3)对两种近似方法作比较、分析。
- 解: (1)由题设可知, χ^2_{18} 分布的均值、方差、偏度分别为

$$\mu = 18, \sigma = 6, \gamma = \frac{2}{3}$$

不妨设 α 分位数为 X_{α} ,则由NP 近似的表达式

$$Pr[S \le x\sigma + \mu] \approx \Phi(\sqrt{\frac{9}{\gamma^2} + \frac{6x}{\gamma} + 1} + \frac{3}{\gamma}) = 1 - \alpha$$

这里 $X_{\alpha} = x\sigma + \mu$,编写用于计算NP近似的R语言函数NP(),代码及说明 附录。得到下表

表 1: 分位数计算结果

		** 3*				
	α	0.05	0.1	0.5	0.9	0.95
近似值 0.267007 10.73804 17.33333 26.11756 20.0061	真实值	9.390455	10.86494	17.3379	25.98942	28.8693
及[[6] 9.201901 10.19094 11.99999 20.11190 29.0001	近似值	9.267907	10.73894	17.33333	26.11756	29.00615

(2)由上小题结论,根据平移伽马近似的参数计算公式得

$$\alpha = \frac{4}{\gamma^2} = 9$$

$$\beta = \frac{2}{\gamma \sigma} = 0.5$$

$$x_0 = \mu - \frac{2\sigma}{\gamma} = 9 - 9 = 0$$

由于这里卡方分布原本就属于伽马分布)故由平移伽马近似所得的结果与真实值一致,结果可见第一小题表格中的真实值。对于平移伽马近似同样可以编写用于计算分位数的R语言函数Tgamma(),代码及函数说明见附录。

(3)NP近似在中心极限定理的基础上对尾部的估计做了修正,从而弥补了中心极限定理往往低估尾部概率的缺点。平移伽马近似的前提是建立在保险中很多风险的分布都较接近于伽马分布的前提之下,如果要近似的风险与伽马分布的吻合度较高,那么效果较好,反之则NP近似更佳。

1.8

设S是复合泊松分布,索赔次数服从参数为 $\lambda = 12$ 的泊松分布,索赔额服从U(0,1),试用中心极限定理、平移伽马近似法和NP 近似法来计算概率Pr[S < 10].

解: 易知此时S服从复合泊松分布,故由复合泊松分布的性质

$$E[S] = E[N]E[X]$$

$$= 12 \cdot \frac{1}{2}$$

$$= 6$$

$$Var[S] = E[N]Var[X] + E[X]^{2}Var[N]$$

$$= 12 \cdot \frac{1}{12} + \frac{1}{4} \cdot 12$$

$$= 4$$

$$\gamma_{S} = \frac{E[X^{3}]}{E[X^{2}]^{\frac{3}{2}}\sqrt{\lambda}}$$

$$= \frac{0.25}{\sqrt{\frac{1}{27} \cdot 12}}$$

$$= 0.3$$

故由中心极限定理

$$Pr[S < 10] = Pr[\frac{S - E[S]}{\sigma} < \frac{10 - 6}{\sqrt{4}}]$$

= $\Phi(2)$
= 0.9772499

所以若采用中心极限定理近似、则Pr[S<10]=0.9772。

由于之前计算已得出均值、方差与偏度、故可以编写用于计算平移伽马近似概率的R语言函数pTgamma(),代码及说明见附录。输入如下代码

- > source("pTgamma.R")
- > P<-pTgamma(6,2,0.3,10);P

计算得

[1] 0.9698043

所以若采用平移伽马近似,则Pr[S<10]=0.9698。 同理对于NP近似,编写用于计算NP近似概率的R语言函数pNP(),代码及说明见附录。输入如下代码

- > source("pNP.R")
- > P<-pNP(6,2,0.3,10);P

计算得

[1] 0.9695584

所以若采用NP近似,则Pr[S < 10] = 0.9696。 综上所述,近似结果可如下表所示:

	表 2: 近似	计算结果		
近似方法	中心极限定理	平移伽马近似	NP近似	
近似值	0.9772	0.9698	0.9696	
'				

12

2 SOA questions

2.1

The aggregate loss S is distributed as a compound binomial distribution, where the primary distribution is B(9,0.2). Claim severity X has pdf: $f_X(1) = 0.4$, $f_X(2) = 0.4$, and $f_X(3) = 0.2$. Calculate Pr[S < 4].

Solution:In this question, we can use the Panjer's recursion. And in order to calculate the answer more quickly, we use R function that is called Panjer(). Input this code:

source("Panjer.R")
PP<-Panjer(N=3,c(9,0.2),c(0,0.4,0.4,0.2),s=3);
PP\$P;</pre>

The result:

[1] 0.5924371

So the value of Pr[S < 4] is 0.5924

2.2

Aggregate claim S can only take positive integer values. If $E[(S-2)_+] = \frac{1}{6}$, $E[(S-3)_+] = 0$, and $f_S(1) = \frac{1}{2}$, calculate the mean of S.

Solution: We can see that

$$E[(S-2)_{+}] = \sum_{s>2} (s-2)f_{S}(s)$$

$$= \frac{1}{6}$$

$$E[(S-3)_{+}] = \sum_{s>3} (s-3)f_{S}(s)$$

$$= 0$$

So we can draw a conclusion that if s > 3, $f_S(s) = 0$. And the value of $f_S(3) = \frac{1}{6}$. Then because of $f_S(1) + f_S(2) + f_S(3) = 1$, we can see that

$$f_S(2) = \frac{1}{3}$$
.

Above all, the mean of S is

$$E[S] = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{3} + 3 \cdot \frac{1}{6}$$
$$= \frac{5}{3}$$

2.3

You are given that $E[(S-30)_+]=9$ and $E[(S-20)_+]=12$,and the only possible aggregate claim in (20,30] is 22, with $f_S(22)=0.1$. Calculate $F_S(20)$.

Solution: As a matter of fact, we only need calculate the value of Pr[S > 20]. And this question indicates that $Pr[S > 20] = f_S(22) + Pr[S > 30]$. Then we calculate the value of Pr[S > 30].

$$\sum_{s>30} (s-20)f_S(s) = 12-2.2$$

$$= 9.8$$

$$\sum_{s>30} (s-30)f_S(s) = 9$$

Then

$$\sum_{s>30} 10 f_S(s) = 0.8$$

Therefore,

$$Pr[S > 30] = \sum_{s>30} f_S(s)$$

= 0.08

So the value of $Pr[S \le 20]$ is 1 - Pr[S > 20] = 1 - 0.1 - 0.08 = 0.82.

2.4

Aggregate losses follow a compound Poisson distribution with parameter $\lambda=3$. Individual losses take values 1,2,3, and 4 with probabilities 0.4,0.3,0.2, and 0.1, respectively. Calculate the probability that the aggregate loss does not exceed 3.

Solution: This question is similar with the question 2.1, so we also can use the R function named Panjer(). We should input this code as below:

```
source("Panjer.R")
PP<-Panjer(N=1,3,c(0,0.4,0.3,0.2,0.1),s=3);
PP$P;</pre>
```

the output

[1] 0.2881676

So the value of $Pr[S \leq 3] = 0.2882$

2.5

Aggregate losses follow a compound distribution. The claim frequency has mean 100 and standard deviation 25. The claim severity has mean 20,000 and standard deviation 5,000. Determine the normal approximation of the probability that the aggregate loss exceeds 150% of the expected loss.

Solution:

$$E[S] = E[N]E[X]$$

$$= 2 \cdot 10^{6}$$

$$Var[S] = E[Var[S|N]] + Var[E[S|N]]$$

$$= E[N] \cdot Var[X] + Var[N] \cdot E[X]^{2}$$

$$= 100 \cdot 5000^{2} + 25^{2} \cdot 20000^{2}$$

So

$$Pr[S < 1.5E[S]] = Pr[\frac{S - E[S]}{\sqrt{Var[S]}} < \frac{1.5E[S] - E[S]}{\sqrt{Var[S]}}]$$

$$\approx \Phi(2)$$
= 0.97725

And

$$Pr[S \ge 1.5E[S]] = 1 - Pr[S < 1.5E[S]]$$

= 1 - 0.97725
= 0.02275

A Panjer递推代码

```
Panjer<-function(N=1,n=c(3,0.5),X=c(0,0.5,0.5),s=1){
#N为理赔次数服从的分布
#可选参数为1,2,3,
#分别对应泊松分布、负二项分布、二项分布
#n为对应参数集
#X为单笔理赔额服从的分布
#这里X只考虑离散分布,即要求输入一向量
#s缺省值为0
library(pracma);
if(!is.numeric(n)){
  stop("please input the parameter!")
}
X<-abs(X);</pre>
 if(sum(X)!=1){
   stop("X is not a pdf!")
 }
 if(N!=1&&N!=2&&N!=3){
   stop("N is wrong!")
 }
 s<-as.integer(s);
  if(s<1){
   stop("s must be larger than 1!")
R=zeros(1,(s+1));
r<-length(X);
   if(N==1){
    lambda=n[1];
    NO=dpois(0,lambda);
     g1<-function(t){</pre>
       exp(lambda*(ext(t)-1));
     }
    a=0;
```

```
b=lambda;
  }
  else if(N==2){
   if(length(n)<2){
     stop("parameters are wrong!")
    r=n[1];
    p=n[2];
    NO=dnbinom(0,r,p);
     g1<-function(t){
       (p/(1-(1-p)*exp(t)))^r;
     }
    a=1-p;
    b=a*(r-1);
  }
  else if(N==3){
   if(length(n)<2){
     stop("parameters are wrong!")
   }
    n1=n[1];
    p=n[2];
    NO=dbinom(0,n1,p);
    g1<-function(t){
       (1-p+p*exp(t))^n1;
     =p/(p-1);
     =-a*(n1+1);
if(X[1]==0){
   R[1]=NO;
  }else{
   R[1]=g1(log(X[1]));
}
```

```
for(i in 2:(s+1)){
   t<-min(i-1,r-1);
   S<-zeros(1,t);</pre>
   for(h in 1:t){
      S[h]=(a+b*h/(i-1))*X[h+1]*R[i-h];
   }
   R[i]=(1/(1-a*X[1]))*sum(S);
 }
   P<-sum(R);</pre>
   output<-list(Pi=R[s+1],P=P);</pre>
}
```

平移伽马近似代

```
#计算分位数
Tgamma <- function (mu, sigma, gamma, alpha)
   if(gamma<=0){</pre>
     stop("gamma is wrong!")
   }
    alpha0<-4/(gamma^2)
    beta <- 2/(gamma * sigma)
    x0<-mu-2*sigma/gamma;
    P<-x0+qgamma(alpha,alpha0,beta);
#计算近似概率
pTgamma<-function(mu,sigma,gamma,p){
   if(gamma<=0){
     stop("gamma is wrong!")
   }
    alpha0<-4/(gamma^2);
    beta<-2/(gamma*sigma);</pre>
    x0<-mu-2*sigma/gamma;
```

C NP近似代码 19

```
P<-pgamma(p-x0,alpha0,beta);
}</pre>
```

C NP近似代码

```
#计算分位数
NP<-function(mu,sigma,gamma,alpha){</pre>
  if(gamma==0){
    stop("gamma is wrong!")
  }
  if(alpha>1&alpha<0){
    stop("alpha is wrong!")
  }
    z<-qnorm(1-alpha,0,1);
    x<-((z+3/gamma)^2-1-(9/(gamma^2)))*gamma/6;
    P<-x*sigma+mu;
}
#计算近似概率
pNP<-function(mu, sigma, gamma, p)
  if(gamma==0){
    stop("gamma is wrong!
  }
    x<-(p-mu)/sigma;
    s<-sqrt((9/(gamma^2))+(6*x/gamma)+1)-(3/gamma);
    P<-pnorm(s,0,1);
```

D actuar包简介

actuar包中收录了大部分较为常用的精算方面的函数,函数主要可分为 三类:

D.1 分布

很多只在保险精算领域常用的分布及其相关函数并未在在R自带的stats包中收录,如pareto分布等,actuar包弥补了这一缺陷,并且用法都与stats包中的对应函数一致²,该包添加的分布见下表³: 需要说明的是,

表 3: act	ual包中的分布
函数	说明
genbeta	广义β分布
genpareto	广义pareto分布
invburr	逆Burr分布
invexp	逆指数分布
invgamma	逆伽马分布
invpareto	逆pareto分布
invparalogis	
invtrgamma	变形逆伽马分布
invweibull	逆威布尔分布
lgamma	对数伽马分布
llogis	
paralogis	
pareto	pareto分布
pareto1	单参数pareto分布
trbeta	变形β分布
trgamma	变形伽马分布

对于上述分布,该包中的函数添加两项功能,分别是计算K阶矩(前面加m)、存在保单限额时的k阶矩(前面加lev)。

D.2 常用量

该包的另一功能是可用于计算一些常用的量,整理如下表:

 $^{^{2}}$ 即 $_{p,d,q,r+}$ 对应分布,分别对应了分布函数、密度函数、分位数、模拟

³表格中函数按字母顺序排列,部分分布也可在VGAM包中找到

表 4: actual包中的分布

	大 4. detdai 已 自力力 中				
adjCoef	计算调节系数				
aggregateDist	总索赔额(提供包括递推、卷积、正态近似、NP近似等五种方法)				
cm	信度模型				
$\overline{\text{CTE}}$	条件尾部期望				
quantile	总索赔额分位数计算				
ruin	计算破产概率				
VaR	计算VaR				

D.3 变换

该包还提供了求矩母函数值的功能等,用法是是在分布后面加Supp 4

D.4 数据集

该包还提供了包括dental等数据集

⁴只针对精算中常用的分布