ACSC/STAT 4703, Actuarial Models II
Fall 2020
Toby Kenney
Homework Sheet 1
Model Solutions

Basic Questions

1. *Aggregate payments have a compund distribution. The frequency distribution is negative binomial with $r = 2$ and $\beta = 1.9$. The severity distribution is a gamma distribution with $\alpha = 0.7$ and $\theta = 18000$. Use a Pareto approximation to aggregate payments to estimate the probability that aggregate payments are more than \$350,000.*

   The frequency distribution has mean $2 \times 1.9 = 3.8$, and variance $2 \times 1.9 \times 2.9 = 11.02$. The severity distribution has mean $0.7 \times 18000 = 12,600$ and variance $0.7 \times 18000^2 = 226,800,000$. The mean of the aggregate loss distribution is therefore $3.8 \times 12600 = 47,880$ and the variance is $3.8 \times 226800000 + 11.02 \times 12600^2 = 2611375200$.

   The Pareto distribution with this variance has

$$\begin{aligned}
\frac{\theta}{\alpha-1} &= 47880 \\
\frac{\theta^2\alpha}{(\alpha-1)^2(\alpha-2)} &= 2611375200 \\
\frac{\alpha}{\alpha-2} &= \frac{2611375200}{47880^2} \\
&= 1.13909774436 \\
0.13909774436\alpha &= 2 \times 1.13909774436 \\
\alpha &= 2 \times 1.13909774436/0.13909774436 \\
&= 16.3783783785 \\
\theta &= 47880 \times 15.3783783785 \\
&= 736316.756763
\end{aligned}$$

   Therefore the probability that the aggregate payment exceeds \$350,000 is $\left(\frac{1}{1+\frac{350000}{736316.756763}}\right)^{16.3783783785} = 0.00171327$.

2. *Loss amounts follow a Pareto distribution with $\alpha = 4$ and $\theta = 120,000$. The distribution of the number of losses is given in the following table:*

Number of Losses	*Probability*
0	*0.47*
1	*0.11*
2	*0.27*
3	*0.15*

Assume all losses are independent and independent of the number of losses. The insurance company buys excess-of-loss reinsurance on the part of the loss above $200,000. Calculate the expected payment for this excess-of-loss reinsurance.

For one loss, the expected reinsurance payment is given by

$$\begin{aligned}
\mathbb{E}((X-200000)_+) &= \int_{200000}^{\infty} \left(\frac{1}{1+\frac{x}{120000}}\right)^4 dx \\
&= 120000 \int_{\frac{200000}{120000}+1}^{\infty} u^{-4}\, du \\
&= 120000 \left[-\frac{u^{-3}}{3}\right]_{\frac{5}{3}+1}^{\infty} \\
&= 40000 \left(\frac{3+5}{3}\right)^{-3} \\
&= 40000 \left(\frac{3}{8}\right)^{3} \\
&= 2109.375
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}((X-a)_+) &= \int_{a}^{\infty} \left(\frac{1}{1+\frac{x}{\theta}}\right)^\alpha dx \\
&= \theta \int_{\frac{a}{\theta}+1}^{\infty} u^{-\alpha}\, du \\
&= \theta \left[-\frac{u^{-3}}{3}\right]_{\frac{5}{3}+1}^{\infty} \\
&= \frac{\theta}{\alpha-1} \left(\frac{\theta+a}{\theta}\right)^{1-\alpha} \\
&= \frac{\theta}{\alpha-1} \left(\frac{\theta}{\theta+a}\right)^{\alpha-1} \\
&= \frac{\theta^\alpha}{(\alpha-1)(\theta+a)^{\alpha-1}}
\end{aligned}$$

For two losses, the expected reinsurance payment is given by

$$\begin{aligned}
&\mathbb{E}((X+Y-200000)_+)\\
&=\mathbb{E}\left(\mathbb{E}((Y-(200000-X))_+|X)\right)\\
&=\mathbb{E}_{X<200000}\left(\int_{200000-X}^{\infty}\left(\frac{1}{1+\frac{y}{120000}}\right)^4\,dy\right)+\mathbb{E}_{X\geqslant 200000}\left(X-200000+\mathbb{E}(y)\right)\\
&=120000\mathbb{E}_{X<200000}\left(\int_{\frac{8}{3}-\frac{X}{120000}}^{\infty}u^{-4}\,du\middle|X\right)+2109.375+40000P(X>200000)\\
&=40000\mathbb{E}_{X<200000}\left(\left(\frac{8}{3}-\frac{X}{120000}\right)^{-3}\right)+2109.375+40000\left(\frac{120000}{120000+200000}\right)^4\\
&=40000\int_0^{200000}\left(\left(\frac{8}{3}-\frac{X}{120000}\right)^{-3}\left(\frac{4\times 120000^4}{(120000+x)^5}\right)\right)\,dx+2109.375+40000\left(\frac{120000}{120000+200000}\right)^4\\
&=40000\int_0^{200000}\frac{4\times 120000^7}{(320000-x)^3(120000+x)^5}\,dx+2109.375+40000\left(\frac{120000}{120000+200000}\right)^4
\end{aligned}$$

It is possible to solve this analytically using partial fractions

$$\begin{aligned}
&\mathbb{E}((X+Y-a)_+)\\
&=\mathbb{E}\left(\mathbb{E}((Y-(a-X))_+|X)\right)\\
&=\mathbb{E}_{X<a}\left(\int_{a-X}^{\infty}\left(\frac{1}{1+\frac{y}{\theta}}\right)^\alpha\,dy\right)+\mathbb{E}_{X\geqslant a}\left(X-a+\mathbb{E}(y)\right)\\
&=\theta\mathbb{E}_{X<a}\left(\int_{\frac{a+\theta}{\theta}-\frac{X}{\theta}}^{\infty}u^{-\alpha}\,du\middle|X\right)+\mathbb{E}((X-a)_+)+\left(\frac{a}{\theta+a}\right)^\alpha\frac{\theta}{\alpha-1}\\
&=\frac{\theta}{\alpha-1}\mathbb{E}_{X<a}\left(\left(\frac{\theta}{\theta+a-X}\right)^{\alpha-1}\right)+\frac{\theta^\alpha}{(\alpha-1)(\theta+a)^{\alpha-1}}+\left(\frac{a}{\theta+a}\right)^\alpha\frac{\theta}{\alpha-1}\\
&=\frac{\theta}{\alpha-1}\int_0^a\left(\left(\frac{\theta}{\theta+a-x}\right)^{\alpha-1}\left(\frac{\alpha\times\theta^\alpha}{(\theta+x)^{\alpha+1}}\right)\right)\,dx+\frac{\theta^\alpha(\theta+a)+\theta a^\alpha}{(\alpha-1)(\theta+a)^\alpha}\\
&=\frac{\alpha\theta^{2\alpha}}{\alpha-1}\int_0^a\frac{1}{(\theta+a-x)^{\alpha-1}(\theta+x)^{\alpha+1}}\,dx+\frac{\theta^\alpha(\theta+a)+\theta a^\alpha}{(\alpha-1)(\theta+a)^\alpha}
\end{aligned}$$

We have

$$
\begin{aligned}
&\frac{1}{(a-x)^3(b+x)^5} \\
&= \frac{1}{(a+b)} \left(\frac{1}{(a-x)^3(b+x)^4} + \frac{1}{(a-x)^2(b+x)^5} \right) \\
&= \dots \\
&= \frac{(a+b)^{-5}}{(a-x)^3} + \frac{(a+b)^{-5}}{(a-x)^2} \left(\frac{1}{(b+x)} + \frac{a+b}{(b+x)^2} + \frac{(a+b)^2}{(b+x)^3} + \frac{(a+b)^3}{(b+x)^4} + \frac{(a+b)^4}{(b+x)^5} \right) \\
&= \frac{(a+b)^{-5}}{(a-x)^3} + \frac{5(a+b)^{-6}}{(a-x)^2} + \frac{(a+b)^{-6}}{a-x} \left(\frac{5}{(b+x)} + \frac{4(a+b)}{(b+x)^2} + \frac{3(a+b)^2}{(b+x)^3} + \frac{2(a+b)^3}{(b+x)^4} + \frac{(a+b)^4}{(b+x)^5} \right) \\
&= \frac{(a+b)^{-5}}{(a-x)^3} + \frac{5(a+b)^{-6}}{(a-x)^2} + \frac{15(a+b)^{-7}}{a-x} + \frac{15(a+b)^{-7}}{(b+x)} + \frac{10(a+b)^{-6}}{(b+x)^2} + \frac{6(a+b)^{-5}}{(b+x)^3} + \frac{3(a+b)^{-4}}{(b+x)^4} \\
&\qquad + \frac{(a+b)^{-3}}{(b+x)^5}
\end{aligned}
$$

Hence

$$
\begin{aligned}
&\int_0^c \frac{1}{(a-x)^3(b+x)}\, dx \\
&= \sum_{n=1}^{3} \binom{7-n}{4} \int_0^c \frac{(a+b)^{n-8}}{(a-x)^n}\, dx + \sum_{n=1}^{5} \binom{7-n}{2} \int_0^c \frac{(a+b)^{n-8}}{(b+x)^n}\, dx \\
&= \sum_{n=2}^{3} \binom{7-n}{4} \left[\frac{(a+b)^{n-8}}{(n-1)(a-x)^{n-1}} \right]_0^c + \sum_{n=2}^{5} \binom{7-n}{2} \left[-\frac{(a+b)^{n-8}}{(n-1)(b+x)^{n-1}} \right]_0^c \\
&\qquad + 15(a+b)^{-7} \left(\log(a) - \log(a-c)\right) + 15(a+b)^{-7} \left(\log(b+c) - \log(b)\right) \\
&= \sum_{n=2}^{3} \binom{7-n}{4} \frac{(a+b)^{n-8}}{(n-1)} \left((a-c)^{1-n} - a^{1-n}\right) + \sum_{n=2}^{5} \binom{7-n}{2} \frac{(a+b)^{n-8}}{(n-1)} \left(b^{1-n} - (b+c)^{1-n}\right) \\
&\qquad + 15(a+b)^{-7} \log\left(\frac{a(b+c)}{b(a-c)} \right) \\
&= \sum_{n=2}^{5} \frac{(a+b)^{n-8}}{(n-1)} \left(\binom{7-n}{4} \left((a-c)^{1-n} - a^{1-n}\right) + \binom{7-n}{2} \left(b^{1-n} - (b+c)^{1-n}\right) \right) + 15(a+b)^{-7} \log \Bigg(\frac{a(}{b(}
\end{aligned}
$$

Plugging this into our formula gives

$$
\begin{aligned}
&\int_0^a \frac{1}{(\theta+a-x)^3(\theta+x)^5}\,dx\\
&=\sum_{n=2}^{5}\frac{(2\theta+a)^{n-8}}{(n-1)}\left(\binom{7-n}{4}\left(\theta^{1-n}-(\theta+a)^{1-n}\right)+\binom{7-n}{2}\left(\theta^{1-n}-(\theta+a)^{1-n}\right)\right)+15(2\theta+a)^{-7}\log\Bigg(\\
&=\sum_{n=2}^{5}\frac{(2\theta+a)^{n-8}}{(n-1)}\left(\binom{7-n}{4}+\binom{7-n}{2}\right)\left(\theta^{1-n}-(\theta+a)^{1-n}\right)+15(2\theta+a)^{-7}\log\left(\frac{(\theta+a)^2}{\theta^2}\right)
\end{aligned}
$$

$$
\begin{aligned}
&\mathbb{E}((X+Y-a)_+)\\
&=\frac{4\theta^8}{3}\left(\sum_{n=2}^{5}\frac{(2\theta+a)^{n-8}}{(n-1)}\left(\binom{7-n}{4}+\binom{7-n}{2}\right)\left(\theta^{1-n}-(\theta+a)^{1-n}\right)+15(2\theta+a)^{-7}\log\left(\frac{(\theta+a)^2}{\theta^2}\right)\right)\\
&\qquad+\frac{\theta^4(\theta+a)+\theta a^4}{3(\theta+a)^4}\\
&=\frac{4\theta^8}{3(2\theta+a)^7}\left(30\log\left(\frac{\theta+a}{\theta}\right)+15\left(\frac{2\theta+a}{\theta}-\frac{2\theta+a}{\theta+a}\right)+\frac{7}{2}\left(\frac{(2\theta+a)^2}{\theta^2}-\frac{(2\theta+a)^2}{(\theta+a)^2}\right)\right.\\
&\qquad\left.+\left(\frac{(2\theta+a)^3}{\theta^3}-\frac{(2\theta+a)^3}{(\theta+a)^3}\right)+\frac{1}{4}\left(\frac{(2\theta+a)^4}{\theta^4}-\frac{(2\theta+a)^4}{(\theta+a)^4}\right)\right)+\frac{\theta^4(\theta+a)+\theta a^4}{3(\theta+a)^4}
\end{aligned}
$$

For 3 losses, the expected payment is given by

$$
\begin{aligned}
\mathbb{E}((X+Y+Z-a)_+)&=\mathbb{E}(\mathbb{E}(X+Y-(a-Z))_+)|Z)\\
&=\int_0^a\frac{4\theta^4}{(\theta+t)^5}\mathbb{E}((X+Y-(a-t))_+)\,dt+P(Z>a)\mathbb{E}(X+Y)+\mathbb{E}((Z-a)_+)
\end{aligned}
$$

We numerically integrate this to get the expected payment

```
  #Crude numerical integration, h=1.
  #Could use trapezium rule or Simpson's rule for more accuracy.
ft<-4*theta^4/(theta+(0:199999))^5
a<-200000:1
#Simplify algebra
c<-(2*theta+a)/theta
d<-(2*theta+a)/(theta+a)
EXYa<-4*theta^8/3/(2*theta+a)^7*(30*log((theta+a)/theta)+15*(c-d)#
+3.5*(c^2-d^2)+c^3-d^3+(c^4-d^4)/4)#
+(theta^4*(theta+a)+theta*a^4)/3/(theta+a)^4
sum(ft*EXYa)+2109.375+(theta/(theta+200000))^4*2*theta/3
```

This allows us to calculate the following table:

Number of Losses	Probability	Conditional Expected reinsurance payment	Expected reinsurance payment
0	0.47	0	0
1	0.11	2109.375	232.031
2	0.27	11718.52	3164.000
3	0.15	16904.735	2535.710

Thus the total expected payment on the excess-of-loss reinsurance is $232.031 + 3164.000 + 2535.710 = \$5,931.74$.

3. *An insurance company models loss frequency as binomial with* $n = 88$, $p = 0.11$, *and loss severity as exponential with* $\theta = 20,000$. *Calculate the expected aggregate payments if there is a policy limit of $80,000 and a deductible of $15,000 applied to each claim.*

   With a policy limit of $80,000 and a deductible of $15,000, the expected payment per loss is $20000\left(e^{-\frac{15000}{20000}} - e^{-\frac{95000}{20000}}\right) = 9274.29715076$. The expected number of losses is $88 \times 0.11 = 9.68$. The expected aggregate payment is therefore $9.68 \times 9274.29715076 = \$89,775.20$.

4. *Claim frequency follows a negative binomial distribution with* $r = 2$ *and* $\beta = 4.1$. *Claim severity (in thousands) has the following distribution:*

Severity	*Probability*
1	*0.4*
2	*0.39*
3	*0.14*
4	*0.05*
5 or more	*0.02*

   Use the recursive method to calculate the exact probability that aggregate claims are at least 5.

   Since the negative binomial distribution is from the $(a, b, 0)$ class with $a = \frac{\beta}{1+\beta} = \frac{4.1}{5.1} = 0.803921568627$ and $b = \frac{(r-1)\beta}{1+\beta} = 0.803921568627$ and the severity distribution is zero-truncated, the recursive formula gives

$$f_S(n) = \sum_{i=1}^{n}\left(a + \frac{bi}{n}\right) f_S(n-i) f_X(i) = \sum_{i=1}^{n} 0.803921568627\left(1 + \frac{i}{n}\right) f_S(n-i) f_X(i)$$

   Since the severity distribution is zero-truncated, the probability that aggregate losses are zero is the probability that there are no losses, which is $f_S(0) = \left(\frac{1}{1+\beta}\right)^r = \frac{1}{5.1^2} = 0.0384467512495$. Applying the recurrence relation gives

$$f_S(1) = 0.80392 \times 2 \times 0.038447 \times 0.4 = 0.0247265380585$$

$$f_S(2) = 0.80392\left(\frac{3}{2} \times 0.024727 \times 0.4 + 2 \times 0.038447 \times 0.39\right) = 0.0360352929646$$

$$f_S(3) = 0.80392\left(\frac{4}{3} \times 0.036035 \times 0.4 + \frac{5}{3} \times 0.024727 \times 0.39 + 2 \times 0.038447 \times 0.14\right) = 0.0370255428058$$

$$f_S(4) = 0.80392\left(\frac{5}{4} \times 0.037026 \times 0.4 + \frac{3}{2} \times 0.036035 \times 0.39 + \frac{7}{4} \times 0.024727 \times 0.14 + 2 \times 0.038447 \times 0.05\right) = 0.0397909781215$$

Therefore, the probability that the aggregate loss is at least 5 is $1 - 0.0397909781215 - 0.0370255428058 - 0.0360352929646 - 0.0247265380585 - 0.0384467512495 = 0.823974896801$.

5. *Use an arithmetic distribution ($h = 1$) to approximate a Pareto distribution with $\alpha = 4$ and $\theta = 60$.*

   (a) Using the method of rounding, calculate the mean of the arithmetic approximation. [You can evaluate this numerically: use 10,000 terms in the sum.]

   Let X be the Pareto distribution and let Y be the arithmetic approximation. We have that

$$P(Y \geqslant i) = P\left(X \geqslant i - \frac{1}{2}\right) = \left(\frac{1}{1 + \frac{i-\frac{1}{2}}{60}}\right)^4$$

   The mean of the arithmetic approximation is therefore given by

$$\sum_{i=1}^{\infty} \left(\frac{120}{119 + 2i}\right)^4$$

   Numerically, we evaluate this as 19.99722.

   (b) Using the method of local moment matching, matching 1 moment on each interval, estimate the probability that the value is larger than 18.5.

   We have that $P(0 \leqslant Y \leqslant 17.5) = P(0 \leqslant X \leqslant 17.5) = 1 - \left(\frac{60}{77.5}\right)^4 = 0.640748829751$. We need to calculate $P(Y = 18)$. We have that

$$
\begin{aligned}
p_{18}+p_{19} &= P(17.5<Y<19.5) = \left(\frac{60}{77.5}\right)^4 - \left(\frac{60}{79.5}\right)^4 \\
&= 0.034809605787 \\
18p_{18}+19p_{19} &= \int_{17.5}^{19.5} x\frac{\alpha\theta^\alpha}{(\theta+x)^{\alpha+1}}\,dx \\
&= 4\times 60^4 \int_{77.5}^{79.5} (u-60)u^{-5}\,dx \\
&= 4\times 60^4 \left[15u^{-4} - \frac{u^{-3}}{3}\right]_{77.5}^{79.5} \\
&= 4\times 60^4 \left(\frac{15}{79.5^4} - \frac{15}{77.5^4} + \frac{1}{3\times 77.5^3} - \frac{1}{3\times 79.5^3}\right) \\
&= 0.643238745556 \\
p_{18} &= 19\times 0.034809605787 - 0.643238745556 \\
&= 0.018143764397
\end{aligned}
$$

So the probability that $Y > 18.5$ is $0.359251170249 - 0.018143764397 = 0.341107405852$.

Standard Questions

6. *The number of claims an insurance company receives follows a negative binomial distribution with $r = 64$ and $\beta = 37$. Claim severity follows a negative binomial distribution with $r = 14$ and $\beta = 1.4$. Calculate the probability that aggregate losses exceed \$32,000.*

   (a) Starting the recurrence 6 standard deviations below the mean [You need to calculate the recurrence up to $f_s(60,000)$.]

   The frequency distribution is from the $(a, b, 0)$ class with $a = \frac{\beta}{\beta+1} = \frac{37}{38}$ and $b = \frac{(r-1)\beta}{\beta+1} = \frac{2331}{38}$. The recurrence is

$$
f_s(n) = \frac{\sum_{i=1}^{n} \frac{37}{38}\left(1+\frac{63i}{n}\right) f_X(i) f_S(n-i)}{1-\frac{37}{38}\times\frac{1}{2.4^{14}}} = \frac{\sum_{i=1}^{n} \left(1+\frac{63i}{n}\right) f_X(i) f_S(n-i)}{\frac{38}{37}-\frac{1}{2.4^{14}}}
$$

   The mean of the aggregate loss distribution is $64 \times 37 \times 14 \times 1.4 = 46412.8$ and the variance is $64 \times 37 \times 14 \times 1.4 \times 2.4 + 64 \times 37 \times 38 \times (14 \times 1.4)^2 = 34679644.16$ so 6 standard deviations below the mean is $46412.8 - 6\sqrt{34679644.16} = 11079.1448004$. We will therefore start the recurrence at $n = 11079$.

```
maxf<-100000
f<-rep(0,maxf)
start<-11079
f[start]<-1

#Calculate severity distribution once, not inside the loop.
fX<-choose((1:maxf)+13,13)*(7/12)^(1:maxf)*(5/12)^14

#the denominator in the recurrence is also constant.
denominator<-(38/37-1/2.4^14)
for(i in seq_len(maxf-start)){
  x<-i+start
  f[x]=sum((1+63*(1:i)/x)*f[x-(1:i)]*fX[1:i])/denominator
  #Vector operations are faster in R than loops
}
```

We then standardise f by dividing by its sum and evaluate the probability:

```
f<-f/sum(f)
sum(f[32001:maxf])
```

This gives the value $P(X > 32000) = 0.9966465$.

(b) Using a suitable convolution.

This distribution is given as a sum of 8 i.i.d. random variables whose distribution is an aggregate loss distribution with frequency following a negative binomial distribution with $r = 8$ and $\beta = 37$ and severity following a negative binomial distribution with $r = 14$ and $\beta = 1.4$. For this aggregate loss distibution, we calculate $f(0) = P_S(0) = P_F(P_X(0)) = P_F(f_X(0)) = (1 + \beta - \beta f_X(0))^{-r} = (1 + 37 - \frac{37}{2.4^{14}})^{-8}$. The recurrence is given by

$$f_s(n) = \frac{\sum_{i=1}^{n} \frac{37}{38}\left(1 + \frac{7i}{n}\right) f_X(i) f_S(n-i)}{1 - \frac{37}{38} \times \frac{1}{2.4^{14}}} = \frac{\sum_{i=1}^{n} \left(1 + \frac{7i}{n}\right) f_X(i) f_S(n-i)}{\frac{38}{37} - \frac{1}{2.4^{14}}}$$

We use this to compute the distribution of this reduced loss

```
g<-rep(0,20001)
g[1]=(1+37-37/(2.4^14))^(-8)

fX<-choose((1:50000)+13,13)*(7/12)^(1:50000)*(5/12)^14

for(x in 2:20001){
    y<-seq_len(x-1)
    temp<-sum((1+7*y/(x-1))*fX[y]*g[x-y])
    g[x]<-temp/(38/37-2.4^(-14))
}
```

Having computed $f_S(x)$ for $x = 0, \ldots, 20000$, we convolve this with itself 3 times to get the aggregate loss distribution

```
  ConvolveSelf<-function(n){
  l<-length(n)
  convolution<-vector("numeric",2*l)
  for(i in seq_len(l)){
    convolution[i]<-sum(n[1:i]*n[i:1])
  }
  for(i in 1:(length(n))){
    convolution[2*l+1-i]<-sum(n[l+1-(1:i)]*n[l+1-(i:1)])
  }
  return(convolution)
}

g2<-ConvolveSelf(g)
g4<-ConvolveSelf(g2)
g8<-ConvolveSelf(g4)

1-sum(g8[1:32001])
```

This gives the probability value 0.9966465.