ACSC/STAT 4703, Actuarial Models II

FALL 2023
Toby Kenney

Homework Sheet 6

Model Solutions

Basic Questions

1. *An insurance company has the following previous data on aggregate claims:*

Policyholder	*Year 1*	*Year 2*	*Year 3*	*Year 4*	*Year 5*	*Mean*	*Variance*
1	*76.04*	*5.33*	*872.93*	*0.00*	*0.00*	*190.860*	*146419.894*
2	*5.55*	*64.77*	*1421.35*	*542.52*	*2140.48*	*834.934*	*853717.609*
3	*33.12*	*10.10*	*108.50*	*0.00*	*173.90*	*65.124*	*5501.682*
4	*12.78*	*494.70*	*1578.16*	*87.65*	*32.21*	*441.100*	*442838.590*
5	*1728.19*	*0.00*	*2898.65*	*1570.46*	*24.30*	*1244.320*	*1528521.572*

Calculate the Bühlmann credibility premium for each policyholder in Year 6.

The book premium is the average of the average claims for each individual, i.e. $\frac{190.860+834.934+65.124+441.100+1244.320}{5} = 555.2676$ The estimated EPV is the average of the variances for the individuals, that is

$$\frac{146419.894 + 853717.609 + 5501.682 + 442838.590 + 1528521.572}{5} = 595399.8694$$

The variance of the observed means is

$$\frac{(190.860 - 555.2676)^2 + (834.934 - 555.2676)^2 + (65.124 - 555.2676)^2 + (441.100 - 555.2676)^2 + (1244.320 - 555.2676)^2}{4} = 234768.59842$$

The part of this due to process variance is $\frac{595399.8694}{5} = 119079.97388$. Therefore, the estimated VHM is $234768.598421 - 119079.97388 = 115688.624541$.

The credibility of 5 years of experience is therefore

$$Z = \frac{5}{5 + \frac{595399.8694}{115688.624541}} = 0.492777250957$$

The premiums are therefore:

Policyholder	Premium
1	$0.492777250957 \times 190.860 + 0.507222749043 \times 555.2676 = 375.70$
2	$0.492777250957 \times 834.934 + 0.507222749043 \times 555.2676 = 693.08$
3	$0.492777250957 \times 65.124 + 0.507222749043 \times 555.2676 = 313.74$
4	$0.492777250957 \times 441.100 + 0.507222749043 \times 555.2676 = 499.01$
5	$0.492777250957 \times 1244.320 + 0.507222749043 \times 555.2676 = 894.82$

2. *The file* `HW6_data.txt` *contains aggregate claim data from 100 policyholders over the past 10 years. Use this data to estimate the book premium and the credibility of 10 years' experience.*

   We use the following code:

```
claims<-read.table("HW6_data.txt")
book.premium<-mean(rowMeans(claims))
EPV<-mean(apply(claims,1,var))
VOM<-var(rowMeans(claims))
VHM<-VOM-EPV/10
Z<-10/(10+EPV/VHM)
```

   This gives the book premium as \$189.2995 and the credibility of 10 years' experience as $Z = 0.2566474$.

3. *An insurance company collects the following numbers of claims from five policyholders over a 5-year period.*

Policyholder	*Year 1*	*Year 2*	*Year 3*	*Year 4*	*Year 5*
1	*6*	*5*	*4*	*5*	*5*
2	*7*	*4*	*5*	*2*	*8*
3	*4*	*2*	*3*	*2*	*0*
4	*6*	*4*	*5*	*5*	*7*
5	*3*	*0*	*2*	*1*	*2*

   The company assumes that the number of claims for each policyholder follows a Poisson distribution. Use Bühlmann credibility to estimate the average number of claims for Policyholder 4 in Year 6.

   There were a total of 97 claims from 25 policyholder-years, so the overall average is $\frac{97}{25} = 3.88$. Since the variance and mean are equal for the Poisson distribution, this is also the EPV. The variance of the observed means is

$$\frac{(5-3.88)^2 + (5.2-3.88)^2 + (2.2-3.88)^2 + (5.4-3.88)^2 + (1.6-3.88)^2}{4} = 3.332$$

   Of this, $\frac{3.88}{5} = 0.776$ is due to process variance, so the estimated VHM is $3.332 - 0.776 = 2.556$. The credibility of 5 years of experience is therefore

$\frac{5}{5+\frac{3.88}{2.556}} = 0.767106842737$, so the expected number of claims for Policyholder 4 is $0.767106842737 \times 5.4 + 0.232893157263 \times 3.88 = 5.046$.

Standard Questions

4. *The file* `HW6_data2.txt` *contains aggregate claim data from 100 policyholders over the past 10 years. Some policyholders did not purchase insurance in all years. Use this data to estimate the book premium and the credibility of 10 years' experience.*

   We use the following code:

```
claims<-read.table("HW6_data2.txt")
book.premium<-mean(as.vector(as.matrix(claims)),na.rm=TRUE)
policy.holder.means<-rowMeans(claims,na.rm=TRUE)

### count the number of years of experience for each policyholder
num.exp<-rowSums(!is.na(claims))
PV<-rowSums((claims-policy.holder.means)^2,na.rm=TRUE)/(num.exp-1)
EPV<-sum(num.exp*PV,na.rm=TRUE)/sum(num.exp)
weighted.SSE<-sum((policy.holder.means-book.premium)^2*num.exp)
VHM<-(weighted.SSE-99*EPV)/(sum(num.exp)-sum(num.exp^2)/sum(num.exp))

Z<-10/(10+EPV/VHM)
```

   This gives the book premium as \$335.47 and the credibility of 10 years' experience as $Z = 0.436981$. [We get $\widehat{\text{EPV}} = 3895191$ and $\widehat{\text{VHM}} = 302321$. If we take a credibility-weighted average for the book premium, we get \$336.12.]

5. *Aggregate claims for a given individual policy are modelled as following a Gamma distribution with $\alpha = 4m_i$ and $\theta = R_i$, where R_i is a risk factor for that policyholder that varies between individuals and m_i is the exposure of the individual.*

   From a dataset of 100 policyholders with different exposures, they find that the total aggregate claim is \$98,236 from a total of 722 units of exposure. They also calculate:

$$\sum m_i^2 = 11,604$$
$$\sum m_i X_i^2 = 16,850,495.1375$$

where X_i is the aggregate claims per unit of exposure for Policyholder i (so $\sum m_i X_i = 98236$). Estimate the EPV and VHM from this data.

If the exposure of Policyholder i is m_i, and that policyholder's average claims per unit of exposure is X_i, then the book premium per unit of exposure is $\widehat{4\mathbb{E}(R)} = \frac{\sum m_i X_i}{\sum m_i} = \frac{98236}{722} = 136.060941828$. The variance of a gamma distribution with parameters α and θ is $\alpha\theta^2$, so the EPV per unit of exposure is $\mathbb{E}(4R_i^2) = 4\left(\mathbb{E}(R_i)^2 + \mathrm{Var}(R_i)\right)$. Using the estimate $\frac{\sum m_i(X_i-\mu)^2-(n-1)\widehat{\mathrm{EPV}}}{m-\frac{\sum m_i^2}{m}}$ for the VHM $\mathrm{Var}(4R_i) = 16\,\mathrm{Var}(R_i)$.

We therefore solve

$$
\begin{aligned}
16\,\mathrm{Var}(R_i) &= \frac{\sum m_i(X_i-\mu)^2 - 4(n-1)\left(\left(\frac{\mu}{4}\right)^2 + \mathrm{Var}(R_i)\right)}{m - \frac{\sum m_i^2}{m}} \\
&= \frac{\sum m_i(X_i-\mu)^2 - 396\,(1157.0362432 + \mathrm{Var}(R_i))}{705.927977839} \\
16.560963742\,\mathrm{Var}(R_i) &= \frac{\sum m_i(X_i - 136.060941828)^2 - 458186.352307}{705.927977839} \\
&= \frac{\sum m_i X_i^2 - 272.121883656 m_i X_i + m_i 136.060941828^2 - 458186.352307}{705.927977839} \\
&= \frac{16850495.1375 - 272.121883656 \times 98236 + 722 \times 136.060941828^2 - 458186.352307}{705.927977839} \\
&= 4286.8765636 \\
16\,\mathrm{Var}(R_i) &= \frac{16}{16.560963742} \times 4286.8765636 = 4141.66869067
\end{aligned}
$$

The EPV per unit of exposure is $\frac{136.060941828^2}{4} + \frac{4141.66869067}{4} = 5663.56214545$.