ACSC/STAT 4703, Actuarial Models II

FALL 2024
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	*0.00*	*37.42*	*0.00*	*0.00*	*10.74*	*9.632*	*262.9317*
2	*245.20*	*2033.31*	*462.66*	*94.24*	*1388.56*	*844.794*	*694280.6027*
3	*695.47*	*1655.90*	*1.97*	*106.79*	*40.06*	*500.038*	*497126.5281*
4	*924.94*	*75.32*	*0.13*	*85.97*	*101.15*	*237.502*	*149193.3216*
5	*326.55*	*0.05*	*183.39*	*48.46*	*38.82*	*119.454*	*18193.9676*

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. $\frac{9.632+844.794+500.038+237.502+119.454}{5} = 342.284$ The estimated EPV is the average of the variances for the individuals, that is

$$\frac{262.9317 + 694280.6027 + 497126.5281 + 149193.3216 + 18193.9676}{5} = 271811.47034$$

The variance of the observed means is

$$\frac{(9.632 - 342.284)^2 + (844.794 - 342.284)^2 + (500.038 - 342.284)^2 + (237.502 - 342.284)^2 + (119.454 - 342.284)^2}{4} = 112173.113536$$

The part of this due to process variance is $\frac{271811.47034}{5} = 54362.294068$. Therefore, the estimated VHM is $112173.113536 - 54362.294068 = 57810.819468$.

The credibility of 5 years of experience is therefore

$$Z = \frac{5}{5 + \frac{271811.47034}{57810.819468}} = 0.515371443706$$

The premiums are therefore:

Policyholder	Premium
1	$0.515371443706 \times 9.632 + 0.484628556294 \times 342.284 = \$170.84$
2	$0.515371443706 \times 844.794 + 0.484628556294 \times 342.284 = \$601.26$
3	$0.515371443706 \times 500.038 + 0.484628556294 \times 342.284 = \$423.59$
4	$0.515371443706 \times 237.502 + 0.484628556294 \times 342.284 = \$288.28$
5	$0.515371443706 \times 119.454 + 0.484628556294 \times 342.284 = \$227.44$

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 \$1,165.391 and the credibility of 10 years' experience as $Z = 0.841466$.

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	*7*	*12*	*9*	*8*	*9*
2	*5*	*2*	*2*	*3*	*4*
3	*2*	*8*	*7*	*6*	*4*
4	*11*	*4*	*9*	*6*	*3*
5	*2*	*2*	*2*	*5*	*1*

   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 2 in Year 6.

   There were a total of 133 claims from 25 policyholder-years, so the overall average is $\frac{133}{25} = 5.32$. Since the variance and mean are equal for the Poisson distribution, this is also the EPV. The variance of the observed means is

$$\frac{(9-5.32)^2 + (3.2-5.32)^2 + (5.4-5.32)^2 + (6.6-5.32)^2 + (2.4-5.32)^2}{4} = 7.052$$

   Of this, $\frac{5.32}{5} = 1.064$ is due to process variance, so the estimated VHM is $7.052 - 1.064 = 5.988$. The credibility of 5 years of experience is therefore

$\frac{5}{5+\frac{5.32}{5.988}} = 0.84912081679$, so the expected number of claims for Policyholder 2 is $0.84912081679 \times 3.2 + 0.15087918321 \times 5.32 = 3.51986386841$.

Standard Questions

4. *The file* `HW6_data2.txt` *contains aggregate claim data from 100 policyholders over the past 10 years. The data are assumed to come from a gamma distribution with shape $\alpha = 4$ and scale parameter varying between policyholders. Calculate the credibility of 10 years of experience, and the premium in Year 11 for Policyholder 15.*

   For the gamma distribution with $\alpha = 4$, we have that the mean is 4θ and the variance is $4\theta^2$. Therefore, the EPV is $\mathbb{E}(4\Theta^2) = 4\left(\mathbb{E}(\Theta)^2 + \text{Var}(\Theta)\right)$. The variance of hypothetical means is $\text{Var}(4\Theta) = 16\,\text{Var}(\Theta)$, so the variance of observed means is $16\,\text{Var}(\Theta) + \frac{4\left(\mathbb{E}(\Theta)^2+\text{Var}(\Theta)\right)}{10} = 16.4\,\text{Var}(\Theta) + 0.4\mathbb{E}(\Theta)^2$.

   We use the following code:

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

VOM<-var(policy.holder.means)
E_Theta<-book.premium/4
Var_Theta<-(VOM-0.4*E_Theta^2)/16.4
VHM<-16*Var_Theta
EPV<-4*(Var_Theta+E_Theta^2)

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

Z*policy.holder.means[15]+(1-Z)*book.premium
```

   This gives the book premium as \$813.0379 and the credibility of 10 years' experience as $Z = 0.524637$, and policyholder 15's premium for Year 11 as 680.1647.

5. *Aggregate claims per unit of exposure for a given individual policy are modelled as following a certain parametric distribution. Each policyholder has a risk parameter Θ. For a policyholder with risk parameter Θ and exposure m, the expected value of total annual claims is Θm, and the variance is $m(\Theta^2 + 8)$.*

From a dataset of 100 policyholders with different exposures, they find that the total aggregate claim is \$432,228 from a total of 1,220 units of exposure. They also calculate:

$$\sum m_i^2 = 19,145$$
$$\sum m_i^3 = 410,153$$
$$\sum m_i X_i^2 = 316,484,432$$

where X_i *is the aggregate claims per unit of exposure for Policyholder i (so* $\sum m_i X_i = 432228$*). Estimate the EPV and VHM from this data.*

[Hint: calculate the expectation of $\sum_{i=1}^{100} m_i \left(X_i - \widehat{\mathbb{E}(\Theta)}\right)^2$*.]*

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 $\widehat{\mathbb{E}\Theta} = \frac{\sum m_i X_i}{\sum m_i} = \frac{432228}{1220} = 354.285245902$. The EPV per unit of exposure is $\mathbb{E}(\Theta^2 + 8) = \left(\mathbb{E}(\Theta)^2 + \mathrm{Var}(\Theta)\right)$. We have

$$
\begin{aligned}
\mathbb{E}\left(X_i - \widehat{\mathbb{E}\Theta}\right)^2 &= \mathbb{E}\left(X_i - \frac{\sum m_i X_i}{\sum m_i}\right)^2 \\
&= \mathbb{E}\left(\left(1 - \frac{m_i}{\sum m_i}\right)^2 X_i^2 - \frac{2\sum_{j\neq i} m_j X_i X_j}{\sum m_j} + \left(\frac{\sum_{j\neq i} m_j X_j}{\sum m_i}\right)^2\right) \\
&= \left(1 - \frac{m_i}{\sum m_i}\right)^2 \mathbb{E}\left(X_i^2\right) - \frac{2\sum_{j\neq i} m_j \mathbb{E}\left(X_i X_j\right)}{\sum m_j} + \frac{\sum_{j,k\neq i} m_j m_k \mathbb{E}\left(X_j X_k\right)}{\left(\sum m_i\right)^2} \\
&= \left(1 - \frac{m_i}{\sum m_i}\right)^2 \left(\mathbb{E}(\Theta_i^2) + \frac{\mathbb{E}(\Theta_i^2 + 8)}{m_i}\right) - \frac{2\sum_{j\neq i} m_j \mathbb{E}\left(\Theta_i \Theta_j\right)}{\sum m_j} \\
&\qquad + \frac{\sum_{j,k\neq i} m_j m_k \mathbb{E}\left(\Theta_j \Theta_k\right) + \sum_{j\neq i} m_j^2 \mathbb{E}\left(\frac{\Theta_j^2 + 8}{m_j}\right)}{\left(\sum m_i\right)^2} \\
&= \left(1 - \frac{m_i}{\sum m_i}\right)^2 \left(\left(1 + \frac{1}{m_i}\right)\left(\mathbb{E}(\Theta)^2 + \mathrm{Var}(\Theta)\right) + \frac{8}{m_i}\right) - 2\frac{\sum_{j\neq i} m_j}{\sum m_j}\mathbb{E}\left(\Theta\right)^2 \\
&\qquad + \frac{\sum_{j,k\neq i} m_j m_k}{\left(\sum m_i\right)^2}\mathbb{E}\left(\Theta\right)^2 + \frac{\sum_{j\neq i} m_j^2 \,\mathrm{Var}(\Theta) + m_j\left(\mathbb{E}(\Theta)^2 + \mathrm{Var}(\Theta) + 8\right)}{\left(\sum m_i\right)^2} \\
&= \left(\left(1 - \frac{m_i}{\sum m_i}\right)^2 \left(1 + \frac{1}{m_i}\right) - 2\left(1 - \frac{m_i}{\sum m_j}\right) + \left(1 - \frac{m_i}{\sum m_i}\right)^2 + \frac{1}{\sum m_j}\left(1 - \frac{m_i}{\sum m_j}\right)\right)\mathbb{E}(\Theta)^2 \\
&\quad + \left(\left(1 - \frac{m_i}{\sum m_i}\right)^2 \left(1 + \frac{1}{m_i}\right) + \frac{\sum_{j\neq i} m_j^2}{\left(\sum m_j\right)^2} + \frac{1}{\sum m_j}\left(1 - \frac{m_i}{\sum m_j}\right)\right)\mathrm{Var}(\Theta) \\
&\qquad + \left(1 - \frac{m_i}{\sum m_i}\right)^2 \frac{8}{m_i} + \frac{8}{\sum m_j}\left(1 - \frac{m_i}{\sum m_j}\right) \\
&= \left(1 - \frac{m_i}{m}\right)\left(\left(1 - \frac{m_i}{m}\right)\left(1 + \frac{1}{m_i}\right) - 2 + \left(1 - \frac{m_i}{m}\right) + \frac{1}{m}\right)\mathbb{E}(\Theta)^2 \\
&\quad + \left(\left(1 - \frac{m_i}{m}\right)^2 \left(1 + \frac{1}{m_i}\right) + \frac{\sum_{j\neq i} m_j^2}{m^2} + \frac{1}{m}\left(1 - \frac{m_i}{m}\right)\right)\mathrm{Var}(\Theta) \\
&\qquad + \left(1 - \frac{m_i}{m}\right)\left(\left(1 - \frac{m_i}{m}\right)\frac{8}{m_i} + \frac{8}{m}\right) \\
&= \left(1 - \frac{m_i}{m}\right)\left(\frac{1}{m_i} - 2\frac{m_i}{m}\right)\mathbb{E}(\Theta)^2 \\
&\quad + \left(\left(1 - \frac{m_i}{m}\right)\left(\left(1 + \frac{1}{m_i} - \frac{m_i}{m}\right) + \frac{\sum_{j\neq i} m_j^2}{m^2} - \frac{m_i}{m^2}\right)\right)\mathrm{Var}(\Theta) \\
&\qquad + \frac{8}{m_i}\left(1 - \frac{m_i}{m}\right)
\end{aligned}
$$

This gives us

$$
\begin{aligned}
\mathbb{E}\left(\sum_i m_i\left(X_i-\widehat{\mathbb{E}\Theta}\right)^2\right) &= \sum m_i\left(1-\frac{m_i}{m}\right)\left(\frac{1}{m_i}-2\frac{m_i}{m}\right)\mathbb{E}(\Theta)^2 \\
&\quad + \sum m_i\left(\left(1-\frac{m_i}{m}\right)\left(\left(1+\frac{1}{m_i}-\frac{m_i}{m}\right)+\frac{\sum_{j\neq i} m_j^2}{m^2}-\frac{m_i}{m^2}\right)\right)\operatorname{Var}(\Theta) \\
&\quad + 8\sum\left(1-\frac{m_i}{m}\right) \\
&= \sum\left(1-\frac{m_i}{m}\right)\left(1-2\frac{m_i^2}{m}\right)\mathbb{E}(\Theta)^2 \\
&\quad + \sum\left(\left(1-\frac{m_i}{m}\right)\left(\left(m_i+1-\frac{m_i^2}{m}\right)+\frac{\sum_{j\neq i} m_i m_j^2}{m^2}-\frac{m_i^2}{m^2}\right)\right)\operatorname{Var}(\Theta) \\
&\quad + 8(n-1) \\
&= \left(n-1-2\frac{\sum m_i^2}{m}+2\frac{\sum m_i^3}{m^2}\right)\mathbb{E}(\Theta)^2 \\
&\quad + \sum\left(m_i+1-2\frac{m_i^2}{m}-\frac{m_i}{m}+\frac{m_i^3}{m^2}+\frac{m_i\sum_j m_j^2}{m^2}-\frac{m_i^3}{m^2}-\frac{m_i^2}{m^2}\right)\operatorname{Var}(\Theta) \\
&\quad + 8(n-1) \\
&= \left(n-1-2\frac{\sum m_i^2}{m}+2\frac{\sum m_i^3}{m^2}\right)\mathbb{E}(\Theta)^2 \\
&\quad + \left(m+n-1-\sum\frac{m_i^2}{m}-\frac{\sum m_i^2}{m^2}\right)\operatorname{Var}(\Theta) \\
&\quad + 8(n-1) \\
&= \left(100-1-2\frac{19145}{1220}+2\frac{410153}{1220^2}\right)\mathbb{E}(\Theta)^2 \\
&\quad + \left(1220+100-1-\sum\frac{19145}{1220}-\frac{19145}{1220^2}\right)\operatorname{Var}(\Theta) \\
&\quad + 8(100-1) \\
&= 68.1658868584\mathbb{E}(\Theta)^2 + 1303.29451424\operatorname{Var}(\Theta) + 792
\end{aligned}
$$

We therefore have that

$$
\frac{\sum m_i\left(X_i-\widehat{\mathbb{E}(\Theta)}\right)^2-792-68.1658868584\widehat{\mathbb{E}(\Theta)}^2}{1303.29451424}
$$

is an unbiased estimate for Var(Θ). We have

$$\begin{aligned}\sum m_i \left(X_i - \widehat{\mathbb{E}(\Theta)}\right)^2 &= \sum m_i \left(X_i^2 - 2X_i\widehat{\mathbb{E}(\Theta)} + \left(\widehat{\mathbb{E}(\Theta)}\right)^2\right)\\ &= \sum m_i X_i^2 - 2\sum m_i X_i \widehat{\mathbb{E}(\Theta)} - n\left(\widehat{\mathbb{E}(\Theta)}\right)^2\\ &= 316484432 - 2 \times 432228 + 100 \times 354.285245902^2\\ &= 328171779.546\end{aligned}$$

Thus our estimator for the VHM per unit of exposure is

$$\frac{328171779.546 - 792 - 68.1658868584 \times 354.285245902^2}{1303.29451424} = 245236.157944$$