ASTAM April 2023 Model Solutions

Question 1

Let N denote the claim frequency and Y_j denote the *j*-th claim amount.

(a)

$$\begin{aligned}\Pr[S \le 350] &= \Pr[N=0] + \Pr[N=1]\Pr[Y_1 \le 350] \\ &\quad + \Pr[N=2]\Pr[Y_1+Y_2 \le 350] \\ &\quad + \Pr[N=3]\Pr[Y_1+Y_2+Y_3 \le 350] \\ &= 0.20 + 0.25(0.9) + 0.4\left(0.5^2 + 2(0.5)(0.4)\right) + 0.15\left(0.5^3\right) \\ &= 0.70375\end{aligned}$$

(b)

$$\mathrm{E}[N] = 0(0.2) + 1(0.25) + 2(0.4) + 3(0.15) = 1.5$$

$$\mathrm{E}[Y_j] = 100(0.5) + 200(0.4) + 500(0.1) = 180$$

$$\mathrm{E}[S] = \mathrm{E}[N]\mathrm{E}[Y_j] = 270$$

$$P = 1.15\,\mathrm{E}[S] = 310.5$$

(c) Let $Y_j^* = Y_j - 100$. Then

$$\mathrm{E}\left[Y_j^*\right] = E[(Y_j - 100)_+] = 0(0.5) + 100(0.4) + 400(0.1) = 80$$

$$P = 1.15\,\mathrm{E}[N]\mathrm{E}\left[Y_j^*\right] = 138$$

(d)

$$S* = \begin{cases} 0 & \text{if } S = 0 \\ S - 100 & \text{if } S > 0 \end{cases}$$

$$\begin{aligned}E[S^*] &= E[S] - E[S \wedge 100] \\ &= 270 - [(0)(0.2) + (100)(0.8)] \\ &= 270 - 80 = 190 \\ \Rightarrow P &= 1.15(190) = 218.5\end{aligned}$$

(e) Let S^R denote the reinsurer's claims.

(i) $$S^R = \left(S-1000\right)^+ = \begin{cases} 0 & \text{if } S \le 1000 \\ (S-1000) & \text{if } S > 1000 \end{cases}$$

The only way that aggregate claims exceed 1000 is if there are 3 claims, and at least 2 of them are equal to 500.

So $$E\left[S^R\right] = \Pr\left[N=3\right]\{100 \Pr\left[Y_1+Y_2+Y_3=1100\right]$$
$$+ 200 \Pr\left[Y_1+Y_2+Y_3=1200\right]$$
$$+500 \Pr\left[Y_1+Y_2+Y_3=1500\right]\}$$
$$=(0.15)\left\{100\left(3(0.1^2)(0.5)\right)+200\left(3(0.1^2)(0.4)\right)+500\left(0.1^3\right)\right\}$$
$$=0.660$$

(ii) Stop Loss insurance protects the insurer in the same way that the policy limit does, but the policy limit may discourage some potential policyholders.

Examiners' Comments

Candidates performed well on this question. The average score was the highest of the exam.

Some candidates just wrote down answers without showing work and stated that the work was done in Excel. If the answer was correct, the candidate received partial credit. Even when using Excel, candidates are required to show their working on their answer sheets. Note that the examiners do not have access to the candidates' Excel workbooks.

(a) Overall candidates did well on this part of the question, with around 50% achieving maximum credit. For those who did not, a common mistake was to calculate the expected value and the variance for the random variable S and use the normal distribution to estimate the probability. However, the distribution is not even approximately normal and there is no reason to think that it should be so.

Other common errors were to omit the probability that S = 0 or to forget that the probability that two claims where one is for 100 and the other is for 200 could occur in two different ways.

(b) and (c) Candidates did exceptionally well on these parts with the majority getting full credit.

(d) Candidates performed acceptably on this part of the question. Quite a few candidates attempted to calculate all possible values of (S – 100) from 0 to 1400 and assign probabilities to each value. While a few candidates were able to complete this calculation, most candidates made an error and received only partial credit.

(e) (i) Overall, very few candidates got the correct answer; most candidates omitted this part. Many candidates who attempted the problem tried to use E(S) – E(S^1000), but this was not generally a feasible calculation.

The approach shown above demonstrates a first principles understanding of Stop Loss insurance and is relatively easy to calculate. Candidates who used this approach generally got the right answer. The most common mistakes was to forget the factors of 3 in th first two terms, allowing for different combinations of claims. These candidates still received the majority of the points for this part of the question.

(ii) While most candidates attempted this part, very few provided a coherent, relevant explanation.

.

Question 2

(a)

(i) Since $p_0 = g(\lambda) = e^{-\lambda}$ is a one-to-one function, by the invariance property of MLE, the MLE of p_0 is $\hat{p}_0 = e^{-\hat{\lambda}} = e^{-0.84} = 0.4317$.

(ii) We know that $\text{Var}\left[\hat{\lambda}\right] = \text{Var}\left[\bar{N}\right] = \dfrac{\lambda}{n}$, and $g'(\lambda) = -e^{-\lambda}$.

Using the delta method, the variance of the maximum likelihood estimator of p_0 is

$$\text{Var}\left(g\left(\hat{\lambda}\right)\right) \approx \left[g'\left(\hat{\lambda}\right)\right]^2 \text{Var}\left(\hat{\lambda}\right) \quad = \left[-e^{-0.84}\right]^2 \times \frac{0.84}{100}$$

$$= 0.001566.$$

Therefore, the 95% linear CI for p_0 is

$$\left(0.4317 \pm 1.96\sqrt{0.001566}\right) = \left(0.3542,\ 0.5093\right)$$

(b)

(i) The likelihood function for the zero-modified Poisson distribution is

$$L\left(\lambda, p_0^M\right) = \prod_{k=0}^{\infty}\left(p_k^M\right)^{n_k} = \left(p_0^M\right)^{n_0} \prod_{k=1}^{4}\left(\frac{1-p_0^M}{1-p_0}\, p_k\right)^{n_k}$$

$$= \left(p_0^M\right)^{n_0}\left(1-p_0^M\right)^{n-n_0} \prod_{k=1}^{4}\left(\frac{\lambda^k e^{-\lambda}}{k!\left(1-e^{-\lambda}\right)}\right)^{n_k}$$

$$= \left(p_0^M\right)^{46}\left(1-p_0^M\right)^{54}\left(1-e^{-\lambda}\right)^{54} e^{-54\lambda}\lambda^{84} K$$

where $K = \prod_{k=1}^{4}\left(\dfrac{1}{k!}\right)^{n_k}$ does not involve λ or p_0^M

(ii) The log-likelihood function is

$$l\left(\lambda, p_0^M\right) = n_0 \log p_0^M + (n-n_0)\log\left(1-p_0^M\right) + C$$

where C is a function of λ only. Therefore,

$$\frac{\partial}{\partial p_0^M} l\left(\lambda, p_0^M\right) = \frac{n_0}{p_0^M} - \frac{n-n_0}{1-p_0^M}$$

Set equal to 0 for the MLE:

$$0 = \frac{n_0}{\hat{p}_0^M} - \frac{n - n_0}{1 - \hat{p}_0^M} \Rightarrow \left(1 - \hat{p}_0^M\right) n_0 = \hat{p}_0^M \left(n - n_0\right)$$

$$\Rightarrow \hat{p}_0^M = \frac{n_0}{n} = 0.46$$

(c)

(i) As a random variable, $\hat{p}_0^M = \frac{N_0}{n}$ where $N_0 \sim Bin(n, p_0^M)$,

$$\text{Var}\left[\hat{p}_0^M\right] = \text{Var}\left[\frac{N_0}{n}\right] = \frac{1}{n^2}\left(np_0^M\left(1 - p_0^M\right)\right) = \frac{p_0^M\left(1 - p_0^M\right)}{n}$$

$$\approx \frac{0.46(1 - 0.46)}{100} = 0.002484$$

(ii) The estimated 95% confidence interval is

$$\left(0.46 \pm 1.96\sqrt{0.002484}\right) = \left(0.3623, 0.5577\right)$$

(d)

(i) The SBC is calculated as $l - \frac{r}{2}\ln n$, where l is the maximum loglikelihood function, r is the number of parameters, and n is the sample size.

SBC for Poisson is $-124.36 - \frac{1}{2}\ln 100 = -126.66$

SBC for ZM Poisson is $-123.91 - \frac{2}{2}\ln 100 = -128.52$

(ii) The statement is incorrect.

The SBC is designed to assess the overall fit of the data to the distribution, taking into consideration the number of parameters involved. It is a very rough method of assessing whether additional parameters provide a significantly better overall fit.

The SBC uses the maximum log likelihood. The MLE method tends to fit to the center of the underlying distribution, and may not provide a good fit in the tails.

The zero-modified Poisson specifically modifies the probability weight at zero to achieve a better fit of p_0 to the data.

Examiners' Comments

Overall, this question proved to be one of the more difficult ones. Candidates often knew what formulas to use, but were unable to derive or explain them.

Candidates did very well on part (a), with most achieving maximum credit.

Part (b) proved much more challenging. Very few candidates were able to write down the likelihood function of the ZM Poisson distribution, and therefore very few could derive the MLE required in (b)(ii). Many candidates omitted this part completely.

Part (c) was also omitted by many candidates. Others correctly wrote down the variance, but few were able to satisfy the "show that" part of the question.

Most candidates correctly evaluated the SBC in part (d)(i), but few recognized that the better fitting model overall might have a worse fit in the left tail.

Question 3

(a)

(i) The $m_{i,j}$ term is a measure of exposure, or volume of data, associated with X_{ij} – that is, associated with from risk group i in year j. It is assumed to be known.

(ii) The θ_i term represents the unknown risk factors associated with risk group i. It may be interpreted in a Bayesian sense as the unknown parameter vector of an underlying distribution. It is treated as a random variable in the Bühlmann-Straub model.

(b)

(i) From the formula sheet:

$$\hat{v} = \frac{\sum_{i=1}^{r}\sum_{j=1}^{n_i} m_{ij}\left(X_{ij} - \bar{X}_i\right)^2}{\sum_{i=1}^{r}\left(n_i - 1\right)}$$

Where

$$X_{1,1} = \frac{39}{200} = 0.195; \quad X_{1,2} = 0.18667; \quad X_{1,3} = 0.0833 \quad X_{2,2} = 0.320; \quad X_{2,3} = 0.3833$$

$$\bar{X}_1 = \frac{82}{530} = 0.15472; \quad \bar{X}_2 = \frac{39}{110} = 0.35454$$

$$n_1 = 3; \quad n_2 = 2$$

$$\sum_{j=1}^{3} m_{1,j}\left(X_{1,j} - \bar{X}_1\right)^2 = 200\left(0.195 - 0.15472\right)^2 + 150(0.18667 - 0.15472)^2 + 180(0.08333 - 0.15472)^2$$

$$= 1.39487$$

$$\sum_{j=2}^{3} m_{2,j}\left(X_{2,j} - \bar{X}_2\right)^2 = 50(0.32 - 0.35454)^2 + 60(0.38333 - 0.35454)^2$$

$$= 0.109394$$

$$\Rightarrow \hat{v} = \frac{1.50427}{3} = 0.5014$$

(ii) From the formula sheet

$$\hat{a} = \frac{\sum_{i=1}^{r} m_i \left(\bar{X}_i - \bar{X}\right)^2 - \hat{v}(r-1)}{m - \frac{1}{m}\sum_{i=1}^{r} m_i^2}$$

$$m_1 = 530; \quad m_2 = 110; \quad m = 640; \quad \bar{X} = \frac{82+39}{530+110} = 0.18906$$

$$\Rightarrow \hat{a} = \frac{530\left(0.15472-0.18906\right)^2 + 110\left(0.35454-0.18906\right)^2 - 0.5014}{640 - \left(\frac{530^2+110^2}{640}\right)}$$

$$= \frac{3.13608}{182.1875} = 0.017213$$

(c) $P = Z_2\bar{X}_2 + \left(1-Z_2\right)\hat{\mu}$

$$Z_2 = \frac{m_2}{m_2 + \hat{v}/\hat{a}} = 0.7906; \quad Z_1 = \frac{m_1}{m_1 + \hat{v}/\hat{a}} = 0.9479$$

From the formula sheet: $\hat{\mu} = \dfrac{\sum_{i=1}^{2} Z_i \bar{X}_i}{\sum_{i=1}^{2} Z_i} = 0.24559$

$$\Rightarrow P = 0.7906 \times 0.35455 + 0.2094 \times 0.24559 = 0.33173$$

(d) Let $\pi\left(A \mid \underline{X}\right)$ and $\pi\left(B \mid \underline{X}\right)$ denote the posterior probabilities for Broker A and B, given a claims history $\underline{X} = (2,2)$.

(i)

$$\hat{\lambda}_A = \frac{82}{530} = 0.15472; \quad \hat{\lambda}_B = \frac{39}{110} = 0.35454$$

$$\pi\left(A \mid \underline{X}\right) = \frac{p(\underline{X} \mid A)\pi(A)}{p(\underline{X})}; \quad \pi\left(B \mid \underline{X}\right) = \frac{p(\underline{X} \mid B)\pi(B)}{p(\underline{X})}$$

where $\pi(A) = 0.75; \quad \pi(B) = 0.25$

$$p(\underline{X} \mid A)\pi(A) = \left(\frac{\hat{\lambda}_A^2 e^{-\hat{\lambda}_A}}{2}\right)^2 \times 0.75 = 0.00010513 \times 0.75 = 7.8844 \times 10^{-5}$$

$$p(\underline{X} \mid B)\pi(B) = \left(\frac{\hat{\lambda}_B^2 e^{-\hat{\lambda}_B}}{2}\right)^2 \times 0.25 = 0.0019439 \times 0.25 = 4.8597 \times 10^{-4}$$

$$\Rightarrow \pi\left(B \mid \underline{X}\right) = \frac{4.8597 \times 10^{-4}}{7.8844 \times 10^{-5} + 4.8597 \times 10^{-4}} = 0.8604$$

(ii) $P = (1-0.8604)\hat{\lambda}_A + 0.8604\hat{\lambda}_B$
$= 0.32665$

Examiners' Comments:
The descriptions in (a) are more comprehensive than was required for full credit. For (i) it was sufficient to say that $m_{i,j}$ *is a measure of volume or exposure for policyholder (or risk group) i in year j. For (ii) it is sufficient to say that* θ_i *is a random risk parameter associated with risk group/policyholder i. Explanations needed to be coherent and correct. Only the best candidates were able to give a correct and coherent interpretation of both parameters.*

Part (b) was omitted by a significant minority of candidates, but those who attempted it generally achieved the majority of the available points.

In part (c) the majority of the candidates mistakenly used $\overline{X}$ *as the estimate of* $\hat{\mu}$ *for part (c). (It is possible that candidates were confusing the Buhlmann and Buhlmann-Straub credibility models.) Partial credit was given to candidates making this mistake.*

Most candidates attempting part (d) did relatively well, earning most or all of the available points, though many candidates omitted this part.

Question 4

(a) Note that 98.4%(751)=738.984 belongs to (738,739). Let $X_{(k)}$ denote the k-th smallest flood loss value.

(i) Using the smoothed empirical method

$$\widehat{Q}_{98.4\%} = 0.016(X_{(738)}) + 0.984(X_{(739)}) = 2778$$

(ii) $$\widehat{ES}_{98.4\%} = \frac{X_{(739)} + ... + X_{(750)}}{12} = \frac{90919 + 2924 + 2780}{12} = 8052$$

(iii) $$\hat{e}(3000) = \left(\text{ave of all values } >3000\right) - 3000 = \frac{90919}{10} - 3000 = 6091.9$$

(b) We expect the distribution of excess loss, $(X - d \mid X > d)$ to converge to the GPD as $d \to \infty$. This means that the MEL function should converge to the MEL of the GPD, (i.e. a straight line) as the excess loss threshold increases. Thus, we are looking for the <u>ultimate</u> gradient of the MEL, which is clearly not represented by the shape change at around 1500.

(c)

(i) $$Q_\alpha = d + \frac{\beta}{\xi}\left(\left(\frac{S_X(d)}{1-\alpha}\right)^{\xi} - 1\right) \quad \text{(formula sheet)}$$

$$d = 2100; \quad \alpha = 0.984; \quad S_X(d) \approx \frac{16}{750}$$

$$\Rightarrow \widehat{Q}_\alpha = 2100 + \frac{2000}{0.68}\left(\left(\frac{16/750}{0.016}\right)^{0.68} - 1\right) = 2735$$

(ii) $$ES_\alpha = \frac{1}{1-\xi}\left(Q_\alpha + \beta - \xi d\right) \quad \text{(formula sheet)}$$

$$\Rightarrow \widehat{ES}_\alpha = \frac{1}{0.32}\left(2735 + 2000 - 0.68(2100)\right) = 10{,}334$$

(d) The VaR estimates are similar, at around 2700. We see from the MEL plot that this is close to the point where the distribution starts to exhibit its tail GPD behaviour; the stretching out of the tail beyond the threshold does not have much effect this close to the threshold, so the empirical VaR will be close to the GPD.

The ES estimates are very different, (8052 compared with 10334).
The MEL plot has a positive right tail gradient, indicating a very fat-tailed distribution.
It is common with fat-tailed distributions for the empirical tail ES to be less than the GPD estimate.

The empirical estimate is limited by the size of the largest 1 or 2 observations. Occasionally these will be jumbo values (i.e. very rare very very large losses) but more often they will not be.
The GPD estimate uses the GPD extrapolation beyond the observed data, and so will always allow for the jumbo losses. Since the ES is an average from the entire right tail of the loss distribution, the impact of the jumbo losses is often significant. So the recommended indicator should be the GPD estimate.

Examiners' Comments:
Performance on part (a) was very mixed. Many candidates lost a small amount of credit for using the raw quantile estimate instead of the smoothed value. Some candidates stated that their calculations were done in Excel. For maximum credit, they needed to write down on their answer sheets the formulas used.
Around 50% of candidates understood the key point in (b), that it is the gradient in the tail that matters.
Part (c) was done well by those who attempted it. The main error was misunderstanding the role of $S_X(d)$ *in the quantile formula.*
The solution for part (d) shown above is far more detailed than would be required for maximum credit. The key points are that the VaR may not be much affected by the tail of a fat-tailed distribution, so GPD estimate is often close to the empirical, but the ES uses the whole tail of the distribution, which includes the very rare, very large claims, and therefore will commonly be greater than empirical (unless the empirical data includes a very rare jumbo claim). Many candidates omitted this part.

ASTAM April 2023 Question Q5 Grading Outline

(a) Premium income in 2022 for Territory A:

$$5000\times 200+2200\times 200\times 0.95=1,418,000 .$$

Premium income in 2022 for Territory B:

$$3100(200)(1.2)+1300(200)(1.2)(0.95)=1,040,400 .$$

So, the loss ratios at current rates for Territories A and B are

A: $842,000/1,418,000=0.5938$

B: $603,000/1,040,400=0.5796$

The new differential for B is $1.20\times\dfrac{0.5796}{0.5938}=1.1713.$

(b) We assume that loss developments continue in the same ratio for both territories, so that using the fully developed data to calculate the new differential would give the same result.

(c) Let $B=$ new base rate:

$$B\left(5000+(0.95)(2200)+1.15(3100)+(1.15)(0.95)(1300)\right)$$

$$=(1.07)\left(1,418,000+1,040,400\right)$$

$$\Rightarrow B=\frac{2,630,488}{12,292.0}=217.84$$

(d)

(i) The territory B premiums will increase by

$$\frac{217.84}{200.00}\times\frac{1.15}{1.20}-1=4.36\%$$

(ii) The Territory A premiums will increase by $\dfrac{217.84}{200.00}-1=8.92\%$.

We would therefore expect more attrition from Territory A policyholders, than Territory B, and possibly more new business from Territory B, if the revised differential results in a more competitive premium in that region. Both factors imply that the Territory A proportion would be expected to decrease.

Examiners' Comments:

Overall, this question proved more challenging than expected, with the lowest average score of all the questions.

Many candidates omitted the question, or made a very brief attempt at one or two parts.

In part (a) , many candidates did not allow for the age differentials within each territory

Few candidates provided a coherent explanation for (b).

Part (c) was done well by those who attempted it, with most achieving maximum credit. However, it is not sufficient for candidates to memorize algorithms or formulas, they are also expected to know the assumptions or theory behind the formulas.

Part (d)(i) was done well by those who attempted it, but few candidates understood the implications of the changes tested in (d)(ii).

Question 6

(a)

(i) Outstanding Claims Reserves are used to provide for insured losses that have occurred, but which have not been fully settled.

(ii) Reasons include:

- Delays in reporting the claim – some claims are reported soon after the loss event (eg auto), bit some may be reported much later, if the damage is not apparent until then (eg medical malpractice).
- Delays in processing the claim – insurer needs to validate the facts, assess the loss, investigate potential recoveries from other parties or from salvage.
- Legal proceedings or disputes about the claim – legal disputes usually involve assignment of responsibility for a loss, or the appropriate amount of loss.

(b)

$$\hat{f}_2 = \frac{355}{325} = 1.09231 \quad \Rightarrow \hat{C}_{2,3} = 225(1.09231) = 245.77$$

$$\Rightarrow \hat{X}_{23} = 245.8 - 225 = 20.77$$

$$\hat{f}_3 = \frac{158}{145} = 1.08966 \quad \Rightarrow \hat{C}_{2,4} = 245.8(1.08966) = 267.80$$

$$\Rightarrow \hat{X}_{24} = 267.8 - 245.8 = 22.03$$

(c)

(i)

$$X_{i,j} \sim \text{Poi}\left(\mu_i \gamma_j\right) \Rightarrow \text{E}\left[X_{i,j}\right] = \mu_i \gamma_j$$

$$C_{i,j} = \sum_{k=0}^{j} X_{i,k} \Rightarrow \text{E}\left[C_{i,j}\right] = \sum_{k=0}^{j} \text{E}\left[X_{i,k}\right] = \sum_{k=0}^{j} \mu_i \gamma_k = \mu_i \sum_{k=0}^{j} \gamma_k$$

(ii) Note that the total claims for AY *i* are $C_{i,4}$, where

$$\text{E}\left[C_{i,4}\right] = \mu_i \sum_{k=0}^{4} \gamma_k = \mu_i \ .$$

Hence, μ_i represents the expected total claims cost from AY *i*.

(iii) Note that the claims paid in DY *j*, arising in AY *i*, are $X_{i,j}$, where $\text{E}\left[X_{i,j}\right] = \mu_i \gamma_j$, and the γ_j's sum to 1.0. Hence, γ_j represents the expected proportion of the ultimate cumulative claims cost that is paid in DY *j*, for any AY.

(d)

(i) The MLE of $E[X_{2,3}]$ is $\hat{\mu}_2 \hat{\gamma}_3$, which is given to be equal to the chain ladder estimate $\hat{X}_{2,3}$. So, from (b), we have $\hat{\mu}_2 \hat{\gamma}_3 = 20.77$.

(ii) Similarly, we have

$$\hat{C}_{2,4} = \hat{\mu}_2 = 267.80 \text{ (from (b))} \quad \Rightarrow \hat{\gamma}_3 = \frac{20.77}{267.80} = 0.07756$$

$$\hat{X}_{2,4} = \hat{\mu}_2 \hat{\gamma}_4 = 22.03 \Rightarrow \hat{\gamma}_4 = \frac{22.03}{267.80} = 0.08226$$

(iii) The sum of Poisson RVs is also Poisson, with parameter equal to the sum of the individual Poisson parameters. The outstanding claims for AY 2 are

$$X_{2,3} + X_{2,4} \sim Poi\left(\mu_2(\gamma_3 + \gamma_4)\right)$$

$$\Rightarrow SD\left[X_{2,3} + X_{2,4}\right] \approx \sqrt{\hat{\mu}_2(\hat{\gamma}_3 + \hat{\gamma}_4)} \quad \approx \sqrt{20.77 + 22.03} = 6.542$$

(iv) The approximate 95% CI is

$$42.80 \pm 1.96(6.542) = [29.98, 55.62]$$

(e) The Poisson model provides a distributional model, which can be used for (eg) confidence intervals and probabilistic inference, and which involves strong assumptions about the distribution and independence of incremental claims in different years. The chain ladder model is a deterministic projection that provides no information on the distribution of outcomes, and which involves much weaker assumptions.

Examiners' Comments:

Parts (a) and (b) were done well, with many candidates achieving maximum points.

Around 1/3 of candidates omitted all or most of parts (c), (d) and (e), indicating that they had mastered the deterministic chain ladder material, but not the extension to stochastic models. Of those who attempted (c) and (d), many scored highly, though very few achieved maximum points. Note that the explanations given in the part (b) and (c) solutions above were not required for full credit. Even fewer identified the main difference between the Poisson and Chain ladder models for part (e).