ASTAM Spring 2024 Model Solutions

ASTAM April 2024 Model Solutions

Question 1

The solution to this question is in the spreadsheet. It should be noted that as stated in the instructions for the exam, only work in the spreadsheet will be graded. Any work on paper is NOT graded for Excel problems.

ASTAM April 2024 Model Solutions

Question 2

(a)

1. The individual loss distribution will generally be highly positively skewed, as most losses are fairly small, but there is a small probability of a very large loss. In contrast, the average loss per policy will be fairly symmetric, because the distribution is approximately Normal, from the Central Limit Theorem.

2. The individual policy loss distribution will have a probability mass at zero, as many policies do not generate any claims. The average loss per policy distribution will not.

3. The individual policy loss distribution will have a much higher standard deviation - i.e. a much larger spread – than the average loss per policy distribution, as the standard deviation of the average is 1/N times the standard deviation of each individual loss (assuming i.i.d).

(b)

The average gross premium rate for 2024 is

$$\frac{\text{New ave loss cost} + \text{fixed expenses per exposure}}{\text{permissible loss ratio}}$$

$$= \frac{(25000 + 30000)}{20} \times \frac{1}{0.85}$$

$$= 1647.06$$

(c)

Since Region B has a 2023 Differential of 1.00, this is considered the base case.

The 2024 differentials then for each region are:

$$\text{Region A:} \quad \text{2023 differential} \times \frac{\text{Region A Loss Ratio}}{\text{Region B Loss Ratio}}$$

$$= 0.80\frac{0.72}{0.80} = 0.72$$

Region B: 1.00 (still base case)

$$\text{Region C:} \quad \text{2023 differential} \times \frac{\text{Region A Loss Ratio}}{\text{Region B Loss Ratio}}$$

$$= 1.20\frac{0.76}{0.80} = 1.14$$

(d)

(i) The differentials calculated from the two methods will not exactly be the same when there are several risk classification characteristics used and the distribution of all cross-variables in any cell (called the cell population) is heterogeneous.

(ii) On the other hand, if the cell population is homogeneous, even with several risk classifications, the differentials from the two methods will yield to the same result.

(e)

From (b), the average gross premium rate for 2024 is 1647.06 so that the target increase is

$$\frac{1647.06}{1500} = 1.098$$

The off-balance factor for all three regions is

$$\frac{0.72(5,000)+1(10,000)+1.14(5,000)}{0.80(5,000)+1(10,000)+1.20(5,000)} = \frac{19,300}{20,000} = 0.965$$

The revised base premium rate for 2024 should therefore be

$$\frac{1.098}{0.965}(1,500) = 1706.74$$

Examiners' Comments:

For Part(a), the candidates did not do great on this question. Many did correctly identify the difference in variance. However, beyond that they tended not really address the question with regard to the difference between the distributions. For Part (b) and Part (c), virtually everyone got these correct. The only common error was that some candidates did not note the B was the Base.

For Part (d), most students received zero or very few points on this question. A common mistake was to try to mathematically equate the two formulas. Very few candidates demonstrated a true understanding of the conditions necessary for the two approaches to be the same.

For Part e, students still did well on this part although not as well as on b and c. A common error was using the inverse of the off-balance factor.

ASTAM April 2024 Model Solutions

Question 3

(a) The mean is

$$E[S] = E[N]E[X] = 10.0$$

The variance is

$$\begin{aligned}\mathrm{Var}[S] &= E\big[\mathrm{Var}[X|N]\big] + \mathrm{Var}\big[E[X|N]\big] \\ &= E[N\mathrm{Var}[X]] + \mathrm{Var}[10N] \\ &= E[60N] + 100\mathrm{Var}[N] = 60 + 90 = 150 \\ \Rightarrow SD[S] &= \sqrt{150} = 12.2474\end{aligned}$$

(b)

(i) There are three (a,b,0) distributions.

The binomial has variance $<$ mean

The Poisson has variance = mean

The negative binomial has variance > mean.

Therefore, since variance $<$ mean, N must be binomially distributed.

(ii) Suppose $N \sim Bin(m, q)$. Then

$$E[N] = mq = 1; \quad Var[N] = mq(1-q) = 0.9$$

$$\Rightarrow q = 0.1 \text{ and } m = 10$$

(iii) We know (from (i)) that the severity distribution is negative binomial.

That is:

$$\begin{aligned}&X_j \sim NB(r, \beta) \\ &\Rightarrow E\left[X_j\right] = r\beta = 10; \; Var\left[X_j\right] = r\beta(1+\beta) = 60 \\ &\Rightarrow \beta = 5; \quad r = 2\end{aligned}$$

(c) The probability generating function of S is

$$
\begin{aligned}
P_S(z) &= P_N(P_X(z)) \\
&= \left(1 + q((1-\beta(z-1))^{-r} - 1)\right)^m \\
&= \left(1 + 0.1\left(\tfrac{1}{(1-5(z-1))^2} - 1\right)\right)^{10} \\
&= \left(0.9 + \frac{0.1}{(6-5z)^2}\right)^{10}
\end{aligned}
$$

(d)

(i) Using the pgf we have $Pr[S=0] = P_S(0) = \left(0.9 + \frac{0.1}{36}\right)^{10} = 0.35959$

Alternatively, let $Pr[X=k] = f_X(k)$, then

$$
\begin{aligned}
&Pr[S=0] = P_N(f_X(0)); \\
&\quad f_X(0) = \left(\frac{1}{6}\right)^2 = 0.027778 \\
&\Rightarrow Pr[S=0] = \left(1 + 0.1(0.027778 - 1)\right)^{10} = 0.35959
\end{aligned}
$$

(ii) Using the pgf we have

$$Pr[S=1] = \frac{d}{dz}P_S(z)\Big|_{z=0}$$
$$= \frac{10}{(6-5z)^3}\left(0.9+\frac{0.1}{(6-5z)^2}\right)^9\Bigg|_{z=0} = 0.018441$$

Hence, $Pr[S>1] = 1 - 0.35959 - 0.01844 = 0.62197$

Alternatively, using Panjer recursions, we have

$$Pr[S=1] = \frac{(a+b)f_X(1)f_S(0)}{1-af_X(0)}$$

Where

$$a = -\frac{q}{1-q} = -0.111111; \quad b = \frac{(m+1)q}{1-q} = 1.222222;$$
$$f_X(1) = r\left(\frac{\beta}{1+\beta}\right)\left(\frac{1}{1+\beta}\right)^r = 0.046296$$
$$\Rightarrow Pr[S=1] = \frac{1.11111(0.046296)(0.35959)}{1+0.11111(0.027778)} = 0.018441$$

Therefore, $Pr[S>1] = 1 - 0.35959 - 0.01844 = 0.62197$

(e) The stop loss premium is:

$$E[(S-2)^+] = E[S] - E[S\wedge 2]$$
$$E[S\wedge 2] = 1f_S(1) + 2(Pr[S>1]) = 1.26238$$
$$\Rightarrow E[(S-2)^+] = 10 - 1.26238 = 8.73762$$

Alternatively:

$$E[(S-2)^+] = E[S] - 2 + 2f_S(0) + f_S(1) = 8.73762$$

Examiners' Comments

a. Was well answered by most candidates. Some candidates applied the formula directly, with a few using an incorrect formula.

Many candidates only calculated the variance and not the standard deviation.

b. Most well prepared candidates did well here.

c. Only a few candidates could answer the question. Many seemed not to understand what a PGF is, and confused it with a PDF.

d. Very few candidates understood the link between the PGF and the PDF. And instead used brute force or alternative approaches.

e. Many candidates were able to write down the stop-loss formula, but were unable to solve it.

Better prepared candidates did well on this question.

ASTAM April 2024 Model Solutions

Question 4

(a) The log likelihood of the random sample is

$$l = \sum_{i=1}^{100} log \quad f(x_i) = 100\, log\, 3 + 2\sum_{i=1}^{100} log\, x_i - \frac{1}{\theta^3}\sum_{i=1}^{100} x_i^3 - 300\, log\, \theta$$

(b) To maximize the log likelihood, we differentiate wrt θ:

$$\frac{\partial l}{\partial \theta} = \frac{3}{\theta^4}\sum_{i=1}^{100} x_i^3 - \frac{300}{\theta}$$

Set equal to 0:

$$\frac{3}{\theta^3}\sum_{i=1}^{100} x_i^3 = 300$$

$$\frac{1}{\theta^3}(109{,}812) = 100$$

$$\hat{\theta}^3 = \frac{1}{100}\sum_{i=1}^{100} x_i^3 = 1098.12$$

$$\Rightarrow \hat{\theta} = 10.31692$$

(c)

$$I = -E\left[\frac{\partial^2 l}{\partial \theta^2}\right] = \frac{12}{\theta^5}\sum_{i=1}^{100} E[X_i^3] - \frac{300}{\theta^2}$$

$$= \frac{12}{\theta^5}(100\theta^3) - \frac{300}{\theta^2} \quad = \frac{900}{\theta^2}$$

The asymptotic variance is I^{-1}, i.e.

$$\mathrm{SD}[\hat{\theta}] \approx \sqrt{\frac{\theta^2}{900}} = \frac{\theta}{30}$$

Approximate θ with $\hat{\theta}$ gives $\mathrm{SD}[\hat{\theta}] \approx \sqrt{0.118265} = 0.34390$

(d) Let V denote Var[X]. Then

$$V = \theta^2\Gamma(5/3) - \theta^2\Gamma(4/3)^2 = 0.10533\theta^2$$

Note that the Gamma function can be calculated in Excel

$\hat{V} = 0.105330\hat{\theta}^2 = 11.2115$

$\mathrm{Var}[\hat{V}] = (g'(\theta))^2\mathrm{Var}[\hat{\theta}]$ where $g(\theta) = 0.10533\theta^2$

$$= \left(0.21066\hat{\theta}\right)^2(0.118265) = 0.558657 = 0.74743^2$$

So a 95% CI for V is

$11.2115 \pm 1.96(0.74743) = (9.7465,\ 12.6765)$

(e)

(i) The invariance property of MLEs means that the MLE of a function of a parameter is the function of the MLE of the parameter; that is, for any function h, $\widehat{h(\theta)} = h(\hat{\theta})$.

We used the invariance property in (d),when we assumed that the MLE of V was a function of the MLE of θ, because V is a function of θ.

(ii) Asymptotically normality means that as the sample size increases, the distribution of the parameter estimator converges to a normal distribution.

We used asymptotic normality in the CI in part (d), where we assumed that 95% of the estimator distribution is captured by subtracting or adding 1.96 standard deviations to the estimated distribution mean.

Examiners' Comments

A. Overall, students did well, though many candidates struggled with the algebra with taking logs (switching log and summation). Additionally, candidates wanted to plug in given sample values instead of giving a general formula.

B. Most students did very well on this part, though some unnecessarily complicated it by plugging in values early.

C. Most students knew what they needed to calculate but struggled with taking expectations.

D. Students showed good knowledge of what formula was needed for the delta method, though some struggled with how to apply the formulas or made small arithmetic mistakes.

E. While most students knew where they used Asymptotic normality assumptions, many struggled with the definition, specifically not mentioning sample size increasing to infinity. Very few were able to explain where Invariance was used (or even what it was), though they did use the property correctly in previous parts.

ASTAM April 2024 Model Solutions

Question 5

(a) Let M denote the mean loss for a randomly selected policy, and let M_Y denote the mean loss from a policy given that the policy is from class Y, $Y \in \{A, B, C\}$. Note that each class is equally likely for a policy drawn at random.

$$M_A = (0)(0.9) + (100)(0.07) + (500)(0.03) = 22$$

$$M_B = (0)(0.5) + (100)(0.3) + (500)(0.2) = 130$$

$$M_C = (0)(0.3) + (100)(0.33) + (500)(0.37) = 218$$

$$\mu = E[M] = \frac{22 + 130 + 218}{3} = 123.33$$

(b) Let V denote the variance of loss for a policy drawn at random, and let V_Y denote the variance for a policy given that the policy class is Y.

$$V_A = (0)^2(0.9) + (100)^2(0.07) + (500)^2(0.03) - 22^2 = 7716$$
$$V_B = (0)^2(0.5) + (100)^2(0.3) + (500)^2(0.2) - 130^2 = 36{,}100$$
$$V_C = (0)^2(0.3) + (100)^2(0.33) + (500)^2(0.37) - 218^2 = 48{,}276$$

$$E[V] = v = \frac{7{,}716 + 36{,}100 + 48{,}276}{3} = 30{,}697.33$$

(c)

$$a = Var[M] = E[M^2] - E[M]^2$$

$$= \frac{22^2 + 130^2 + 218^2}{3} - 123.33^2$$

$$= 21636 - 123.33^2 = 6{,}425.71$$

(d)

$$\bar{X}_i = \frac{500 + 100 + 100}{3} = 233.33$$

$$Z = \frac{n}{n + \frac{v}{a}} = \frac{3}{3 + \frac{30{,}697.33}{6{,}425.71}} = \frac{3}{3 + 4.77727} = 0.3857$$

The credibility premium is

$$\Rightarrow P^{cred} = Z\bar{X}_i + (1 - Z)\mu$$

$$= 0.3857(233.33) + 0.6143(123.33) = 165.77$$

(e) The posterior probabilities for each policy class are as follows. *K* is a constant of proportionality.

$$\pi(A) = \frac{1}{3}\big((p(500|\text{class } A) \times p(100|\text{class } A)^2\big)K$$

$$= \frac{0.03 \times 0.07^2 \times K}{3} = 0.000049K$$

$$\pi(B) = \frac{0.20 \times 0.30^2 \times K}{3} = 0.00600K$$

$$\pi(C) = \frac{0.37 \times 0.33^2 \times K}{3} = 0.013431K$$

$$\pi(A) + \pi(B) + \pi(C) = 1 \Rightarrow K = 51.3347$$

$$\Rightarrow \pi(A) = 0.002515; \quad \pi(B) = 0.308008; \quad \pi(C) = 0.689476$$

$$\Rightarrow P^{Bayes} = 0.002515(22) + 0.308008(130) + 0.689476(218) = 190.40$$

(f)

(i) When *v* increases, *Z* decreases. This is reasonable, because a larger process variance indicates more uncertainty in the data from the individual risk, and therefore less credibility assigned to the individual risk.

(ii) When *a* increases, Z also increases. This is reasonable because a larger *a* indicates more uncertainty in the collective estimate of the mean, indicating that we should place less weight on the collective and more weight on the individual risk experience.

(iii) When the number of years, *n,* increases, Z also increases. This is reasonable, as more data from the individual risks indicates more information, which allows a higher credibility for the individual risks relative to the collective.

Examiners' Comments

a) Candidates did very well on this part, and most earned full credit

b) Candidates did very well on this part, and many earned full credit

c) Candidates did well on this part. A few candidates did not know how to interpret a and used an incorrect or overly complicated formula to calculate variance.

d) Candidates did well on this part; the most common error was not completing the final step with the premium calculation.

e) Many candidates struggled with or omitted this part and had difficulty interpreting the Bayesian estimate. The most common error was not calculating the correct posterior probability.

f) Most candidates were able to correctly identify the direction of the impacts to the credibility factor but were unable to explain why. Solely referencing the formula for z was not sufficient justification for why the change was reasonable.

In general, candidates should remember to show their work; simply stating that a calculation was done in Excel or their calculator without explanation is not considered showing work. Please remember that both the work and the answer is graded and partial credit can only be given for work shown. For questions 2 through 6, only the work on the written paper is graded. Work done in Excel is not reviewed.

Question 6

(a) We have that, conditional on $\Lambda = \lambda$, the distribution of X is exponential, i.e.

$$S_{X|\Lambda}(x) = e^{-\Lambda x}$$

Also, the density function of Λ is $f_\Lambda(\lambda) = \frac{\gamma^\alpha \lambda^{\alpha-1} e^{-\gamma\lambda}}{\Gamma(\alpha)}$

(from the formula sheet, with $\gamma = \theta^{-1}$).

Then:

$$S_X(x) = E\left[S_{X|\Lambda}(x)\right] = \int_0^\infty e^{-\lambda x} f_\Lambda(\lambda)\, d\lambda$$
$$= \int_0^\infty e^{-\lambda x} \frac{\gamma^\alpha \lambda^{\alpha-1} e^{-\gamma\lambda}}{\Gamma(\alpha)}\, d\lambda$$

(Now multiply and divide by $(\gamma + x)^\alpha$)

$$= \int_0^\infty \frac{e^{-\lambda(\gamma+x)} \lambda^{\alpha-1} (\gamma + x)^\alpha}{\Gamma(\alpha)}\, d\lambda \left(\frac{\gamma^\alpha}{(\gamma + x)^\alpha}\right) = \left(\frac{\gamma}{\gamma + x}\right)^\alpha$$

As the integrand is the density function of a gamma distribution for λ, with parameters α and $\gamma + x$, which must integrate to 1.0.

Alternatively,

We have that, conditional on Λ = λ, the distribution of X is exponential, i.e.

$S_{X|\Lambda}(x) = e^{-\Lambda x}$.

Also, the distribution of Λ is Gamma(α, γ^{-1})

(from the formula sheet, with $\theta = \gamma^{-1}$).

Then:

$$S_X(x) = E_\Lambda\left[S_{X|\Lambda}(x)\right] = E_\Lambda\left[e^{-\Lambda x}\right] = M_\Lambda(-x)$$

$$= \left(1 + \frac{x}{\gamma}\right)^{-\alpha} = \left(\frac{\gamma}{\gamma + x}\right)^\alpha$$

(b) Let Q denote the 97.5% quantile of X.

(i) $S_X(Q) = 0.025 \Rightarrow \left(\frac{\hat{\gamma}}{\hat{\gamma}+\hat{Q}}\right)^{\hat{\alpha}} = 0.025$

$\Rightarrow \hat{Q} = \hat{\gamma}\left((0.025)^{-1/\hat{\alpha}} - 1\right) = 9497.34$

(ii) From the formula sheet:

$$E[S - Q|S > Q] \simeq \left(\frac{\lambda + Q}{\alpha - 1}\right)$$

$$\Rightarrow E[S|S > Q] \simeq \hat{Q} + \left(\frac{\hat{\lambda} + \hat{Q}}{\hat{\alpha} - 1}\right) = 14{,}583.3$$

(c) Using the normalized distribution of the n-sample maximum, we have:

$$\begin{aligned}
\lim_{n\to\infty} Pr\left(\frac{M_n - d_n}{c_n} \leq x\right) &= \lim_{n\to\infty} Pr[M_n \leq c_n x + d] \\
&= \lim_{n\to\infty} Pr[max(X_1, X_2, \ldots, X_n) \leq (c_n x + d_n)] \\
&= \lim_{n\to\infty} \left(F_X(c_n x + d_n)\right)^n \\
&= \lim_{n\to\infty} \left(1 - \left(\frac{\gamma}{\gamma + \gamma n^{1/\alpha} x - \gamma}\right)^{\alpha}\right)^n \\
&= \lim_{n\to\infty} \left(1 - \left(\frac{1}{n^{1/\alpha} x}\right)^{\alpha}\right)^n \\
&= \lim_{n\to\infty} \left(1 - \frac{x^{-\alpha}}{n}\right)^n \\
&= \exp(-x^{-\alpha})
\end{aligned}$$

And since $exp(-x^{-\alpha})$ is the distribution function of the Fréchet distribution, the Pareto distribution is confirmed to be in the Fréchet MDA.

Alternatively:

It is sufficient to prove that $\lim_{n\to\infty} nS(c_n x + d_n) = -\log H(x)$

Where $H(x)$ is a GEV distribution function with parameter $\xi > 0$.

We have

$$nS(c_n x + d_n) = n\left(\frac{\gamma}{\gamma + \gamma n^{1/\alpha}x - \gamma}\right)^{\alpha}$$
$$= x^{-\alpha}$$
$$= -\log H(x)$$

where $H(x)$ is a GEV distribution with parameters $\xi = \frac{1}{\alpha},\ \mu = 1,\ \theta = \xi$
OR where $H(x)$ is a Frechet distribution with parameters $\alpha, \mu = 0,\ \theta = 1$

Since $\xi > 0$ this is a Fréchet distribution, as required.

(d) Substituting into the formula sheet equations, we have:

(i) $Q_q = d + \frac{\beta}{\xi}\left(\left(\frac{S_X(d)}{1-q}\right)^{\xi} - 1\right)$

$$\approx 8900 + \frac{1595.6}{0.85}\left(\left(\frac{17/500}{0.025}\right)^{0.85} - 1\right) = 9460.71$$

(ii) $ES_q = \frac{1}{1-\xi}\left(Q_q + \beta - \xi d\right)$

$$= \frac{1}{0.15}(9460.7 + 1595.6 - 0.85(8900)) = 23{,}275.3$$

(e)

- We notice that in this case, the frailty/Pareto gives a result which is very close to the GPD model for the VaR, but significantly underestimates the ES compared with the GPD.
- The MLE fit may be very good in the middle of the distribution, but may not capture the tails adequately.
- The ES uses the full tail of the distribution; the VaR ignores the top 2.5%. The ES is therefore a better measure of the extreme tail risk.
- The right tail of the loss distribution is critical for assessing reinsurance requirements.
- Based on the difference between the two ES values, the frailty/Pareto model may not adequately represent the tail risk. Using the extreme value theoretic approach typically provides a more tailored fit of tail behaviour than the MLE.
- Therefore, I would recommend basing the reinsurance strategy on the EVT values rather then the MLE.

Examiners' Comments

Overall, this question was not done well.

For Part (a), many candidates started from the unconditional survival function of X being the moment generating function of the frailty random variable by replacing x by -x. However, as a proof question, candidates are expected to show why this is valid.

Candidates did very well in Part (b).

Part (c) was a challenging question and was poorly done. Candidates were expected to identify the corresponding parameters of the GEV distribution and explain their relationships with the limiting distributions of normalized block maxima. We noticed that many numerical errors prevented candidates from identifying the proper parameters.

Candidates did very well in Part (d).

In Part (e), many candidates provided very generic answers rather than examine how the calculations done in the previous parts can contribute to this question. The question asks candidates to compare the fitting outcomes of the fat-tailed frailty model and the GPD distribution, in terms of the tail. However, a fair number of candidates provided unrelated answers.