ASTAM October 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 October 2024 Model Solutions

Question 2

(a) In 000's, the payouts under the different deductibles are

D = 250: $(95,120-150)-(20,000+6,000+2,000)(0.25)=87,970$

D = 500: $(90,000+4,320)-(20,000+6,000)(0.50)=81,320$

D = 1000: $90,000-20,000(1.0)=70,000$

(i) The deductible relativity of 250:500 is $\dfrac{87,970}{81,320}=1.082$

(ii) The deductible relativity of 1000:500 is $\dfrac{70,000}{81,320}=0.861$

(iii) Different deductible levels are likely associated with different wealth levels – wealthier policyholders will tend to self-insure to a higher level than less wealthy policyholders. Wealthy policyholders also tend to have higher ground up losses as they drive more expensive cars, and hence the collision coverage would tend to be more expensive.

(b) In millions, the payouts under the different limits are:

$L = 0.25$: $180+1000(0.25)=430.00$

$L = 1.0$: $180+361.25+150(1)=691.25$

$L = 5.0$: 728.75

So the ILFs, relative to the L=0.25 limit, are

(i) $L = 1.0$: $\dfrac{691.25}{430.00}=1.608$

(ii) $L = 5.0$: $\dfrac{728.75}{430.00}=1.695$

(c)

- The insurer may not know the total amount of the claim. If the claim exceeds the upper limit, the amount above the upper limit may not be reported to the company.
- The amount of the loss may be influenced by the amount of the limit. Attorneys often sue for the amount of the limit.
- Adverse or self-selection can occur with policy limits, which means that changing the limit might also change the ground-up loss distribution.

Examiners' Comments

Overall, this question was very poorly done. Mathematically, this question was pretty simple. Candidates may have been expecting higher difficulty, thus they appeared to be overthinking the question

Many mistakes were likely made due to hasty reading and interpreting of the question and table. Spending a few extra minutes to read and understand the table would result in fewer minor mistakes.

There was some confusion with the table buckets and interpretation of the data.

For Part B, many candidates were still thinking in terms of deductible, and did not shift to think about policy limits.

ASTAM October 2024 Model Solutions

Question 3

(a)

(i) Let X_j, N, and S denote the jth loss severity, the number of losses, and the aggregate loss, respectively. Then

$$E\left[X_j\right]=e^{7+.5^2/2}=1242.65$$
$$E[N]=0(0.5)+1(0.3)+2(0.1)+3(0.1)=0.8$$
$$\Rightarrow E[S]=0.8\times 1242.65=994.12$$

(ii) We know that:

$$\text{Var}[S]=\text{E}\left[\text{Var}\left[S|N\right]\right]+\text{Var}\left[\text{E}\left[S|N\right]\right]$$
$$\text{where Var}\left[S|N\right]=N\text{Var}\left[X_j\right]=N\left(\left(e^{7+0.5^2/2}\right)^2\left(e^{0.5^2}-1\right)\right)$$
$$438,584.8N$$

$$\Rightarrow \text{E}\left[\text{Var}\left[S|N\right]\right]=438,584.8(0.8)=350,867.8$$

$$\text{and Var}\left[\text{E}\left[S|N\right]\right]=\text{Var}\left[N\left(\text{E}\left[X_j\right]\right)\right]$$
$$=\left(\text{E}\left[X_j\right]\right)^2\text{Var}[N]$$

$$\text{Var}[N]=1^2(0.3)+2^2(0.1)+3^2(0.1)-0.8^2=0.96$$

$$\Rightarrow \text{Var}\left[\text{E}\left[S|N\right]\right]=1242.6^2(0.96)=1,482,407.5$$

$$\Rightarrow \text{Var}[S]=350,867.8+1,482,407.5=1,833,275.3=1354.0^2$$

(iii) We use the normal approximation when there are a substantial number of expected losses (typically >30), such that the central limit theorem becomes relevant.

In the example above, with an expected frequency of less than 1 loss, and with a positively skewed severity distribution, the compound distribution will also be substantially positively skewed, and the normal approximation will not provide a good fit.

(b) The expected amount of the *j*th loss that is not covered under the policy is

$$\mathrm{E}\left[X_j \wedge d\right] = \mathrm{E}\left[X_j\right]\ \Phi\left(\frac{\log d - \mu - \sigma^2}{\sigma}\right) + d\left(1 - \Phi\left(\frac{\log d - \mu}{\sigma}\right)\right)$$

$$= 1242.65\ \Phi(-2.071) + 500\left(1 - \Phi(-1.571)\right) = 494.79$$

Note: The formula for $E\left[X_j \wedge d\right]$ is given in the formula sheet.

(c)

$$\mathrm{E}[I_j] = \Pr[X_j > d] = S(500) = 1 - F(500) = 1 - \Phi\left(\frac{\log(500) - 7}{0.5}\right) = 0.9419 = v$$

$$P_{N_P}(z) = \mathrm{E}\left[z^{N_p}\right] = \mathrm{E}\left[\mathrm{E}\left[z^{N_P} | N_L\right]\right]$$
$$= \mathrm{E}\left[\mathrm{E}\left[z^{I_1 + I_2 + \ldots + I_{N_L}} | N_L\right]\right]$$
$$= \mathrm{E}\left[\left(\mathrm{E}\left[z^{I_j}\right]\right)^{N_L}\right] (\text{as } I_j \text{ are iid})$$
$$= \mathrm{P}_{N_L}\left(\mathrm{E}\left[z^{I_j}\right]\right) (\text{from definition of PGF})$$

$$\mathrm{E}\left[z^{I_j}\right] = z^0(1 - v) + z^1 v = 1 - v(1 - z)$$

$$\Rightarrow P_{N_P}(z) = P_{N_L}\left(1 - v(1 - z)\right)$$

(d) Using PGFs:.

$$P_{N_L}(z) = 0.5z^0 + 0.3z + 0.1z^2 + 0.1z^3$$

$$\Rightarrow P_{N_P}(z) = 0.5 + 0.3(1 - v(1 - z)) + 0.1(1 - v(1 - z))^2 + 0.1(1 - v(1 - z))^3$$

(i) $\Pr[S=0] = \Pr\left[N_p = 0\right] = P_{N_P}(0) = P_{N_L}(1 - v)$

$= 0.5 + 0.3(1 - v) + 0.1(1 - v)^2 + 0.1(1 - v)^3$

$= 0.51779$

(ii) $\Pr[N_P = 1] = P'_{N_P}(z=0) = vP'_{N_L}(1 - v)$

$= v\left(0.3 + 0.2(1 - v) + 0.3(1 - v)^2\right)$

$= 0.29447$

OR

From first principles. Let $p_k = \Pr[N=k]$:

(i) $\Pr[S=0] = p_0 + p_1(1 - v) + p_2(1 - v)^2 + p_3(1 - v)^3 = 0.51779$

(ii) $\Pr\left[N_p = 1\right] = p_1 v + 2p_2 v(1 - v) + 3p_3 v(1 - v)^2 = 0.29447$

(e)

(i)

$$X_j^* = \begin{cases} X_j - 2500 & X_j > 2500 \quad \leftarrow 2500 \textit{ is insurer oayment plus deductible} \\ 0 & X_j \le 2500 \end{cases}$$

$$\mathrm{E}\left[X_j^*\right] = \mathrm{E}\left[X_j\right] - \left(\mathrm{E}\left[X_j \wedge 2500\right]\right)$$

$$\mathrm{E}\left[X_j \wedge 2500\right] = e^{\mu + \sigma^2/2} \Phi\left(\frac{\ln(2500) - \mu - \sigma^2}{\sigma}\right) + 2500\left(1 - \Phi\left(\frac{\ln(2500) - \mu}{\sigma}\right)\right) = 1210.91$$

$$\Rightarrow G = 1.2E[N]\left(\mathrm{E}\left[X_j\right] - \left(\mathrm{E}\left[X_j \wedge 2500\right]\right)\right) = 1.2(0.8)(1242.65 - 1210.91) = 30.472$$

OR

$$E\left[X_j^*\right] = \int_{2500}^{\infty} x f(x)\, dx - 2500(1 - F(2500))$$

$$= e^{\mu + \sigma^2/2}\left(1 - \Phi\left(\frac{\ln(2500) - \mu - \sigma^2}{\sigma}\right)\right) - 2500\left(1 - \Phi\left(\frac{\ln(2500) - \mu}{\sigma}\right)\right)$$

$$= 1242.65(0.12547) - 2500(0.04967) = 31.742$$

$$\Rightarrow G = 1.2(0.8)\, 31.742 = 30.472$$

(ii) From the definition of X_j^*, we see that if X_j increases by 10%, but the 2500 limit stays the same, then X_j^* will increase by more than 10%, so the premium will also increase by more than 10%.

OR

$E[X_j]$ will increase by 10%, but $E[X_j \wedge 2500]$ will increase by less than 10%. Effectively, the retention point has decreased in real terms, so the excess cost over the retention has increased in real terms.

Examiners' Comments

Part A

Candidates did not do well. Most candidates understood how to calculate the mean and standard deviation of aggregate loss. However, the majority candidates didn't realize that they should work on ground-up losses and tried to solve a much harder question on aggregate payments after applying deductible. Candidates might have misread questions if they find that it takes extraordinarily longer time to solve than the assigned points.

Part B

Candidates overall did well. Some candidates calculated the ratio of the expected insurance payment (rather than the expected loss eliminated by deductible) over the expected ground-up losses. If the candidate did part of the work in Part A, it would be very helpful to point out here for exam graders.

Part C

Candidates did poorly in Part C. Candidates should understand important relations among basic concepts (i.e., the number of ground-up losses and the number of payments) rather than just remember the formulas.

Part D

Candidates did fairly well. Most candidates understood how to solve the problem using the first principles, although some pieces were missed.

Part E

Candidates did well. A common mistake is that the reinsurer covers losses above 2500 rather than 2000 due to the deductible.

ASTAM October 2024 Model Solutions

Question 4

(a)

(i)

$$h(x) = -\frac{d\log S(x)}{dx};$$

$$\log S(x) = \left(\frac{1}{\theta}\right)(\log\lambda - \log(\lambda + \theta x))$$

$$\Rightarrow h(x) = \left(\frac{1}{\theta}\right)\left(\frac{\theta}{\lambda + \theta x}\right) = \frac{1}{\lambda + \theta x}$$

OR

$$h(x) = \frac{f(x)}{S(x)};$$

$$f(x) = -\frac{d}{dx}S(x) = \lambda^{\frac{1}{\theta}}(\lambda + \theta x)^{-\frac{1}{\theta} - 1} = \left(\frac{\lambda}{\lambda + \theta x}\right)^{\frac{1}{\theta}}\left(\frac{1}{\lambda + \theta x}\right)$$

$$\Rightarrow h(x) = \frac{\left(\frac{\lambda}{\lambda + \theta x}\right)^{\frac{1}{\theta}}\left(\frac{1}{\lambda + \theta x}\right)}{\left(\frac{\lambda}{\lambda + \theta x}\right)^{\frac{1}{\theta}}} = \left(\frac{1}{\lambda + \theta x}\right)$$

(ii) We see that the hazard rate is a decreasing function of x which indicates a fat-tailed distribution.

(b)

ES Advantage: It is a coherent risk measure, which means it is less easy to game, i.e. does not incentivize risk seeking behavior.

ES Advantage: It takes into consideration how large the rare tail losses may be. These do not impact the VaR.

ES Disadvantage: for some fat tailed distributions the ES may not be defined.

ES Disadvantage: requires more information to calculate; VaR can be calculated without information on the far right tail of the loss distribution.

(c)

(i) $S(Q) = 0.01 \Rightarrow \left(\frac{2000}{2000 + 0.4Q}\right)^{2.5} = 0.01$

$\Rightarrow 2000 + 0.4Q = 2000(0.01)^{-0.4}$

$\Rightarrow Q = \frac{2000}{0.4}\left(0.01^{-0.4} - 1\right) = 26,547.9$

(ii) $ES_\alpha = Q_\alpha + e\left(Q_\alpha\right) = Q_\alpha + \frac{\lambda + \theta Q_\alpha}{1-\theta} = \frac{Q_\alpha + \lambda}{1-\theta}$

Set this equal to26, 547.9from (i)

$26{,}547.9 = \frac{Q_\alpha + \lambda}{1-\theta} \Rightarrow Q_\alpha = 26,547.9(0.6) - 2000 = 13,928.7$

To find α, set $S\left(Q_\alpha\right) = 1 - \alpha$

$\Rightarrow 1 - \alpha = \left(\frac{2000}{2000 + 0.4(13,928.7)}\right)^{2.5} \Rightarrow \alpha = 0.9641$

(iii) Note that both Q_α and ES_α are increasing functions of α.
Also, for continuous loss distributions, ES_α is the expected value of loss given that it is greater than Q_α, so $ES_\alpha \geq Q_\alpha$
Hence, if $ES_\alpha = Q_\beta$ for some α and β, then $Q_\beta \geq Q_\alpha \Rightarrow \beta \geq \alpha$

(d)

(i) $$nS(c_n x + d_n) = n\left(\frac{\lambda}{\lambda + \theta\left(c_n x + d_n\right)}\right)^{\frac{1}{\theta}} = \left(\frac{n^\theta \lambda}{\lambda + \theta\left(c_n x + d_n\right)}\right)^{\frac{1}{\theta}}$$

$$= \left(\frac{\lambda + \theta c_n x + \theta d_n}{n^\theta \lambda}\right)^{-\frac{1}{\theta}} = \left(\frac{\lambda + \theta d_n}{n^\theta \lambda} + \frac{c_n}{n^\theta \lambda}(\theta x)\right)^{-\frac{1}{\theta}}$$

For this to be equal to $(1 + \theta x)^{-\frac{1}{\theta}}$, we need

$$\frac{c_n}{\lambda n^\theta} = 1 \Rightarrow c_n = \lambda n^\theta$$

and $$\frac{\lambda + \theta d_n}{\lambda n^\theta} = 1 \Rightarrow d_n = \frac{\lambda(n^\theta - 1)}{\theta}$$

(ii) This result shows that the distribution is in the MDA of the GEV distribution, with parameter $\xi = \theta$; if $\theta > 0$ this is the Fréchet distribution, which is very fat-tailed.

Examiners' Comments:

Part A

Candidates did very well on part A. The biggest error was incorrectly determining the derivative of the $\log S(x)$ *or the derivative of* $S(x)$. *Additionally, to get full credit on a "show" question, you must actually show the steps which lead to the given answer. It is best to show too much work than not enough.*

Part B

Candidates also did well on this part although many candidates did not state that ES was a coherent risk measure.

Part C

Many candidates earned full credit on part (i) but many candidates did not earn credit on part (ii) as their work showed that they did not really understand the concepts and calculation of ES. Part (iii) was done acceptably.

Part D

Candidates who attempted part (i) did quite well on this part. Also, many candidates were able to describe the significance of the results in part (ii).

ASTAM October 2024 Model Solutions

Question 5

(a) $\mu = E[\mu(\Theta)]$; $\quad v = E[v(\Theta)]$; $\quad a = Var[\mu(\Theta)]$

(b) Note that $X|\beta$ has a Geometric distribution with parameter β so that:

$\Theta = \beta,\ \mu(\Theta) = E[X|\beta] = \beta, v(\theta) = Var[X|\beta] = \beta(1+\beta)$

(i) $\mu = E[\mu(\theta)] = E\big[E[X|\beta]\big] = E[\beta]$

(ii) $a = Var\big[E[X|\beta]\big] = Var[\beta]$

(c)

(i) $Var(X) = E\big[Var[X|\beta]\big] + Var\big[E[X|\beta]\big]$

$= E[\beta + \beta^2] + Var[\beta]$

$= E[\beta] + E[\beta^2] + Var[\beta]$

$= E[\beta] + (Var[\beta] + E[\beta]^2) + Var[\beta]$

$= 2Var[\beta] + E[\beta] + E[\beta]^2$

$= 2Var[\beta] + E[\beta](1 + E[\beta])$

(ii) $a = Var[\beta]\ and\ E[\beta] = E[X]$

So $Var(X) = 2a + E[X](1 + E[X])$

So that $a = \frac{Var(X) - E[X](1+E[X])}{2}$

(d)

(i) $\hat{\mu} = \bar{X} = \frac{(0)(2230)+1(209)+2(41)+3(12)+4(8)}{2500} = 0.1436$

(ii) $s^2 = \frac{(0^2)(2230)+1^2(209)+2^2(41)+3^2(12)+4^2(8)}{2500} - (0.1436)^2\left(\frac{2500}{2499}\right) = 0.22307$ so that the standard deviation is $\sqrt{0.22307} = 0.4723$

(iii) $\hat{a} = \frac{0.22307 - 0.1436(1.1436)}{2} = 0.02942$

(iv) $\hat{v} + \hat{a} = s^2 \quad so\ that\ \ \hat{v} = s^2 - \hat{a} = 0.22307 - 0.02942 = 0.193645$

Alternatively, $\hat{v} = E[X](1 + E[X]) + \hat{a} = (0.1436)(1.1436) + 0.02942 = 0.19364$

(e)

(i) $Z = \frac{1}{1+\frac{\hat{v}}{\hat{a}}} = \frac{1}{1+\frac{0.193645}{0.02942}} = 0.1319$

$$P = Z \times \bar{X} + (1 - Z) \times \hat{\mu} = 0.1319(1) + (1 - 0.1319)(0.1436) = 0.2566$$

(ii) $Z = \frac{5}{5 + \hat{v}/\hat{a}} = \frac{5}{5 + \frac{0.193645}{0.02942}} = 0.4317$

$$P = Z \times \bar{X} + (1 - Z) \times \hat{\mu} = 0.4317(1) + (1 - 0.4317)(0.1436) = 0.5133$$

(iii) A claim frequency of 1 per year is much higher than the portfolio average, of around 1 every 7 years. After only 1 year, this claim frequency is not given much weight, it might just be bad luck. With 5 claims in 5 years, the high claim frequency now looks more systematic, indicating that this driver is particularly high risk. More data means higher weight/credibility on their individual experience, lass on the prior.

Examiners' Comments:

Parts (a) and (b) were generally done well by candidates. The only common error was to mistakenly assume that β had a Geometric distribution, rather than $X|\beta$. Some candidates lost credit in part (b) by simply writing the result rather than showing any steps or assumptions.

Some candidates omitted one or both parts of (c), but those who attempted it generally did fairly well. Most candidates were able to write down the variance decomposition formula and substitute in the correct values for the mean and variance of $X|\beta$. The most common mistake was to confuse $E[\beta^2]$ with $E[\beta]^2$ or confuse $E[\beta(1+\beta)]$ with $E[\beta](1 + E[\beta])$. On both subparts, some candidates failed to show enough work for full credit.

Part (i) of (d) was in general done very well, as was, to a lesser extent, part (ii). The most common errors for (ii) were giving the variance but not the standard deviation, and not showing sufficient work. No points were deducted for omitting the factor of (2500/2499). For full credit, it is expected to show any formulas used. It is insufficient to write "Used calculator" or "Used Excel".

For (d) (iv), many candidates tried to use $\nu = E[\beta(1+\beta)]$, but approximated this quantity as $\hat{v} = \hat{\beta}(1 + \hat{\beta}) = 0.1436(1.1436)$.

A significant minority of candidates omitted part (e), though the candidates who attempted it generally did well on this part.

Some candidates misunderstood the data to be one claim total over five years rather than one claim in each of the last five years (five claims total).

ASTAM October 2024 Model Solutions

Question 6

(a)

(i) The business is short-tailed, as around 98% of claims are settled within 4 years, and all claims are settled within around 6 years.

(ii)

$$\widehat{C}_{6,7} = 9182\left(\prod_{j=1}^{7} f_j\right) = 9182(1.99147) = 18,285.7$$

$$\widehat{R}_6 = \widehat{C}_{6,7} - 9,182 = 9103.7$$

(b)

(i)

$$\widetilde{C}_{6,7} = \widehat{\beta}_1 \widehat{C}_{6,7} + \left(1 - \widehat{\beta}_1\right)\mu_6$$

$$\widehat{C}_{6,7} = 18,285.7 \text{(from (a))}$$

$$\widehat{\beta}_1 = \frac{1}{\widehat{\lambda}_1} = \frac{1}{1.99147} = 0.502141$$

$$\mu_6 = 0.85(18,000) = 16,200$$

$$\Rightarrow \widetilde{C}_{6,7} = 0.502141(18,285.7) + 0.497849(16,200) = 16,799$$

$$\Rightarrow \widetilde{R}_{6,7} = 16,799 - 9,182 = 7,617$$

(ii) Advantage: Less weight placed on very early claims information, i.e. from the latest accident years.

Disadvantage: The value for μ is rather arbitrary. Using the loss ratio method, if the premiums are too low, then the OCR will also be too low.

(c)

(i) A1. and A3. are the same as the Chain Ladder assumptions;. A2. and A4 are specific to the Mack Model.

(ii) We have that $\mathrm{E}\left[C_{i,j+1} \middle| C_{i,j}\right] = f_j \; C_{i,j}$ (A3) and $\mathrm{Var}\left[C_{i,j+1} \middle| C_{i,j}\right] = \sigma_j^2 \; C_{i,j}$ (A4).

For $C_{i,j+2}$ we condition on $C_{i,j+1}$ (as well as $C_{i,j}$) and then use (A3) and (A4), for the moments of $C_{i,j+2}$ given $C_{i,j+1}$.

(Note that given $C_{i,j+1}$, we gain no extra information from $C_{i,j}$ because the process is Markov (A2).)

That is $\mathrm{Var}\left[C_{i,j+2} \middle| C_{i,j}\right] = \mathrm{E}\left[\mathrm{Var}\left[C_{i,j+2} \middle| C_{i,j+1}, C_{i,j}\right]\right] + \mathrm{Var}\left[\mathrm{E}\left[C_{i,j+2} \middle| C_{i,j+1}, C_{i,j}\right]\right]$

Now $\mathrm{Var}\left[C_{i,j+2} \middle| C_{i,j+1}, C_{i,j}\right] = \sigma_{j+1}^2 C_{i,j+1}$ (A4)

and $\mathrm{E}\left[C_{i,j+2} \middle| C_{i,j+1}, C_{i,j}\right] = f_{j+1} C_{i,j+1}$ (A3)

So $E\left[\mathrm{Var}\left[C_{i,j+2} \middle| C_{i,j+1}, C_{i,j}\right]\right] = E\left[\sigma_{j+1}^2 C_{i,j+1} | C_{i,j}\right] = \sigma_{j+1}^2 f_j \, C_{i,j}$ (A3)

and $\mathrm{Var}\left[E\left[C_{i,j+2} \middle| C_{i,j+1}, C_{i,j}\right]\right] = Var\left[f_{j+1} C_{i,j+1} | C_{i,j}\right] = f_{j+1}^2 \, \sigma_j^2 \, C_{i,j}$ (A4)

$\Rightarrow \mathrm{Var}\left[C_{i,j+2} \middle| C_{i,j}\right] = \sigma_{j+1}^2 f_j \, C_{i,j} + f_{j+1}^2 \, \sigma_j^2 \, C_{i,j}$

$= C_{i,j}\left(\sigma_{j+1}^2 f_j + f_{j+1}^2 \, \sigma_j^2\right)$

(d)

(i)

- If the Mack Model assumptions are valid, the residuals should be independent and approximately N(0,1) distributed.
    - There are 26 residuals, so we would expect around 16% of them, i.e. around 4, to be greater than 1.0 in absolute value (as $\Phi(1.0) = 0.84$).
    - The residuals plots show that around 10 of the residuals are less than –1.0 or greater than 1.0, which is very unlikely if the residuals are N(0,1) distributed.
    - This is evidence that the data are not consistent with the model.
- On the other hand, all the residuals are within 2 standard deviations of the mean, which is consistent with the expectation that only 0.05(26)=0.3 residuals would have absolute value greater than 2.
- There appears to be an increasing trend of residuals by AY and CY, indicating that the observed values are systematically higher than the fitted values for later AYs, and also for later CYs. This points to a possible calendar year effect, which is inconsistent with the model assumptions.
- There is no systematic residual trend by DY, indicating that the Markov assumption is OK.

(ii) They are the same., (assuming that the development factors are estimated the same way, i.e. minimum variance).

(e)

(i) The AY's are only conditionally independent, conditional on knowing the f_j and σ_j parameters.

The total MSEP is the sum of the process variance and the (squared) estimation error. The process variance is additive by AY.

The estimation error involves uncertainty in the estimated $\hat{C}_{i,J}$ values, i.e. uncertainty from parameter estimation. Since the parameters are the same for each AY, these errors are not independent, and the total MSEP must allow for the contribution of covariances between AYs in the estimation error.

(ii) I would recommend the Mack Model estimate because

1. It is fully empirical, that is, it involves no subjective inputs, unlike B-F where the μ parameter is subjective.
2. It is higher than the B-F estimate, so is a more safe-side selection.

Reasons for not using Mack:

1. The output shows that the data may not comply with the MM assumptions.

Examiners' Comments

Part A

Even though many candidates technically knew the definitions of short-tailed and long-tailed insurance, they applied the examples in the textbook dogmatically to decide that the business was long-tailed, instead of applying own their judgment. The calculation in part (ii) was well answered by most candidates.

Part B

The calculation part was well answered by most candidates. Many candidates offered valid alternatives solutions. For part (ii), many candidates were able to provide an advantage. For the disadvantage, many did not clearly make the point that is arbitrary.

Part C

Many candidates scored well here. More candidates were comfortable with the chain ladder assumptions than with the Mack model assumptions. Only the best candidates could complete the proof, which came directly from the source material.

Part D

Few candidates could interpret the residual graphs. Many candidates made general comments, but failed to refer to a specific graph. Graders were flexible in awarding points for different interpretations of the plots in part (i). For part (ii) it is sufficient to note that the estimates are the same (no need to include the comment in parentheses).

Part E

For part (i), few candidates grasped the main point, which was dependency through parameter uncertainty.

For part (ii), either answer was acceptable if justified. Many candidates were able to make and justify their recommendations.