Advanced Short Term Actuarial Mathematics

Fall 2025

Model Solutions and Examiners' Comments

These solutions and examiners' comments are intended to be used by candidates preparing for future ASTAM exams. In many places they are more detailed than would be required for full credit under examination conditions. In some cases there also be valid alternative methods or derivations that are not included in these solutions, but that were given appropriate credit in grading.

Question 1

See Excel workbook for solutions and comments

Question 2

(a) (*2.5 points*) The log-likelihood (LL) is $l(r,\beta) = \sum_{k=0}^{3} n_k \log p_k$.

The maximum value of the LL is the LL evaluated using the MLE parameter. That is, given $r = 2$ and $\hat{\beta} = 0.21825$, we have:

$$p_0 = \left(\frac{1}{1+\hat{\beta}}\right)^r = 0.67379; \qquad p_1 = r\left(\frac{1}{1+\hat{\beta}}\right)^r \left(\frac{\hat{\beta}}{1+\hat{\beta}}\right) = 0.24142$$

$$p_2 = \frac{r(r+1)}{2}\left(\frac{1}{1+\hat{\beta}}\right)^r \left(\frac{\hat{\beta}}{1+\hat{\beta}}\right)^2 = 0.06488$$

$$p_3 = \frac{r(r+1)(r+2)}{6}\left(\frac{1}{1+\hat{\beta}}\right)^r \left(\frac{\hat{\beta}}{1+\hat{\beta}}\right)^3 = 0.01550$$

$\Rightarrow$ max LL $= 1325 \times \log 0.67379 + \cdots + 39 \times \log 0.01550 = \mathbf{-1747.25}$

(b) (*2.5 points*) The table for calculating the test statistic is

k	p_k	O	E	$(O-E)^2/E$
0	0.67379	1325	1347.59	0.379
1	0.24142	516	482.84	2.277
2	0.06488	120	129.75	0.732
3	0.01550	39	30.99	2.068
≥4	0.00441	0	8.82	8.824

(i) The test statistic is $T = 0.379 + \cdots + 8.824 = \mathbf{14.2814}$.

(ii) Degrees of freedom = 5-1-1 =3.

$\Rightarrow$ critical value $= Q_{5\%}$: $\Pr\left[\chi_3^2 > Q_{5\%}\right] = 0.05$

$\Rightarrow Q_{5\%} = \mathbf{7.81}$ (Excel: CHISQ.INV.RT(0.05,3))

(iii) The null hypothesis is that the data are consistent with the NB$\left(r,\hat{\beta}\right)$ distribution.

As $T > Q$ we reject the null hypothesis and conclude that the data are not consistent with the NB$\left(r,\hat{\beta}\right)$ distribution.

(c) (*3 points*)

(i) The AIC statistic is $LL_{\max} - k$ where $LL_{\max}$ is the max log likelihood, and k is the number of parameters.

Model 1: $-1747.25 - 1 = -1748.25$

Model 2: $-1744.43 - 2 = -1746.43$

Model 2 is preferred as the AIC statistic is higher.

(ii) The SBC statistic is $LL_{\max} - \dfrac{k \log n}{2}$ where $n = 2000$ is the sample size.

Model 1: $-1747.25 - 0.5\left(\log 2000\right) = -1751.05$

Model 2: $-1744.43 - 1\left(\log 2000\right) = -1752.03$

Model 1 is preferred as the SBC statistic is higher.

(iii) The LRT tests the null hypothesis that the more parsimonious model (Model 1) is preferred.

The test statistic is

$Q = 2\left(LL_{\max}(\text{model 2}) - LL_{\max}\ (\text{model 1})\right) = 5.636$

p – value is $\Pr\left[\chi^2_{\nu=1} > Q\right] = 0.018$

At 5% significance level we reject the null hypothesis and select Model 2.

(iv) The p-value for the Model 1 GOF test is $\Pr\left[\chi^2_3 > 14.28\right] = 0.25\%$

The p-value for the Model 2 GOF test is $\Pr\left[\chi^2_{\nu=2} > 12.98\right] = 0.15\%$

Model 1 is preferred based on a slightly higher p-value.

(d) (*1 point*) A zero modified (ZM) distribution might improve the fit of observed vs expected for the zero claims category, but the fit for both sets of NB parameters is already pretty good in that cell, and so improving it will not reduce the test statistic very much, and will add a parameter and therefore reduce the degrees of freedom.

Examiners' Comments

1. *The examiners expected this question to be high scoring. Most candidates did achieve a passing score on the question, but overall the performance was below expectations.*
2. *In part (a), candidates were expected to plug the given parameter values into the LL equation to find the maximum, since the MLE, by definition, is the value that maximizes the LL. Instead, many candidates started by finding $\hat{\beta}$ even though it was given in the question. This made the question much longer than it needed to be, and led to more mathematical errors. Candidates are advised that they should re-read the question as*

they work through it, to ensure that they have followed instructions and used all the given information appropriately.

3. *Part (b) was done well.*
4. *In part (c), most candidates correctly applied the AIC and SBC criteria. The Likelihood Ratio test was also fairly well done, although a significant number of candidates used the wrong distribution for the critical value. In part (c)(iv), only a few candidates recognized that a direct comparison of the test statistics was not valid for ranking the two cases, as the difference in degrees of freedom needed to be allowed for.*
5. *In part (d), the examiners were looking for a recognition from the χ^2 table that the problem with the fit was at the right tail of the distribution, not the left tail. Using a ZM distribution will reduce the degrees of freedom, without much reducing the test statistic. Many candidates omitted this part completely. Those that attempted it largely identified that the fit at 0 claims was quite good. Only a few also mentioned the reduced degrees of freedom.*

Question 3 (*9 points*)

(a) (*2 points*)

$$S_Y(y) = \Pr\left[X - d > y \mid X > d\right] = \frac{S_X(y+d)}{S_X(d)} \quad \text{for } y \le b-d$$

$$= \left(\frac{b-(y+d)}{b-a}\right)\left(\frac{b-a}{b-d}\right) = \frac{(b-d)-y}{(b-d)}$$

$$\Rightarrow Y \sim U(0, b-d)$$

(b) (*3 points*) Now $d = 2$ so $Y \sim U(0,8) \Rightarrow F_Y(y) = \frac{y}{8} \quad 0 < y \le 8$. Let $H(y)$ denote the d.f of the GEV distribution with $\xi = -1,$ i.e. $H(y) = \exp\left(-(1-y)\right)$ (for $y > 1$). Then we can use either of two approaches:

Either:

$$\lim_{n\to\infty} F_n(c_n y + d_n) = H(y)$$

$$\text{Now } F_n(y) = \Pr\left[\max(Y_1, Y_2, ..., Y_n) \le y\right] = \left(\frac{y}{8}\right)^n$$

$$\Rightarrow \lim_{n\to\infty} F_n(c_n y + d_n) = \lim_{n\to\infty}\left(\frac{c_n y + d_n}{8}\right)^n = \lim_{n\to\infty}\left(1 + \frac{nc_n y + n\left(d_n - 8\right)}{8n}\right)^n$$

$$\text{So } \lim_{n\to\infty} F_n(c_n y + d_n) = H(y) \iff \exp\left(\frac{nc_n y + n\left(d_n - 8\right)}{8}\right) = \exp\left(-(1-y)\right)$$

Equate terms in y in the exponents:

$$\frac{nc_n y}{8} = y \Rightarrow c_n = \frac{8}{n} \quad ; \quad \frac{n\left(d_n - 8\right)}{8} = -1 \Rightarrow d_n = 8\left(1 - \tfrac{1}{n}\right)$$

Or:

$$\lim_{n\to\infty} nS\left(c_n y + d_n\right) = -\log H(y) = -\log e^{-(1-y)} = 1 - y \quad (\text{for } y > 1)$$

$$\text{Now } \lim_{n\to\infty} nS\left(c_n y + d_n\right) = n\left(1 - \frac{c_n y + d_n}{8}\right)$$

$$\text{So } \lim_{n\to\infty} nS\left(c_n y + d_n\right) = -\log H(y) \Rightarrow n\left(1 - \frac{c_n y + d_n}{8}\right) = 1 - y$$

Equate terms in y^1 and y^0: $\quad -\frac{nc_n}{8} = -1 \Rightarrow c_n = \frac{8}{n}$

$$n\left(1 - \frac{d_n}{8}\right) = 1 \Rightarrow d_n = 8\left(1 - \frac{1}{n}\right)$$

(c) (*2 points*)

(*i*) $\Pr\left[M_{10} \le 7\right] = \left(\frac{7}{8}\right)^{10} = \mathbf{0.26308}$

(*ii*) $\Pr\left[M_{10} \le 7\right] \approx \Pr\left[\frac{M_{10} - d_{10}}{c_{10}} \le \frac{7 - d_{10}}{c_{10}}\right]$

$$\mathrm{c}_{10} = \frac{8}{10} = 0.8 \quad \text{and} \quad \mathrm{d}_{10} = 8\left(1 - \frac{1}{10}\right) = 7.2$$

$$\Rightarrow \frac{7 - d_{10}}{c_{10}} = -\frac{0.2}{0.8} = -0.25$$

$$\Rightarrow \Pr\left[M_{10} \le 7\right] \approx \Pr\left[H_{\xi=-1} \le -0.25\right] \quad = e^{-1.25} = \mathbf{0.28651}$$

(d) (*2 points*)

(i) As $\xi < 0$, the GEV is the EV Weibull distribution.

(ii) Weibull is bounded above. Gumbel is unbounded above.

Examiners' Comments

1. *Overall, this question was not high scoring, though around 10% of candidates did very well.*
2. *The examiners expected strong performance on the first part, which was intended to be a relatively simple conditional distribution calculation, similar to problems seen in introductory courses in probability and statistics. In fact, part (a) was done very poorly or not at all by most candidates. Many candidates tried to find E[Y], which was not relevant to the question; a distribution is uniquely defined by, e.g. its survival function, or distribution function, or pdf, but not by its mean or variance.*
3. *The result in part (a) could be derived using the distribution or density function, instead of the survival function, but the survival function approach is the shortest and simplest, which is why the examiners set the question up in terms of S(x). Candidates who tried to derive the result using pdf's often failed to specify bounds appropriately, i.e. proving that Y was uniformly distributed on some interval of length b-d, rather than specifically on 0 to b-d.*
4. *Part (b) was done well by most candidates who attempted it. However, some candidates failed to show sufficient work for a derivation question. In particular, many candidates wrote down equations for c_n and d_n, but gave little or no derivation.*
5. *Part (c)(i) was done well by most candidates who attempted it. Part (c)(ii) was more mixed, though most candidates who attempted this part received at least partial credit. If incorrect values for c_n and d_n taken from the previous part were used, candidates did not lose additional credit.*

   Again, many candidates did not show sufficient work to receive full credit.

   A minority of candidates used incorrect distributions for their approximations, such as the normal distribution.

6. *Some candidates' answers in part (c) were clearly unreasonable - typically in this case a probability that is not in [0, 1]. A candidate who writes down an impossible value and does not comment that they know it is impossible will generally receive less partial credit than a candidate who gives the same answer, but understands and comments on the fact that it cannot be correct.*
7. *Part (d) was done well by those who attempted it, though in (d)(ii) a significant number of candidates just mentioned parameter values or numbers of parameters in the distributions. Other candidates wrongly proposed that one distribution has all of its moments while the other does not. In fact, both distributions have all finite moments. In the context of extreme value theory, the fact that one distribution is right bounded while the other is unbounded is particularly significant.*

Question 4 (*11 Points*)

(a) (*3 points)*

(i) $LER = \frac{E\left[X \wedge 60\right]}{E[X]} \quad = \frac{60}{310} = \mathbf{0.1935}$

(ii) $EPPL = E\left[X\right] - E\left[X \wedge 60\right] = \mathbf{250}$

(iii) First, assume $d < 100$:

$LER = 0.4 \Rightarrow \frac{d}{310} = 0.4 \Rightarrow d = 124$ which contradicts the assumption.

Now, assume $100 < d \leq 500$:

$LER = 0.4 \Rightarrow \frac{100 \times 0.6 + d \times 0.4}{310} \Rightarrow d = \mathbf{160}$

(b) (*2 points*)

(i) Ordinary deductible, *d*:

The insured pays the full loss if it is less than *d*, and pays a flat amount of *d* for losses greater than *d*. That is, let *X* denote the loss, and *Y* denote the amount covered by insurance, for an ordinary deductible *d*,

$$Y = \begin{cases} 0 & X \leq d \\ X - d & X > d \end{cases}$$

Franchise deductible:

The insured pays the full loss if it is less than d, and the insurer pays the full loss if it is greater than d. That is,

$$Y = \begin{cases} 0 & X \leq d \\ X & X > d \end{cases}$$

(ii) One example of a franchise deductible is short term disability income where the insured covers all losses if the they are disabled for less than the elimination period (such as 7 days) but if the disability extends longer than the elimination period, then the insurance covers all losses including the elimination period,

(c) (*4 points*) Loss frequency is represented by $N \sim \text{Poisson}(\lambda = 1)$. The claim per loss distribution is now

$$X^* = \begin{cases} 0 & \text{w.p. } 0.6 \\ 400 & \text{w.p. } 0.3 \\ 800 & \text{w.p. } 0.1 \end{cases}$$

(i) Let p_k denote the probability function for the Poisson(1) loss frequency. Then

$$\Pr[S=0] = p_0 + p_1(0.6) + p_2(0.6)^2 + \ldots = \text{E}\left[0.6^N\right]$$

$$= P_N(0.6) = e^{\lambda(0.6-1)} = \mathbf{0.67032}$$

where P_N is the Poisson$(\lambda = 1)$ probability generating function.

Alternatively, let *N'* denote the number of claims, where *N* denotes losses. From the Poisson properties, $N' \sim \text{Poisson}(\lambda q)$ where $\lambda = 1$, and $q = \Pr[X^* > 0] = 0.4$. Then

$$\Pr[S=0] = \Pr[N'=0] = e^{-\lambda q} = e^{-0.4} = \mathbf{0.67032}$$

(ii) Work in units of 400, and apply the (a,b,0) recursion formula from the formula sheet.

Let $f_j = \Pr[X^* = 400j]$ for $j = 0,1,2,$ and let $g_s = \Pr[S = 400s]$ for $s = 0,1,2,\ldots$

For the Poisson distribution, $a = 0, \quad b = \lambda$, so the recursion equation can be written

$$g_s = \sum_{j=1}^{\min(s,2)} \left(\frac{j}{s}\right) f_j \, g_{s-j} \quad for \ s = 1,2,\ldots$$

$$\Rightarrow g_1 = f_1 \, g_0 = (0.3)(0.67032) = 0.20110$$

$$g_2 = f_1 \, g_1 + f_2 \, g_0 = \left(\frac{1}{2}(0.3)(0.20110) + (0.1)(0.67032)\right)$$

$$= 0.09720$$

So $\Pr[S < 1000] = g_0 + g_1 + g_2 = \mathbf{0.96861}$

(d) (*2 points*) The net premium is

$$P^{SL} = E[S] - E[S \wedge 1000]$$

$$E[S] = \lambda E\left[X^{**}\right] = 1(400 \times 0.3 + 800 \times 0.1) = 200$$

$$E[S \wedge 1000] = 0g_0 + 400g_1 + 800g_2 + 1000\left(1 - (g_0 + g_1 + g_2)\right)$$

$$= 400(0.20110) + 800(0.09720) + 1000(0.03139)$$

$$= 189.58$$

$$\Rightarrow P^{SL} = \mathbf{10.42}$$

Examiners' Comments

1. *Overall, the grades for this question were very much lower than expected.*
2. *Part (a) was done reasonably well, with around half of the candidates achieving full credit. The LER is a common feature in ASTAM exams, and candidates are expected to be very familiar with the calculations.*
3. *Part (b) was also fairly well, with more than half of the candidates correctly describing the ordinary and franchise deductible. Only a few candidates were able to provide a relevant example.*
4. *In part (c), candidates were asked to calculate the zero claim probability, and then to use recursion to calculate the probabilities associated with total claims being 400 or 800. Many candidates omitted this part. Many others were able to calculate the zero aggregate claim probability, but could not implement the recursion formula from the formula sheet, even though the application in this case was quite straightforward. Partial marks were awarded for candidates who recognized the correct formula, and made some attempt to use it. Many candidates omitted this part.*
5. *In part (d) candidates were expected to use the probabilities calculated in (c) to determine the stop loss premium. Candidates who did not calculate the correct probabilities in (c) were not penalized in (d) for using their incorrect values. Candidates who did not have any probabilities from (c) could gain partial credit by writing down the relevant formulas. Some candidates used arbitrarily chosen probabilities to calculate the net premium. If they explained their approach, and if their calculations were correct, they received full credit. Overall however, the points for this part were very low, with many candidates omitting it.*

Question 5 (10 points)

(a) (*4 points*) In the following, subscript $i=0$ corresponds to Q3-2023, $i=1$ corresponds to Q4-2023, and so on. The 2024 accident quarters then correspond to $i=2,3,4,5$.

(i) The development factors are:

$$f_0=\frac{1190}{420}=2.8333; \quad f_1=\frac{1156}{900}=1.2844; \quad f_2=\frac{966}{816}=1.1838;$$

$$f_3=\frac{670}{606}=1.1056; \quad f_4=\frac{316}{310}=1.0194$$

$$\lambda_j=\prod_{k=j}^{4} f_k$$

$$\Rightarrow \lambda_0=4.8554; \quad \lambda_1=1.7137; \quad \lambda_2=1.3342; \quad \lambda_3=1.1270; \quad \lambda_4=1.0194$$

So 2024 Projected cumulative claims are

$$C_{2,3}\lambda_3+C_{3,2}\lambda_2+C_{4,1}\lambda_1+C_{5,0}\lambda_0$$

$$=360(1.1270)+340(1.3342)+290(1.7137)+100(4.8554)$$

$$=\mathbf{1841.85}$$

(ii) Claim severities are fixed amounts, so there is no inflationary impact.

(iii) Let $\tilde{X}_{i,j}=\tilde{C}_{i,j}-\tilde{C}_{i,j-1}$ denote the projected claim payments made in Development Quarter (DQ) *j* arising in Accident Quarter (AQ) *i*.

Q2-2025 projected payments are $\tilde{X}_{i,j}$ where $i+j=7$.

So the total projected Q2-2025 payments from 2024 claims (that is, for $i=2,3,4,5$) are:

$$\tilde{X}_{2,5}+\tilde{X}_{3,4}+\tilde{X}_{4,3}+\tilde{X}_{5,2};$$

$$\tilde{X}_{2,5}=\tilde{C}_{2,5}-\tilde{C}_{2,4}=C_{2,3}\left(f_3 f_4-f_3\right)=7.70$$

$$\tilde{X}_{3,4}=\tilde{C}_{3,4}-\tilde{C}_{3,3}=C_{3,2}\left(f_2 f_3-f_2\right)=42.51$$

$$\tilde{X}_{4,3}=\tilde{C}_{4,3}-\tilde{C}_{4,2}=C_{4,1}\left(f_1 f_2-f_1\right)=68.47$$

$$\tilde{X}_{5,2}=\tilde{C}_{5,2}-\tilde{C}_{5,1}=C_{5,0}\left(f_0 f_1-f_0\right)=80.59$$

Total: **199.28**

(b) (*1 point*) Calendar year effect means that claim payment patterns are dependent on the calendar year (or quarter) of payment, rather than just AQ and DQ.

(c) (*4 points*)

(i) The null hypothesis is

H_0: Development factors are independent of calendar quarter of payment.

(ii) We consider the off diagonals of the triangle, which correspond to payment quarter-years. We ignore the first diagonal ($k = 0$) which only has a single observation.

For each diagonal we find the following statistics:

S_k is the number of S's in QY k; L_k is the number of L's in QY k;

$n_k = S_k + L_k$; $m_k = \lfloor (n_k - 1)/2 \rfloor$; $Z_k = \min\left(S_k, L_k\right)$

That is:

$k = 1$ (Q4-2023): $S_1 = 2$; $L_1 = 0$; $n_1 = 2$; $m_1 = 0$; $Z_1 = 0$

$k = 2$ (Q1-2024): $S_2 = 2$; $L_2 = 0$; $n_2 = 2$; $m_2 = 0$; $Z_2 = 0$

$k = 3$ (Q2-2024): $S_3 = 0$; $L_3 = 3$; $n_3 = 3$; $m_3 = 1$; $Z_3 = 0$

$k = 4$ (Q3-2024): $S_4 = 1$; $L_4 = 3$; $n_4 = 4$; $m_4 = 1$; $Z_4 = 1$

(iii) From the formula sheet:

$$E\left[Z_4\right] = \frac{n_4}{2} - \binom{n_4 - 1}{m_4}\frac{n_4}{2^{n_4}} = 2 - \binom{3}{1}\frac{1}{4} = 2 - \frac{3}{4} = \mathbf{1.25}$$

(iv) The test statistic is $Z = \sum Z_k = \mathbf{1}$

(v) $$p = 2\left(1 - \Phi\left(\frac{|Z - E[Z]|}{\sqrt{Var[Z]}}\right)\right) = 2\left(1 - \Phi\left(\frac{2}{\sqrt{1.125}}\right)\right) = \mathbf{5.93\%}$$

The test is not significant at the 5% level, and so at this significance we would accept the null hypothesis that payment patterns in different quarters are not significantly different.

(Note though that the result is significant at the 10% level, and with so little data, the test is not very powerful.)

(d) (*1 point*) The first chain ladder assumption is that expected run-off patterns are the same for different accident quarter years. If there is a calendar year effect, that implies that the run-off rates might change over time – for example, if the insurer implements improved claims management systems at the start of 2025, then the development factors in 2025 will be different from those before 2025.

Examiners' Comments

1. *The calendar year test used in the second part of this question has not been tested before in ASTAM, and the examiners were aware that many candidates might not have studied it. However, all the necessary formulas were available from the formula sheet. Candidates are advised to be very familiar with all the contents of the formula sheet in order to do well on this exam.*
2. *Part (a)(i) was done well. Part (a)(ii) required candidates to read and understand the context of the question, and was missed by many. Part (a)(iii) required candidates to*

understand how accident quarter, development quarter, and calendar quarter are represented in a claims triangle, as well as the difference between cumulative claims and paid claims. This part was more problematic. Candidates should expect the exam to test their understanding of context and processes.

3. *Part (b) was hit or miss. The assumption that payment patterns are independent of calendar year/quarter is key to the chain ladder and associated methods and models. Around 40% of candidates achieved all or most credit for this part. Around the same number scored no points for this part.*
4. *Part (c) was also hit or miss – candidates either achieved close to maximum points, or 0 points.*
5. *Part (d) was omitted by most candidates. Most of those who did attempt the question offered valid reasons and earned the full exam point.*

Question 6 (9 *points*)

(a) (*2 points*)

(*i*) $E[S]=E[N]E[Y_j]$

$E[N]=r\beta=15; \qquad E[Y_j]=\alpha\theta=1260$

$\Rightarrow E[S]=15(1260)=\mathbf{18{,}900}$

(*ii*) $V[S]=E[N]V[Y_j]+V[N]E[Y_j]^2$

$V[Y_j]=\alpha\theta^2=45{,}360{,}000; \quad V[N]=r\beta(1+\beta)=60$

$\Rightarrow V[S]=775{,}656{,}000 \Rightarrow SD[S]=\mathbf{27{,}850.60}$

(b) (*2 points*)

(*i*) $S\approx Z$, say, where $Z\sim N\left(\mu,\sigma^2\right)$ with $\mu=18{,}900, \quad \sigma=27{,}850.60$

$VaR_{95\%}\left[Z\right]=Q$, say, where $\Pr\left[Z\le Q\right]=0.95$

$$\Rightarrow \Phi\left(\frac{Q-\mu}{\sigma}\right)=0.95 \;\Rightarrow\; \frac{Q-\mu}{\sigma}=1.64485 \qquad \left(\text{Excel: NORM.S.INV(0.95)}\right)$$

$\Rightarrow Q=\mu+1.64485\sigma=\mathbf{64{,}710.16}$

(ii) From the formula sheet

$$ES_{95\%}\left(Z\right)=\mu+\frac{\sigma}{\left(1-0.95\right)}\phi(z_{0.95})$$

where ϕ is the standard normal pdf and $z_\alpha=\Phi^{-1}(\alpha)$

$\Rightarrow \qquad z_{0.95}=1.64485$

$\Rightarrow \phi(z_{0.95})=0.103136$ (Excel: NORM.S.DIST(1.64485,FALSE))

$$\Rightarrow ES_{95\%}\left(Z\right)=18{,}900+\frac{27850.6}{0.05}(0.103136)=\mathbf{76{,}347.8}$$

(c) (*3 points*)

(i) Now we assume $S\approx W$ where $W\sim \text{LogN}\left(\mu^*,\sigma^*\right)$. We find μ^* and σ^* by matching two moments of S to two moments of L. From the formula sheet we have:

$E\left[W\right]=e^{\mu^*+(\sigma^*)^2/2}=18{,}900 \Rightarrow \mu^*+\left(\sigma^*\right)^2/2=\ln(18900)=9.84692$

$E\left[W^2\right]=e^{2\mu^*+2(\sigma^*)^2}=V[S]+E[S]^2=1{,}132{,}866{,}000 \Rightarrow 2\mu^*+2\left(\sigma^*\right)^2=20.84801$

$\Rightarrow \left(\sigma^*\right)^2=20.84801-2(9.84692)=1.15418$

$\Rightarrow \sigma^*=\mathbf{1.074329}$

(ii) We need Q such that $\Pr\left[W\le Q\right]=0.95$, which means we first need to find μ^*.

We know that $\mu^* + (\sigma^*)^2 = 9.84692 \Rightarrow \mu^* = 9.84692 - 1.15418/2 = 9.2699$. So we have

$$\Pr[W \le Q] = 0.95 \Rightarrow \Phi\left(\frac{\log Q - \mu^*}{\sigma^*}\right) = 0.95$$

$$\Rightarrow \frac{\log Q - \mu^*}{\sigma^*} = 1.64485$$

$$\Rightarrow Q = \exp(\mu^* + 1.64485\sigma^*) = \mathbf{62{,}127.2}$$

(iii) From the formula sheet:

$$ES_{95\%}(W) = \frac{E[W]}{0.05}\Phi\left(z_{0.05} + \sigma^*\right) \quad = \frac{18900}{0.05}\Phi\left(-1.64485 + 1.07433\right)$$

$$= \mathbf{107{,}412.8}$$

(d) (*2 points*)

(i) The Central Limit Theorem states that the aggregate distribution will converge to a normal distribution as the expected claim frequency increases. In this case we only have 15 expected claims, which is low for applying the Normal Approximation.

We see that the aggregate loss distribution is positively skewed (the mean is less than the standard deviation, even though the distribution is bounded below by 0), but the normal distribution has 0 skewness. This means that the normal distribution will underestimate the right tail probabilities of the aggregate claims.

The lognormal distribution is positively skewed, and is therefore likely to give a better approximation in the tail.

The lognormal estimate of the ES is also significantly larger than the normal approximation, so selecting the lognormal would be a safe side assumption.

(ii) If we discretize the severity distribution, i.e. approximate the gamma distribution with a discrete distribution, then we can use recursions (or Fast Fourier transforms) to determine an approximate full probability distribution for *S*. This method allows for fast calculation and very good accuracy when the discretization is sufficiently fine.

Examiners' Comments

1. *The examiners expected candidates to find this question to be one of the easiest on the exam. Matching moments of distributions is an elementary exercise in introductory statistics, and all of the formulas required for the moments, and for the Expected Shortfall (ES) calculations were provided in the formula sheet. However, most candidates struggled with all parts of this question other than (a) and (b)(i). Candidates*

were not familiar with the formulas on the formula sheet, and only a minority were able to fit a lognormal distribution using the first two moments of S.

2. *Part (a) and part (b)(i) were done well by almost all candidates who attempted them.*
3. *Part (b)(ii) required candidates to apply the ES formula from the Normal distribution section of the formula sheet. Most candidates identified the correct formula, but did not know the meaning of the* ϕ *term, even though it is defined on the formula sheet as the pdf of the N(0,1) distribution. Familiarity with the formula sheet is an important part of the preparation for this exam.*
4. *Part (c)(i) required candidates to find the two lognormal parameters given the distribution mean and variance. The formulas for the lognormal moments are given on the formula sheet. Nevertheless, only a minority of candidates were able to solve for* σ^**. Candidates who were not able to do so could have used given value of* σ^* *to find* μ^* *for parts (c)(ii) and (c)(iii), but very few did so.*
5. *Although the solution above uses E[W] and E[W*^{2}*] to find* σ^**, it is also quite straightforward to find it using E[W] and Var[W].*
6. *A common mistake in (c)(i) was to assume that the mean of the lognormal distribution is* e^{μ^*}*. In some contexts, the lognormal distribution may be parameterized in this way, but in this exam we always use the more common parameterization, such that the mean is* $e^{\mu^*+(\sigma^*)^2/2}$*, which is the version given in the formula sheet, and which is the parametrization of the distribution of* e^Z*, where* $Z \sim N\left(\mu, \sigma^2\right)$.
7. *Many candidates who attempted part (c)(iii) made the error of using* $z_{0.95} = 1.64485$ *in the formula instead of* $z_{0.05} = -1.64485$.
8. *Part (d)(i) was not done well. It is important to have some understanding of when the normal distribution is likely to be an adequate approximation, and when it is not. Luckily, the central limit theorem helps with that – when we aggregate a sufficiently large number of independent random variables, then the normal distribution will become increasingly accurate. Here though, we have a small number of expected losses, together with a positively skewed distribution, and an instruction that the insurer is particularly interested in the right tail. The lognormal approximation is positively skewed, and the ES value is significantly larger than the ES estimated using the normal distribution. Both facts point to the lognormal likely being a more appropriate approximating distribution in this case.*

   Note that full credit for this part did not require the level of explanation given in the model above. It was sufficient for a candidate to note that the underlying distribution is positively skewed, and hence that the right tail risk measure estimated with the lognormal distribution, which is also positively skewed, is likely to be more accurate than the normal distribution, which is not.

9. *In practice, neither of these approximations will be very accurate. The point of part (d)(ii) was for candidates to show that they understand that the process of discretizing the severity distribution, and then using recursions (or fast Fourier transforms) to construct an a discretized aggregate distribution is another approach to approximating the underlying aggregate distribution. Very few candidates made this connection.*