ASTAM Spring 2026 Model Solution

Question 1 – See Excel

Question 2

(a) (i) Expected Loss: $\dfrac{5{,}125{,}000}{2{,}500} = 2{,}050.0$

(ii) 500 deductible: $\dfrac{5{,}125{,}000 - 172{,}000 - (500)(2500 - 770)}{2{,}500} = \dfrac{4{,}088{,}000}{2{,}500} = 1{,}635.2$

(iii) 1000 Deductible:

$$\frac{5{,}125{,}000 - 172{,}000 - 328{,}000 - (1000)(2500 - 770 - 450)}{2{,}500} = \frac{3{,}345{,}000}{2{,}500} = 1{,}338.0$$

(b) Since 500 is the base deductible:

(i) $IDR_0 = \dfrac{2050}{1635} = 1.254$

(ii) $IDR_{1000} = \dfrac{1338}{1635} = 0.818$

Examiners' Comments:

1. *Candidates generally did well on parts (a) and (b).*
2. *It is important for candidates to understand the difference between cost per loss and cost per payment.*

(c) (i) $e(6000) = E\left[X \mid X > 6000\right] - 6000$

$$= \frac{1{,}914{,}000}{170} - 6000 \quad = 5{,}258.82$$

(ii) The expected value of all the claims is 2050. The expected excess amount of claims over 6000 is 5259. When the tail mean excess loss is greater than the overall mean, that signals a heavy tailed distribution.

Examiners' Comments:

1. *Performance on this part was mixed. Many candidates achieved full credit. A significant number forgot to deduct the threshold value in the mean excess loss calculation.*
2. *The Examiners were pleased that most of the candidates who were able to calculate the MEL were also able to interpret the results.*

(d) (i) Contract A:

$$\frac{\left(687,000+1,914,000-\left(140+170\right)\left(4,000\right)\right)}{2,500}=\frac{1,361,000}{2,500}=544.4$$

Contract B:

$$\frac{0.5\left(\left(1,276,000+687,000\right)-\left(450+140\right)\left(2,000\right)\right)+(170)(2,000)}{2,500}=\frac{731,500}{2,500}=292.6$$

(ii) NED should select the excess loss insurance (option A) as that largely eliminates the tail risk since the maximum that NED would have to payout on each claim is 4000. The max for option B is unlimited, as the insurer is responsible for all but 2000 of any loss over 6000.

Examiners' Comments:

1. *The main point of (i) is to test whether candidates can correctly interpret the reinsurance cover, and extract the relevant information from the data table. Contract A was done fairly well, Contract B proved more of a challenge.*
2. *In part (ii), a coherent explanation was necessary for any credit.*
3. *The Examiners are looking for candidates to demonstrate not only that they can do calculations, but also that they can understand and interpret the information given, and numerical results, in an insurance context.*

(e) Advantages (only one required):

- This eliminates the tail risk which we can see from above is significant.
- This would eliminate the need for reinsurance.

Disadvantages (only one required):

- Having no deductible means that NED will need to deal with all claims, even those that are very small. This will increase the expenses per unit of claims settled.
- There will be a significant portion of the insureds (about 170/2500) who will have claims exceeding the upper limit. It is probable that many of these insureds will not have understood the implications of the upper limit and will be very dissatisfied with the policy and with NED as the insurance company. There is a reputational risk involved here.

Examiners' Comments:

This part was generally well done.

Question 3

(a) (i) $E[S]=E[N]E[X]=(2)(1000)=2000$

(ii) $V[S]=E[N]V(X)+V[N]\left(E\left[X\right]\right)^2$

$$=\lambda\theta^2+\lambda\theta^2=2(2)\left(\theta^2\right)$$

$$\Rightarrow SD\left[S\right]=2\theta=2000$$

$\Big($Or: $V\left[S\right]=\lambda E\left[X^2\right]=2\left(2\theta^2\right)=4\theta^2\Big)$

(b) Let S_P denote the aggregate claims over the whole portfolio. Then

$$E[S_P]=(10{,}000)(2000)=20{,}000{,}000$$

$$V[S_P]=(10{,}000)(4{,}000{,}000)=(200{,}000)^2$$

$$S_P\approx \mathrm{N}\left(\mu=20{,}000{,}000,\ \sigma^2=(200{,}000)^2\right)$$

$$\Rightarrow Q_{0.95}\left(S_P\right)\approx\mu+1.64485\sigma=20{,}328{,}971$$

Examiners' Comments:

Almost all candidates earned full credit for (a), and most also earned full credit for (b).

(c) (i) Let G denote the gross premium for an individual policy. From the information given, we know that over the whole portfolio, premiums minus claims at the 95th percentile minus expenses = 5% of gross premiums, that is, for a portfolio of $n=10{,}000$ policies:

$$nG-(0.08)nG-0.12nG-Q_{0.95}\left(S_P\right)-300{,}000=0.05nG$$

$$\Rightarrow G=\frac{Q_{0.95}\left(S_P\right)+300{,}000}{10{,}000\left(0.8-0.05\right)}=\frac{20{,}628{,}971}{7500}$$

$$=2750.53$$

(ii) The expected profit over the whole portfolio is

$$(10{,}000)(0.8G)-E\left[S_P\right]-300{,}000=1{,}704{,}200$$

$\Rightarrow$ Expected profit per policy $=170.42$

Examiners' Comments:

1. *Almost all candidates calculated the premium correct but virtually no candidate correctly calculated the expected profit.*
2. *The expected profit includes both the 5% of explicit profit but also the profit expected from using the 95th percentile of claims as opposed to the expected value of claims. The term "expected" signals that the claims are included at their expected value, not at the quantile value used as a margin in the premium calculation.*

(d)

Method of rounding/mass dispersal:

Advantage:

- The method is intuitive and easy to implement

Disadvantage:

- The moments of the discretized distribution will not match the moments of the original distribution.

Method of local moment matching:

Advantage:

- Greater accuracy for a given value of *h* when used to approximate a compound distribution, compared with the method of rounding.

Disadvantage (one required)

- Far more complex calculation of probabilities.
- Formulas are not intuitive.
- For $p \geq 2$ the discretized "probabilities" may be negative.

Examiners' Comments:

1. *The question asks for the advantages and disadvantages and not a description of the methods. No credit was given for the description of the methods.*
2. *The Examiners' are looking for evidence of understanding, rather than just doing a key word search. Candidates are encouraged to explain advantages and disadvantages in their own words (but briefly). The verbal questions are an opportunity for candidates to demonstrate their knowledge to the Examiners.*

(e) (i) $f_1 = F(400+200) - F(400-200) = 0.451188 - 0.181269 = 0.26992$

(ii) $$f_1 = \frac{2E[X \wedge 400] - E[X \wedge 0] - E[X \wedge 800]}{400}$$

$$= \frac{2\left(2000\left[1-e^{-\frac{400}{1000}}\right]\right) - 2000\left[1-e^{-\frac{0}{1000}}\right] - 2000\left[1-e^{-\frac{800}{1000}}\right]}{400}$$

$$= \frac{2(329.68) - 0 - 550.67}{400}$$

$$= 0.27172$$

Examiners' Comments:

1. *The formulas for the method of local moment matching and for* $E[X \wedge x]$ *for the exponential distribution are given in the formula sheet.*

2. *Candidates who attempted this part generally did well. The most common error was assuming that* $E[X \wedge 0]$ *is* $E[X]$. *In fact,* $E[X \wedge 0] = E[\min(X,0)] = 0$.

Question 4

(a) $N \mid \Lambda \sim \text{Poi}(\Lambda) \Rightarrow E[N \mid \Lambda] = \Lambda$ and $V[N \mid \Lambda] = \Lambda$

$\Lambda \sim \text{Gamma}(\alpha = 2, \theta = 0.2) \Rightarrow E[\Lambda] = 0.4$ and $V[\Lambda] = 0.08$

(i) $E[N] = E[E[N \mid \Lambda]] = E[\Lambda] \quad = 0.4$

(ii) $$V[N] = E[V[N \mid \Lambda]] + V[E[N \mid \Lambda]]$$
$$= E[\Lambda] + V[\Lambda]$$
$$= 0.4 + 0.08 = 0.48$$
$$\Rightarrow \text{SD}[N] = 0.6928$$

Examiners' Comments:

1. *Part (a) was generally done very well by almost all candidates, especially part (a)(i), for which most candidates received full credit.*
2. *Part (a)(ii) was mostly done well, though some candidates omitted one of the two variance terms. A small number of candidates also overlooked the need to convert the variance to a standard deviation.*
3. *Some candidates knew that the unconditional distribution of N is Negative Binomial and used this to get the mean and variance. This earned full credit if done correctly.*

(b) Let $\pi(\lambda)$ denote the prior pdf for Λ, and let $L(\boldsymbol{x} \mid \lambda)$ denote the likelihood function for the data assuming that $\Lambda = \lambda$. The posterior pdf for Λ is proportional to the product of the prior and the likelihood, i.e.

$$\pi_{\Lambda|\underline{x}}(\lambda) \propto \pi(\lambda) L(\boldsymbol{x} \mid \lambda)$$
$$\pi(\lambda) = \frac{\lambda^{\alpha-1} e^{-\lambda/\theta}}{\theta^{\alpha}\Gamma(\alpha)} = \frac{\lambda e^{-5\lambda}}{0.2^2}$$
$$L(\boldsymbol{x} \mid \lambda) = \left(\frac{e^{-\lambda}\lambda^0}{0!}\right)\left(\frac{\lambda^2 e^{-\lambda}}{2!}\right)\left(\frac{e^{-\lambda}\lambda^0}{0!}\right) = \frac{\lambda^2 e^{-3\lambda}}{2}$$
$$\Rightarrow \pi_{\Lambda|\underline{x}}(\lambda) \propto \lambda^3 e^{-8\lambda} \text{ (ignoring terms not involving } \lambda)$$

By comparison with the Gamma pdf, we see that the posterior distribution is Gamma with

$$\lambda^{\alpha^*-1} = \lambda^3 \Rightarrow \alpha^* = 4 \text{ and } e^{-\lambda/\theta^*} = e^{-8\lambda} \Rightarrow \theta^* = \frac{1}{8} = 0.125.$$

Examiners' Comments:

1. *The goal of this part was for candidates to demonstrate that they could derive the conjugacy result, as shown by the use of the words "Show that" in the question. Only a minority of candidates did so. Many others simply stated the conjugacy result which earned very little credit.*

(c) (i) The predictive distribution is the unconditional distribution of *N* using the posterior distribution for Λ. That is

$$p_k = \int_0^\infty \Pr\left[N = k \mid \Lambda = \lambda\right] \pi(\lambda)\, d\lambda$$

$$= \int_0^\infty \frac{e^{-\lambda}\lambda^k}{k!} \frac{\lambda^3 e^{-8\lambda}}{\left(\tfrac{1}{8}\right)^4 \Gamma(4)}\, d\lambda = \int_0^\infty \frac{\lambda^{k+3} e^{-9\lambda}}{k!\left(\tfrac{1}{8}\right)^4 \Gamma(4)}\, d\lambda$$

We note that the integrand is proportional to a gamma pdf, evaluated at λ, with parameters $\alpha^+ = k+4$ and $\theta^+ = \frac{1}{9}$. Using the fact that the gamma pdf must integrate to 1 over the interval from 0 to ∞, we have

$$p_k = \int_0^\infty \frac{\lambda^{k+3} e^{-9\lambda}}{\left(\tfrac{1}{8}\right)^4 k!\Gamma(4)}\, d\lambda$$

$$= \int_0^\infty \frac{\lambda^{k+3} e^{-9\lambda}}{\left(\tfrac{1}{9}\right)^{k+4} \Gamma(k+4)}\, d\lambda \, \frac{\left(\tfrac{1}{9}\right)^{k+4} \Gamma(k+4)}{k!\left(\tfrac{1}{8}\right)^4 \Gamma(4)}$$

$$= \frac{\left(\tfrac{1}{9}\right)^{k+4} \Gamma(k+4)}{k!\left(\tfrac{1}{8}\right)^4 \Gamma(4)} \quad = \left(\frac{1}{9}\right)^k \left(\frac{8}{9}\right)^4 \frac{\Gamma(k+4)}{\Gamma(4)k!}$$

This is a negative binomial probability function, with parameter $r = 4$, and where $\frac{\beta}{1+\beta} = \frac{1}{9} \Rightarrow \beta = \frac{1}{8} = 0.125$

(ii) Under the predictive distribution $E[N] = r\beta = \frac{4}{8} = 0.5$

Examiners' Comments:

1. *The average points earned for this part were very low. Very few candidates made a substantive effort to do part (i), though those that did tended to get a majority of the credit.*
2. *Many candidates noted that the Gamma distribution is a conjugate prior for the Poisson rate parameter; this is true and relates to the result in part (b), but isn't relevant for part (c) and earned no credit.*
3. *Part (c)(ii) was generally done well by those who attempted it, with the only common mistake being to mix up parameter values.*

(d) (i) $v = E\left[V\left[N\right]\right] = E\left[\Lambda\right] = 0.4$

$a = V\left[E\left[N\right]\right] = V\left[\Lambda\right] = 0.08$

$\mu = E\left[N\right] = 0.4$

(ii) $Z = \dfrac{3}{3 + \dfrac{v}{a}} = \dfrac{3}{8} = 0.375$

$$Cred\ \ Est = Z\bar{X} + \left(1 - Z\right)\mu = 0.375\left(\frac{2}{3}\right) + 0.625(0.4) = 0.50$$

Examiners' Comments:

1. *Part (d) was generally done quite well, with most candidates getting full (or nearly full) credit.*
2. *Some candidates mixed up the Bühlmann parameters, typically swapping a and v, and some others used posterior parameter values instead of prior values.*
3. *The Bühlmann estimate of the claim frequency, part (d)(ii), was done well by the bulk of the candidates.*

(e) We only have 3 years of data. This means that the estimate from the policyholder's own experience is not likely to be very accurate.
By using a credibility approach, we can combine the information from the policyholder's experience with the prior expectation based (presumably) on substantially more data.

Examiners' Comments:

This part was generally done well by those who attempted it.

Question 5

(a) $f_0 = \dfrac{317+390}{235+320} = 1.2739;\ f_1 = \dfrac{349}{317} = 1.100946$

(i) $\widehat{C}_{1,2} = 390\, f_1 = 429.37$

(ii) $\widehat{C}_{2,2} = 416\, f_0\, f_1 = 583.43$

(b) The γ_j values represent the proportion of claims settled in DY *j*.

Examiners' Comments:

1. *Part (a) was done very well, with almost all candidates earning full credit.*
2. *Part (b) was also done well, though a number of candidates simply described the formula for γ_j in words, rather than giving the interpretation. The Examiners are looking for evidence that candidates can see beyond the formulas.*

(c) From the formula sheet:

$$\hat{v} = \frac{1}{I}\sum_{i=0}^{I-1} \hat{s}_i^2$$

$$\hat{s}_i^2 = \frac{1}{I-i}\sum_{j=0}^{I-i} \hat{\gamma}_j \left(\frac{X_{i,j}}{\hat{\gamma}_j} - \widehat{C}_{i,J} \right)^2$$

$$I = 2 \quad X_{1,0} = C_{1,0} = 320 \quad X_{1,1} = C_{1,1} - C_{1,0} = 70 \quad \widehat{C}_{1,2} = 429.37$$

$$\Rightarrow \hat{s}_1^2 = \hat{\gamma}_0 \left(\frac{X_{1,0}}{\hat{\gamma}_0} - \widehat{C}_{1,2} \right)^2 + \hat{\gamma}_1 \left(\frac{X_{1,1}}{\hat{\gamma}_1} - \widehat{C}_{1,2} \right)^2$$

$$= 0.71303(19.420)^2 + 0.19528(-70.909)^2 = 1250.80$$

$$\Rightarrow \hat{v} = (625.41 + 1250.80)/2 = 938.11$$

Examiners' Comments:

1. *This question required candidates to understand how to use the formula given on the formula sheet.*
2. *More than half of the candidates earned full credit on this part.*
3. *Candidates who lost credit generally did so through misinterpreting the formula inputs. For example, a common error was to use the cumulative claims (i.e. $C_{1,1}$) in the second term of the formula, instead of the incremental claims (i.e. $X_{1,1}$). This demonstrates the value for all candidates in being very familiar with the formula sheet,*
4. *Note that $\dfrac{X_{1,j}}{\hat{\gamma}_j}$ is an estimate of $C_{1,2}$, and so each of the two terms in the formula is a squared error term.*

5. *A few candidates also calculated* $\hat{s}_0^2$, *even though it was given in the question. There is a lot of value in reading the question through several times before embarking on any calculations.*
6. *Many candidates used Excel to do the calculations for this question. Candidates who wrote down the formulas used and the inputs earned full credit if the calculation was done correctly. A few candidates wrote down the answer with no explanation. Because the question used the wording "Show that" these answers received little credit.*
7. *Some candidates used wrong inputs or formulas and magically ended up with the rounded answer given. However, the examiners do read every line of the candidates' work. A candidate who honestly records an answer that is incorrect may receive more partial credit for the same error than a candidate who tries to bluff the examiners.*

(d) (i) From the formula sheet:

$$Z_i = \frac{\hat{\beta}_{I-i}}{\hat{\beta}_{I-i} + \hat{v}/\hat{a}}$$

where $\hat{\beta}_0 = \gamma_0 = 0.71303$; $\hat{\beta}_1 = \gamma_0 + \gamma_1 = 0.90831$; $\hat{\beta}_2 = \gamma_0 + \gamma_1 + \gamma_2 = 1$

$$\frac{\hat{v}}{\hat{a}} = 0.07670$$

$$\Rightarrow Z_0 = \frac{1}{1.07670} = 0.92877;$$

$$Z_1 = \frac{0.90831}{0.98501} = 0.922134;$$

$$Z_2 = \frac{0.71303}{0.78973} = 0.902879$$

(ii) Z_i measures the credibility of the chain ladder estimate of ultimate claims for AY *i*. For AY0, we have 3 years of data, for AY1 we have 2 years, and for AY2 we have 1 year, so the credibility of the estimates is highest for AY0 and lowest for AY2.

(iii) From the formula sheet, $\hat{\mu}$ is the average of the estimated ultimate claims, weighted by the credibility factors; that is

$$\hat{\mu} = \frac{Z_0 \widehat{C}_{0,2} + Z_1 \widehat{C}_{1,2} + Z_2 \widehat{C}_{2,2}}{Z_0 + Z_1 + Z_2} = \frac{1246.84}{2.7538} = 452.77$$

(iv)

$$\tilde{C}_{2,2}^{\text{BS}} = Z_2\, \hat{C}_{2,2} + (1 - Z_2)\,\hat{\mu}$$

$$= (0.90288)(583.43) + (0.09712)(452.77) = 570.74$$

$$\tilde{C}_{2,2}^{\text{BS2}} = Z_2^* \widehat{C}_{2,2} + \left(1 - Z_2^*\right)\mu$$

$$Z_2^* = 1 - \left(1 - \beta_0\right)\left(1 - Z_2\right) = 0.97213$$

$$\Rightarrow \tilde{C}_{2,2}^{\text{BS2}} = 0.97213(583.43) + 0.02787(452.77) = 579.78$$

Examiners' Comments:

1. *This part was done well by most candidates.*
2. *Most of the required formulas were on the formula sheet.*
3. *In part (iv), full credit was awarded for either the first or second iterated Bühlmann Straub estimate.*

(e) Advantages (2 required):

- Buhlmann-Straub (B-S) provides an objective, empirical estimate for the 'prior' mean μ. B-F typically uses the loss ratio estimate, which is subjective.
- The B-S estimate of μ is automatically updated as more data is available.
- The B-S credibility factors reflect both the within-AY volatility (v) and the between AY volatility (a). The B-F credibility factors do not reflect either, based solely on the average development pattern.

Disadvantages (1 required):

- The B-S approach assumes that the μ parameter is the same for all AY. The B-F method allows different μ values for different AY's.
- The B-F method requires fewer calculations.
- The B-F method is more widely used and recognized.

Examiners' Comments:

1. *This part was done less well than the rest of the question.*
2. *Many candidates listed advantages and disadvantages of B-F compared with the Chain Ladder, which did not answer the question.*
3. *Many candidates claimed that B-S used credibility, and B-F did not, which is incorrect. B-F is implicitly a credibility based method, where the credibility factor for an AY with data up to DY j is β_j. Similarly, many candidates claimed that the B-S method gave more weight where there is more data, but the B-F also does this, through the β_j credibility factors.*
4. *Many candidates claimed that B-S required more information or data than B-F which is not correct. Both use the run-off triangle data, but B-F requires an external candidate for the prior mean. The BS method requires no other information.*

Question 6

(a) Let $X_{(j)}$ denote the ranked losses.

(i) The 99% VaR is $X_{(0.99\times 500)} = X_{(495)} = 7530$

(ii) The 99% ES is the average of the 5 largest values, i.e.

$$ES = \frac{7548+...+21355}{5} = 12{,}119.8$$

Examiners' Comments:

1. *Question 6 was omitted by a significant number of candidates.*
2. *Candidates who attempted this question generally did well on this part.*

(b) (i) From the formula sheet:

$$Q_{0.99} = d + \frac{\beta}{\xi}\left(\left(\frac{S(d)}{0.01}\right)^{\xi} - 1\right)$$

where $d = 4953$, and $S(d)$ is the empirical survival function at d, i.e.

$$S(d) = \frac{16}{500} = 0.032$$

$$\Rightarrow Q_{0.99} = 4953 + \frac{1100}{0.74}\left(\left(\frac{0.032}{0.01}\right)^{0.74} - 1\right) = 6981.89$$

(ii) From the formula sheet

$$ES_{0.99} = \frac{1}{1-\xi}\left(Q_{0.99} + \beta - \xi d\right)$$

$$= \frac{1}{0.26}\left(6981.89 + 1100 - 0.74(4953)\right) = 16{,}987.20$$

Examiners' Comments:

1. *This part was done well by the candidates who did not omit the question..*

(c) We have $d = X_{(464)}$. This means that in the Hill estimator formula from the formula sheet, we have $j = 484$, so that

$$\hat{\alpha}^H = \left(\sum_{k=j}^{n} \frac{\log\left(x_{(k)}\right)}{n-j+1} - \log\left(x_{(j)}\right)\right)^{-1} = \left(\sum_{k=484}^{500} \frac{\log\left(x_{(k)}\right)}{17} - \log\left(4953\right)\right)^{-1}$$

$$= \left(\frac{175.391-(8.405+8.441+8.448)}{17} - 8.508\right)^{-1} = 3.1130$$

Examiners' Comments:

1. *Candidates performance was fair in part (c), on average. Many candidates correctly identified the given Hill estimator formula but missed the parameter inputs.*
2. *In particular, understanding that in this case, j = 484, was not recognized by a significant umber of the candidates..*

(d) (i) The q-VaR for the distribution is Q_q, say where $\hat{S}^H(Q_q) = 1-q$, so

$$\hat{S}^H(Q_q) = 1-q \Rightarrow 0.032\left(\frac{Q_{0.99}}{d}\right)^{-\hat{\alpha}^H} = 0.01$$

$$\Rightarrow Q_{0.99} = \left(\frac{0.032}{0.01}\right)^{1/\hat{\alpha}^H} \times d \quad = 7196.81$$

(ii) There are several ways to calculate the ES.

For simplicity, let $\hat{\alpha} = \hat{\alpha}^H = 3.113$, so that $S(x) = kx^{-\hat{\alpha}}$ where $k = 0.032d^{\hat{\alpha}}$.

Also let $Q = Q_{0.99} = 7196.81$ from (i) above, so that $S(Q) = 0.01$.

Method 1: As the random variable is continuous, we have

$$ES = E[X \mid X > Q] = \frac{1}{S(Q)}\int_Q^\infty x\, f(x)dx$$

$$= \frac{1}{S(Q)}\left(\left[-xS(x)\right]_Q^\infty + \int_Q^\infty S(x)dx\right) \qquad \text{(by parts)}$$

$$= Q + \frac{1}{S(Q)}\int_Q^\infty S(x)dx = Q + \frac{k}{0.01}\int_Q^\infty x^{-\hat{\alpha}}dx$$

$$= Q + \frac{k}{0.01}\frac{1}{(\hat{\alpha}-1)}Q^{-(\hat{\alpha}-1)} = Q + \frac{0.032\,d^{3.113}}{0.01(2.113)Q^{2.113}}$$

$$= 10{,}602.81$$

Method 2: As above, but using the density function

$$ES = \frac{1}{S(Q)}\int_Q^\infty x\ f(x)\,dx \quad \text{where} \quad f(x) = -\frac{d}{dx}S^H(x) = \hat{\alpha}k\,x^{-\hat{\alpha}-1}$$

$$\Rightarrow ES = \frac{1}{S(Q)}\int_Q^\infty \hat{\alpha}k\,x^{-\hat{\alpha}}dx \quad = \frac{\hat{\alpha}\,k}{0.01}\left[\frac{1}{\hat{\alpha}-1}Q^{-(\hat{\alpha}-1)}\right]$$

$$= \frac{3.113\,(0.032)d^{3.113}}{(0.01)(2.113)Q^{2.113}} \quad = 10{,}602.81$$

Method 3:

Another definition for ES is $ES_q = \frac{1}{1-q}\int_q^1 Q_p \, dp$

From (i) we see that $Q_p = d\left(0.032\right)^{1/\hat{\alpha}} (1-p)^{-(1/\hat{\alpha})}$

$$\Rightarrow ES_{0.99} = \frac{d(0.032)^{1/\hat{\alpha}}}{0.01}\int_{0.99}^{1}\left(1-p\right)^{-1/\hat{\alpha}}dp \quad = \frac{d(0.032)^{1/\hat{\alpha}}}{0.01\left(1-\frac{1}{\hat{\alpha}}\right)}\left[-\left(1-p\right)^{1-1/\hat{\alpha}}\right]_{0.99}^{1}$$

$$= \frac{4953\left(0.33098\right)(0.99)^{0.678765}}{0.01\left(0.678765\right)} = 10{,}602.81$$

Examiners' Comments:

1. *Although a substantial number of candidates achieved full credit on part (i), only a handful were able to apply first principles to determine the ES in part (ii).*

(e) I would recommend the MLE approach

- The MLE approach utilizes the full specification of GPD assumptions, while the Hill estimator is a semi-parametric approach that estimates the tail index only.
- The MLE approach is likely to be more stable for small data size. In this case, only about 17 extreme values are used for the MLE and the Hill estimators.
- The generated ES from Hill estimator is less conservative (smaller than the empirical ES).
- The empirical ES is constrained by the range of the data. The MLE and Hill can extrapolate the distribution beyond the data.

Examiners Comments::

1. *Although the MLE method is generally preferred, for the reasons stated, appropriate credit was also given if candidates chose one of the alternatives with sensible justification.*
2. *No credit was awarded for picking the Hill estimator "because it's the smallest". That's bad risk management.*