Advanced Short-Term Actuarial Mathematics

Solutions to Sample Questions

This Study Note contains the solutions to the sample questions of the Advanced Short-Term Actuarial Mathematics exam. There may be alternative solution methods that are not presented here.

Different solutions show different levels of accuracy in intermediate results. These model solutions are not intended to imply that this is the best rounding for each question. Graders do not penalize rounding decisions, unless an answer is rounded to too few digits in the context of the problem and the given information. In particular, if a problem in one step asks you to calculate something to the nearest 1, and you calculate it as (for example) 823.18, you need not bother saying “that’s 823 to the nearest 1”, and you may use 823.18 or 823 in future steps

In the numerical solutions presented, there may be small rounding differences arising from the fact that values used in the calculations are typically more accurate than the intermediate values recorded.

Versions:

Nov 9, 2022	Original set of 20 sample questions published for the ASTAM exam
Feb. 6, 2024	Added two new sample questions (Questions 21-22); Question 22 is a sample Excel question.

Current Version Dated Feb. 6, 2024

Question 1

(a) $\hat{\lambda} = \overline{X} = \dfrac{3400 + (2)(2400) + (3)(1500) + (4)(300)}{10,000} = 1.39$

$Var[\hat{\lambda}] = \dfrac{\hat{\lambda}}{n} = \dfrac{1.39}{10,000} = 0.000139$

90% CI : $1.39 \pm 1.645\sqrt{0.000139} = (1.3706, 1.4094)$

(b) Using $\hat{\lambda} = 1.39$ from (a):

Number of Claims	**Observed Number of Policies** (O_j)	**Expected Number of Policies** E_j	$\dfrac{(E_j - O_j)^2}{E_j}$
0	2400	$10,000e^{-1.39} = 2490.753$	3.307
1	3400	$10,000e^{-1.39}(1.39) = 3462.147$	1.116
2	2400	$10,000e^{-1.39}(1.39)^2 / 2 = 2406.192$	0.016
3+	1800	$10,000 - 2490.753 - 3462.147 - 2406.192 = 1640.908$	15.425
		χ^2 ➔	**19.864**

Degrees of freedom: 4–1–1=2

The *p*-value is $\Pr\left[\chi_2^2 > 19.864\right] = 5 \times 10^{-5}$. Since this is very small, it is extremely unlikely that this data came from the Poisson distribution with $\hat{\lambda} = 1.39$, and we reject the null hypothesis.

(c) The K-S test statistic calculations are:

x	$F_n(x^-)$	$F_n(x)$	$F^*(x) = 1 - \left(\dfrac{75,000}{75,000 + x}\right)^4$	Absolute Value of Maximum Difference
200	0	0.2	$1 - \left(\dfrac{75,000}{75,200}\right)^4 = 0.01060$	0.18940
1000	0.2	0.4	$1 - \left(\dfrac{75,000}{76,000}\right)^4 = 0.05160$	0.34840

5000	0.4	0.6	$1-\left(\frac{75,000}{80,000}\right)^4 = 0.22752$	0.37248
10,000	0.6	0.8	$1-\left(\frac{75,000}{85,000}\right)^4 = 0.39387$	0.40613
100,000	0.8	1.0	$1-\left(\frac{75,000}{175,000}\right)^4 = 0.96626$	0.16626

The test statistic D is the maximum absolute difference, i.e., $D = 0.40613$.

The critical value is $1.36/\sqrt{5} = 0.6082$. Since $D < 0.6082$, we do not reject H_0.

(d)

(i) Depending on the purpose for which model is to be used, the simplest model that reflects reality is generally preferred.

(ii) If you try enough models, one will look good, even if it is not.

(iii) Choose the model with the highest *p*-value under the chi-square goodness-of-fit test. This is consistent with parsimony because the more complex tests have lesser degrees of freedom.

Question 2

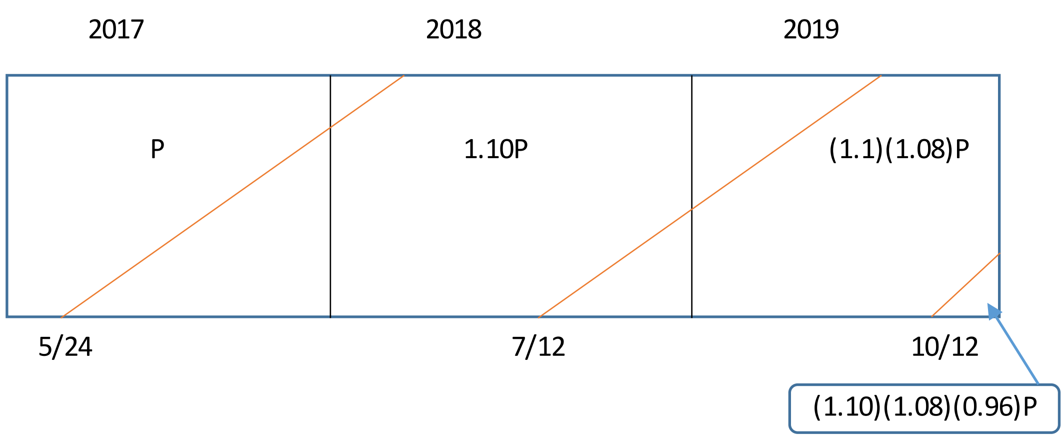
(a) The rates must cover the expected losses and expenses. Clearly, for the insurer to stay in business, income (premiums and investment income) must at least equal outgo (losses and expenses).

Ratemaking should produce rates that make adequate provision for contingencies. While the rates should cover the expected losses, there should be the cost of the unexpected (unusual weather patterns, wildfires, 100-year floods, etc.) also built into the rates. This is difficult to do while maintaining competitive rates.

The rates should encourage loss control. An appropriate risk categorization process will reduce claim frequency and/or severity. Such process not only allows that insurer to lower rates but to provide a service to society by reducing accidents, injuries, and property damage.

Rates must satisfy rate regulators. Almost all rates must be filed with and approved by state insurance department or other agencies. The basic requirement of the rate regulators is that rates must be adequate, not excessive, and not unfairly discriminatory.

(b)



2017

$$\text{Weighted Premium} = \left[\frac{\left(\frac{19}{24}\right)\left(\frac{19}{24}\right)}{2}\right]1.10P + \left\{1 - \left[\frac{\left(\frac{19}{24}\right)\left(\frac{19}{24}\right)}{2}\right]\right\}P = 1.031337P$$

$$\text{Current Rate Earned Premium} = (10{,}000)\left(\frac{(1.10)(1.08)(0.96)}{1.031337}\right) = 11{,}058.27$$

2018

$$\text{Weighted Premium} = \left[\frac{\left(\frac{5}{24}\right)\left(\frac{5}{24}\right)}{2}\right]P + \left[\frac{\left(\frac{5}{12}\right)\left(\frac{5}{12}\right)}{2}\right](1.10)(1.08)P$$

$$+\left\{1-\left[\frac{\left(\frac{5}{24}\right)\left(\frac{5}{24}\right)}{2}\right]-\left[\frac{\left(\frac{5}{12}\right)\left(\frac{5}{12}\right)}{2}\right]\right\}1.10P = 1.105469P$$

$$\text{Current Rate Earned Premium} = (12{,}000)\left(\frac{(1.10)(1.08)(0.96)}{1.105469}\right) = 12{,}380.05$$

2019

$$\text{Weighted Premium} = \left[\frac{\left(\frac{7}{12}\right)\left(\frac{7}{12}\right)}{2}\right](1.10)P + \left[\frac{\left(\frac{2}{12}\right)\left(\frac{2}{12}\right)}{2}\right](1.10)(1.08)(0.96)P$$

$$+\left\{1-\left[\frac{\left(\frac{7}{12}\right)\left(\frac{7}{12}\right)}{2}\right]-\left[\frac{\left(\frac{2}{12}\right)\left(\frac{2}{12}\right)}{2}\right]\right\}(1.10)(1.08)P = 1.172368$$

$$\text{Current Rate Earned Premium} = (8000)\left(\frac{(1.10)(1.08)(0.96)}{1.172368}\right) = 7782.40$$

(c) See the proof on Page 127 and 128 of Brown and Lennox.

(d)

We need to calculate the expected loss for each Type which = $E(S) = E(N)E(S)$

Safe ==> $(0.04)(3000) = 120$

Not So Safe ==> (0.10)(4000) = 400

Reckless ==> $(0.25)(5000) = 1250$

Differential for Safe $= 1.00$ since it is the base type

Differential for Not So Safe = $\frac{400}{120} = 3.33333$

Differential for Reckless $= \frac{1250}{120} = 10.41666$

(e)

Indicated Differential for Not So Safe = $(3.33333)\left(\frac{0.63}{0.60}\right) = 3.50000$

Indicated Differential for Class C = $(10.41666)\left(\frac{0.51}{0.60}\right) = 8.85416$

Question 3

(a) The earned part of the premium paid in January is 1,600

The earned part of the premium paid in March is $10/12 \times 1,800 = 1,500$

The earned part of the premium paid in May is $8/12 \times 1,200 = 800$

The earned part of the premium paid in July is $6/12 \times 1,200 = 600$

The total earned premium paid in 2021 is 4,500

The premium paid in 2020 and earned in 2021 is 1,300

The total earned premium in 2021 is 5,800

Expected total losses are $0.80 \times 5,800 = 4,640$

Reserve = Expected total losses – Claims Already Paid $= 4,640 - 2,500 = 2,140$

(b)

$$\hat{f}_1 = \frac{1500+1750+1900+2200+2900}{1000+1100+1200+1500+2000} = 1.50735$$

$$\hat{f}_2 = \frac{1700+1775+2200+2500}{1500+1750+1900+2200} = 1.11224$$

$$\hat{f}_3 = \frac{1800+1825+2350}{1700+1775+2200} = 1.05286$$

$$\hat{f}_4 = \frac{1850+1870}{1800+1825} = 1.02621$$

$$\hat{f}_5 = \frac{1875}{1850} = 1.01351$$

AY	0	1	2	3	4	5	OCR
16	1,000	1,500	1,700	1,800	1,850	1,875	
17	1,100	1,750	1,775	1,825	1,870	1,895.3	25.3
18	1,200	1,900	2,200	2,350		2,444.2	94.2
19	1,500	2,200	2,500			2,737.7	237.6

20	2,000	2,900				3,532.1	632.1
21	2,500					4,589.8	2089.8

The total OCR is 3,079.0

The total sum paid in 2021 is :

$$2,500 + (2,900 - 2,000) + (2,500 - 2,200) + (2,350 - 2,200) + (1,870 - 1,825) + (1,875 - 1,850) = 3,920$$

(c) The BF estimate of projected 2021 AY claims is

$$\tilde{C}_{2021,5} = \hat{\beta}_0 \, \hat{C}_{2021,5} + (1 - \hat{\beta}_0)\mu_{2021}$$

$$\hat{\beta}_0 = \frac{1}{\hat{f}_0 \times \hat{f}_1 \times \cdots \hat{f}_4} = 0.54469$$

$\hat{C}_{2021,5}$ is the Chain Ladder estimated projected claims = 4589.8

μ_{2021} is the LR estimated projected claims = 4640.0

$$\Rightarrow \tilde{C}_{2021,5} = 4,612.7 \Rightarrow \text{OCR for AY } 2021 = 4,612.7 - 2500 = 2112.7$$

(d) BF Advantage – much less reliance on a single data point in the most recent AY
BF Disadvantage – relies on subjective estimate of loss ratio (and adequacy of premiums).

Question 4

Type	propn	λ	Severity	$E[X]=\alpha\theta$	$Var[X]=\alpha\theta^2$
S	0.5	0.04	$\alpha=3,\theta=1000$	3000	$(3)(1000)^2=3,000,000$
NSS	0.3	0.10	$\alpha=4,\theta=1000$	4000	$(4)(1000)^2=4,000,000$
R	0.2	0.25	$\alpha=5,\theta=1000$	5000	$(5)(1000)^2=5,000,000$

(a) We can calculate the probability of a claim arising from a random policyholder as:

$$\Pr(S)\Pr(claim)=(0.5)(0.04)=0.02$$
$$\Pr(NSS)\Pr(claim)=(0.3)(0.10)=0.03$$
$$\Pr(R)\Pr(claim)=(0.2)(0.25)=0.05$$
$$Total=0.02+0.03+0.05=0.10$$

Given that a claim arises, we have

$$\Pr(S\mid Claim)=0.02/0.1=0.2$$
$$\Pr(NSS\mid Claim)=0.03/0.10=0.3$$
$$\Pr(R\mid Claim)=0.05/0.10=0.5$$

So $E\left[Var\left[X\right]\right]=(0.2)(3,000,000)+(0.3)(4,000,000)+(0.5)(5,000,000)=4,300,000$

(b) $E[X]=(0.2)(3000)+(0.3)(4000)+(0.5)(5000)=4300$

$$E[X^2]=(0.2)(3000)^2+(0.3)(4000)^2+(0.5)(5000)^2=19,100,000$$
$$\Rightarrow Var\left[E\left[X\right]\right]=19,100,000-(4300)^2=610,000$$

(c) $K=\dfrac{E[Var]}{V[E]}=\dfrac{4,300,000}{610,000}=7.04918$

$$\Rightarrow Z=\frac{N}{N+K}=\frac{1}{1+7.04918}=0.12424$$

$$Estimated\ \ Severity=(0.12424)(20,000)+(1-0.12424)(4300)=6250.51$$

(d) $E[S]=E[N]E[X]$ and $Var[S]=E[N]Var[X]+\left(E[X]\right)^2 Var[N]$

Type of Driver	E(S)	Var(S)
Safe	(0.04) (3000)=120	$(0.04)(3{,}000{,}000)+(3000)^2(0.04)=480{,}000$
Not So Safe	(0.1) (4000) =400	$(0.1)(4{,}000{,}000)+(4000)^2(0.1)=2{,}000{,}000$
Reckless	(0.25) (5000) =1250	$(0.25)(5{,}000{,}000)+(5000)^2(0.25)=7{,}500{,}000$

$$EPV=(0.5)(480,000)+(0.3)(2,000,000)+(0.2)(7,500,000)=2,340,000$$

(e) $E[S]=E[N]E[X]$ and $Var[S]=E[N]Var[X]+\left(E[X]\right)^2 Var[N]$

Type of Driver	E(S)	Var(S)
Safe	(0.04)(3000)=120	$(0.04)(3{,}000{,}000)+(3000)^2(0.04)=480{,}000$
Not So Safe	(0.1)(4000)=400	$(0.1)(4{,}000{,}000)+(4000)^2(0.1)=2{,}000{,}000$
Reckless	(0.25)(5000)=1250	$(0.25)(5{,}000{,}000)+(5000)^2(0.25)=7{,}500{,}000$

$$EPV=(0.5)(480,000)+(0.3)(2,000,000)+(0.2)(7,500,000)=2,340,000$$

Question 5

(a) $\Pr[M_{20} > Q_{0.95}(X)] = 1 - \Pr[M_{20} \le Q_{0.95}(X)] = 1 - F_X\left(Q_{0.95}(X)\right)^{20}$

$= 1 - (0.95)^{20} = 1 - 0.35849 = 0.64151$

(b) $\alpha(n) = 1 - \dfrac{1}{n} = \Pr\left[X \le Q_{\alpha(n)}(X)\right] = 1 - e^{-Q_{\alpha(n)}(X)}$

$\Rightarrow \dfrac{1}{n} = e^{-Q_{\alpha(n)}(X)} \Rightarrow n = e^{Q_{\alpha(n)}(X)}$

$\Rightarrow \log n = Q_{\alpha(n)}(X)$

(c) $\mathrm{ES}_{\alpha(n)}(X) = \mathrm{E}[X \mid X > Q_{\alpha(n)}(X)]$ as X is continuous

$= Q_{\alpha(n)}(X) + E[X - Q_{\alpha(n)}(X) \mid X > Q_{\alpha(n)}(X)]$

$= Q_{\alpha(n)}(X) + E[X]$ from the memoryless property

$= \log n + 1$

(d) The exponential distribution is in the MDA of the Gumbel EV distribution means that there exist deterministic functions c_n and d_n, such that as $n \to \infty$, the distribution of $\dfrac{M_n - d_n}{c_n}$ converges to the standard Gumbel distribution.

Let $H(x)$ denote the d.f. of the Gumbel distribution.

$$\Pr\left[\frac{M_n - d_n}{c_n} \le x\right] \approx H(x) \Rightarrow E\left[\frac{M_n - d_n}{c_n}\right] \approx 0.5772$$

$$\Rightarrow E\left[M_n - \log n\right] \approx 0.5772$$

$$\Rightarrow E[M_n] \approx 0.5772 + \log n$$

So we have

$$\left(Q_{\alpha(n)} = \log n\right) < \left(\mathrm{E}[M_n] = \log n + 0.5772\right) < \left(\mathrm{ES}_{\alpha(n)} = \log n + 1\right)$$

(e) Suppose we have a sample of $20N$ values of X_i, split into N blocks of 20, where N is a large number.

We would approximate $\mathrm{E}[M_{20}]$ by taking the average of the N block maxima. As $N \to \infty$ this approximation will converge to the true value.

We would approximate $\mathrm{ES}_{0.95}$ by taking the average of the largest N values. As $N \to \infty$ this approximation will converge to the true value. This cannot be smaller than the $\mathrm{E}[M_{20}]$

estimate, (as the boundary case is that the *N* largest values are also block maxima) but it could be bigger, if one or more blocks have several values that are larger than the smallest block maximum. Hence, $ES_{0.95}$ must be greater than $E[M_{20}]$, unless both are equal to the maximum possible value of X.

Also, for each block, the estimated 95% quantile of X lies between the 19th and 20th values (the smoothed empirical estimate would be the 19.95th value, estimated by linear interpolation). That means that the expected value of the 95% quantile is less than or equal to the expected value of the block maximum, with a block size of 20, i.e. that $E[M_{20}]$, must be greater than $Q_{0.95}$, unless both are equal to the maximum possible value of X.

Question 6

(a)

$$E\left[\min(Y,M)\right]=\int_0^M y\,f(y)dy+M(1-F(M))$$

$$=\int_0^M y\frac{\alpha\lambda^\alpha}{(\lambda+y)^{\alpha+1}}dy+M\left(\frac{\lambda}{\lambda+M}\right)^\alpha$$

$$=\left[-y\left(\frac{\lambda}{\lambda+y}\right)^\alpha\right]_0^M+\int_0^M\left(\frac{\lambda}{\lambda+y}\right)^\alpha dy+M\left(\frac{\lambda}{\lambda+M}\right)^\alpha$$

$$=-M\left(\frac{\lambda}{\lambda+M}\right)^\alpha+\left[-\frac{\lambda^\alpha}{(\alpha-1)(\lambda+y)^{\alpha-1}}\right]_0^M+M\left(\frac{\lambda}{\lambda+M}\right)^\alpha$$

$$=\frac{\lambda^\alpha}{\alpha-1}\left(\frac{1}{\lambda^{\alpha-1}}\right)-\frac{\lambda^\alpha}{\alpha-1}\left(\frac{1}{(\lambda+M)^{\alpha-1}}\right)$$

$$=\frac{\lambda}{\alpha-1}\left(1-\left(\frac{\lambda}{\lambda+M}\right)^{\alpha-1}\right)$$

(b)

$$L(\alpha,\lambda)=\left(\alpha^{97}\lambda^{97\alpha}\prod_{i=1}^{97}(\lambda+x_i)^{-\alpha-1}\right)\left(\frac{\lambda}{\lambda+M}\right)^{3\alpha}$$

$$l=\log L(\alpha,\lambda)=97\log\alpha+100\alpha\log\lambda-(\alpha+1)\sum_{i=1}^{97}\log(\lambda+x_i)-3\alpha\log(\lambda+M)$$

$$\frac{\partial l}{\partial\alpha}=\frac{97}{\alpha}+100\log\lambda-\sum_{i=1}^{97}\log(\lambda+x_i)-3\log(\lambda+M)$$

$$\frac{\partial l}{\partial\lambda}=\frac{100\alpha}{\lambda}-(\alpha+1)\sum_{i=1}^{97}\frac{1}{(\lambda+x_i)}-\frac{3\alpha}{\lambda+M}$$

Set equal to zero:

$$0=\frac{97}{\hat\alpha}+100\log\hat\lambda-\sum_{i=1}^{97}\log\left(\hat\lambda+x_i\right)-3\log\left(\hat\lambda+M\right)$$

$$0=\frac{100\hat\alpha}{\hat\lambda}-(\hat\alpha+1)\sum_{i=1}^{97}\frac{1}{\left(\hat\lambda+x_i\right)}-\frac{3\hat\alpha}{\hat\lambda+M}$$

(c) Substituting the given values gives $\hat{\alpha} = 3.3836$.

(d) The estimated covariance matrix is found by taking second partial derivatives of the log-likelihood, multiply by -1, substitute the unknown parameters with the MLEs, and invert, as follows:

$$\frac{\partial l}{\partial \alpha} = \frac{97}{\alpha} + 100\log\lambda - \sum_{i=1}^{97}\log\left(\lambda + x_i\right) - 3\log\left(\lambda + M\right)$$

$$\frac{\partial l}{\partial \lambda} = \frac{100\alpha}{\lambda} - (\alpha+1)\sum_{i=1}^{97}\frac{1}{\left(\lambda + x_i\right)} - \frac{3\alpha}{\lambda + M}$$

$$\frac{\partial^2 l}{\partial \alpha^2} = -\frac{97}{\alpha^2} \approx -\frac{97}{\hat{\alpha}^2} = -8.4725$$

$$\frac{\partial^2 l}{\partial \lambda^2} = -\frac{100\alpha}{\lambda^2} + (\alpha+1)\sum_{i=1}^{97}\frac{1}{\left(\lambda + x_i\right)^2} + \frac{3\alpha}{\left(\lambda + M\right)^2} \approx -5.483\times 10^{-7}$$

$$\frac{\partial^2 l}{\partial \lambda \partial \alpha} = \frac{100}{\lambda} - \sum_{i=1}^{97}\frac{1}{\left(\lambda + x_i\right)} - \frac{3}{\lambda + M} \approx 0.002113$$

$$\Rightarrow V \approx \begin{pmatrix} 8.4725 & -0.002113 \\ -0.002113 & 5.483\times 10^{-7} \end{pmatrix}^{-1} = \begin{pmatrix} 8.4725 & -0.002113 \\ -0.002113 & 5.483\times 10^{-7} \end{pmatrix}^{-1}$$

$$\approx \begin{pmatrix} 3.034 & 11,693 \\ 11,693 & 4.689\times 10^{7} \end{pmatrix}$$

Question 7

(a) From the formula sheet

$$E\left[\left(X\wedge x\right)^k\right]=e^{k\mu+k^2\sigma^2/2}\left\{\Phi\left(\frac{\log x-\mu-k\sigma^2}{\sigma}\right)\right\}+x^k\left(1-F(x)\right)$$

$$\Rightarrow\int_0^d x^k\ f(x)dx+d^k\left(1-F(d)\right)=e^{k\mu+k^2\sigma^2/2}\left\{\Phi\left(\frac{\log d-\mu-k\sigma^2}{\sigma}\right)\right\}+d^k\left(1-F(d)\right)$$

$$\Rightarrow\int_0^d x^k\ f(x)dx=e^{k\mu+k^2\sigma^2/2}\left\{\Phi\left(\frac{\log d-\mu-k\sigma^2}{\sigma}\right)\right\}$$

Also, we have

$$\int_0^\infty x^k f(x)dx=E\left[X^k\right]=e^{k\mu+k^2\sigma^2/2}$$

$$\Rightarrow\int_d^\infty x^k\ f(x)dx=e^{k\mu+k^2\sigma^2/2}\left\{1-\Phi\left(\frac{\log d-\mu-k\sigma^2}{\sigma}\right)\right\}$$

(b) The expected number of claims involving the reinsurer is

$$\lambda\left(1-F(3000)\right)=50\left(1-\Phi\left(\frac{\log 3000-6}{1.22475}\right)\right)=2.53$$

The expected value of an individual reinsurer claim is

$$\frac{1}{S(d)}\int_d^\infty\left(x-d\right)f(x)dx=\frac{1}{S(d)}\int_d^\infty\left(x\right)f(x)dx-d$$

$$=\frac{1}{\left(1-\Phi\left(\frac{\log d-\mu}{\sigma}\right)\right)}e^{\mu+\sigma^2/2}\left\{1-\Phi\left(\frac{\log d-\mu-\sigma^2}{\sigma}\right)\right\}-d=\frac{854.06\times 0.66036}{0.05069}-3000$$

$$=2722.4$$

Let S denote the aggregate claim cost before reinsurance and let S^* denote the reinsurer's aggregate claims cost. Let Y and Y^* denote the gross and reinsurer's claim severity random variables, respectively.

Then the aggregate cost of claims paid by the reinsurer is $E\left[S^*\right]=\lambda S(d)E\left[Y^*\right]=2.5345\times 2722.38=6{,}899.98$.

The aggregate cost of claims paid before reinsurance is $E\left[S\right]=\lambda E\left[Y\right]=50\times 854.06=42{,}702.94$.

Hence, the average cost of claims to the direct insurer, net of reinsurance, is $E\left[S-S^*\right]=35{,}803.0$

(c)

$$Var\left[S^*\right] = \lambda S(d) E\left[\left(Y^*\right)^2\right] = 2.5345 \times E\left[\left(Y^*\right)^2\right]$$

$$E\left[\left(Y^*\right)^2\right] = \frac{1}{S(d)}\int_d^\infty \left(x-d\right)^2 f(x)dx = \frac{1}{S(d)}\int_d^\infty x^2 f(x)dx - 2d\frac{1}{S(d)}\int_d^\infty xf(x)dx + d^2$$

$$= \frac{1}{S(d)}\left(e^{2\mu+2\sigma^2}\left(1-\Phi\left(\frac{\log d - \mu - 2\sigma^2}{\sigma}\right)\right)\right) - 2d\left(2722.8+3000\right) + d^2$$

$$= 51,037,050 - 34,334,283 + 9,000,000 = 25,702,767 = 5,069.79^2$$

$$\Rightarrow Var\left[S^*\right] = 8,071.2^2$$

(d) Now we have

$$Y^* = \left(1-\alpha\right)Y; \quad S = \sum_1^N Y; \quad S^* = \sum_1^N Y^*$$

So $E\left[S^*\right] = (1-\alpha)E\left[S\right]$ i.e. $6899.94 = (1-\alpha)42,702.94$

$$\Rightarrow \alpha = 0.8384$$

(e) The variance of reinsurer claims is

$$Var\left[S^*\right] = Var\left[(1-\alpha)S\right] = (1-\alpha)^2 Var\left[S\right] = \left(1-\alpha\right)^2 \lambda E\left[Y^2\right]$$

$$\Rightarrow SD\left[S^*\right] = (1-\alpha)SD\left[S\right]$$

$$E\left[Y^2\right] = e^{2\mu+2\sigma^2} = 3,269,017$$

$$\Rightarrow SD[S] = 12,784.7$$

$$\Rightarrow SD\left[S^*\right] = 2065.8$$

(f) The two reinsurance contracts have the same expected cost, but the standard deviation of costs under the excess loss policy is significantly larger than under the proportional reinsurance contract. As the reinsurance premium will reflect risk, we would expect a higher premium loading for the excess loss policy than for the proportional policy.

Question 8

(a) The hazard rate is $\lambda(x)$, say, where

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

$$\log S(x) = -2x + \log\left(x^2 + x + 1\right)$$

$$\Rightarrow \frac{d}{dx}\log S(x) = -2 + \frac{2x+1}{x^2+x+1} = \frac{-2x^2-1}{x^2+x+1}$$

$$\Rightarrow \lambda(x) = \frac{2x^2+1}{x^2+x+1}$$

(b) The hazard rate is increasing for all $x > 1$. For the exponential distribution, the hazard rate is constant, and the MEL is flat. This indicates that the given distribution is lighter tailed than the exponential.

(c)

$$e(d) = E\left[X - d \mid X > d\right] = \frac{1}{S(d)}\int_d^\infty S(x)dx$$

$$\int_d^\infty S(x)dx = \int_d^\infty e^{-2x}\left(x^2+x+1\right)dx = \left[-\frac{1}{2}e^{-2x}(x^2+x+1)\right]_d^\infty + \frac{1}{2}\int_d^\infty e^{-2x}\left(2x+1\right)dx$$

$$= (d^2+d+1)\frac{e^{-2d}}{2} + \left[-\frac{1}{4}e^{-2x}(2x+1)\right]_d^\infty + \frac{1}{2}\int_d^\infty e^{-2x}dx$$

$$= (d^2+d+1)\frac{e^{-2d}}{2} + \left(2d+1\right)\frac{e^{-2d}}{4} + \left[-\frac{e^{-2x}}{4}\right]_d^\infty$$

$$= (d^2+d+1)\frac{e^{-2d}}{2} + \left(2d+1\right)\frac{e^{-2d}}{4} + \frac{e^{-2d}}{4}$$

$$= \frac{e^{-2d}}{4}\left(2d^2+4d+4\right) = e^{-2d}\left(\tfrac{1}{2}d^2+d+1\right)$$

$$\Rightarrow e(d) = \frac{e^{-2d}\left(\frac{1}{2}d^2+d+1\right)}{e^{-2d}\left(d^2+d+1\right)} = \frac{\frac{1}{2}d^2+d+1}{d^2+d+1}$$

(d) As $d \to \infty, e(d) \to \frac{1}{2}$, and the gradient of the MEL function tends to 0. This indicates that the distribution is in the MDA of the Gumbel GEV distribution, which means that $\xi = 0$.

Question 9

(a) The assumptions are

1. $E\left[X_{ij} \mid \theta_i\right] = \mu(\theta_i)$ independent of *j*
2. $Var\left[X_{i,j} \mid \theta_j\right] \times m_{i,j} = v(\theta_i)$ independent of *j*

In this case:

$$E\left[X_{ij} \mid \theta_i\right] = E\left[\frac{1}{m_{i,j}} \sum_{l=1}^{m_{i,j}} Y_{i,j,l} \middle| \theta_i\right] = \frac{1}{m_{i,j}}\left(m_{i,j} E\left[Y_{i,j,l} \mid \theta_i\right]\right) = \mu(\theta_i)$$

$$Var\left[X_{ij} \mid \theta_i\right] = Var\left[\frac{1}{m_{i,j}} \sum_{l=1}^{m_{i,j}} Y_{i,j,l} \middle| \theta_i\right] = \frac{1}{m_{i,j}^2}\left(m_{i,j} Var\left[Y_{i,j,l} \mid \theta_i\right]\right) = \frac{v(\theta_i)}{m_{i,j}}$$

$$\Rightarrow Var\left[X_{i,j} \mid \theta_i\right] m_{i,j} = v(\theta_i)$$

(b)

(i) $Z_i = \frac{m_i}{m_i + \frac{\hat{v}}{\hat{a}}}$, where $m_i = \sum_{j=1}^{n} m_{i,j}$, and $\hat{a}$ and $\hat{v}$ are estimates of $\mathrm{Var}\left[\mu(\theta_i)\right]$ and $\mathrm{E}\left[v(\theta_i)\right]$ respectively.

(ii) The parameter *a* is a measure of the uncertainty in the estimation of $\mu = E\left[\mu(\theta_i)\right]$. Greater uncertainty in the prior mean estimate indicates that it should have less weight, and that the data, represented by $\overline{X}_i$, should have more weight. As Z_i is the weight given to the data, a larger value of *a* leads to a larger value of Z_i.

(iii) The parameter *v* is a measure of the variability within the data from the individual risk. More variability means that there is more uncertainty associates with the accuracy of $\overline{X}_i$ as an estimate of the risk premium. That indicates that we should put more weight on the prior mean, and less weight on $\overline{X}_i$, which is achieved through a smaller value for Z_i.

(iv) The m_i parameter is a measure of the volume of data available in the estimation of $\overline{X}_i$. If we have more data, then the estimate is more reliable, so we expect a larger value of Z_i.

(c) (i) From (a) we have $E\left[X_{ij} \mid \theta_i\right] = \mu(\theta_i)$

$$\Rightarrow E\left[\bar{X}_i \mid \theta_i\right] = \frac{\sum_{j=1}^{n} m_{ij} E\left[X_{ij} \mid \theta_i\right]}{\sum_{j=1}^{n} m_{ij}} = \mu(\theta_i)$$

$$\Rightarrow E[\hat{\mu}] = \frac{\sum_{i=1}^{r} Z_i E[\mu(\theta_i)]}{\sum_{i=1}^{r} Z_i} = E[\mu(\theta_i)] = \mu$$

(ii) Using a weighted average of unbiased estimates, where the weights are approximately equal to the variance of each estimate generates (approximately) the minimum variance unbiased estimator, compared with any other weighted average.

Question 10

(a)

$$L(\lambda)=\prod_{i=1}^{n}\lambda^2 x_i e^{-\lambda x_i} \Rightarrow \log L(\lambda)=2n\log\lambda-\lambda\sum_{i=1}^{n}x_i+\sum_{i=1}^{n}\log x_i$$

$$\Rightarrow \frac{\partial \log L(\lambda)}{\partial\lambda}=\frac{2n}{\lambda}-\sum_{i=1}^{n}x_i. \quad \text{Set equal to zero for} \quad \hat{\lambda}=\frac{2}{\overline{x}}$$

(b)

$$M_Y(t)=E\left[e^{2\lambda n\overline{X}t}\right]=E\left[e^{2\lambda t\sum_{i=1}^{n}x_i}\right]=\left(E\left[e^{2\lambda Xt}\right]\right)^n$$

as the X_i are i.i.d.

$$E\left[e^{2\lambda Xt}\right]=\int_0^\infty e^{2\lambda xt}\,\lambda^2 xe^{-\lambda x}dx=\int_0^\infty e^{-\lambda x(1-2t)}\,\lambda^2 xe^{-\lambda x}dx$$

$$=\frac{\lambda^2}{\left(\lambda(1-2t)\right)^2}\int_0^\infty e^{-\lambda x(1-2t)}\left(\lambda(1-2t)\right)^2 xe^{-\lambda x}dx$$

If (1-2t)>0, the integrand is the pdf of the original distribution, with a new λ parameter, so the integral is equal to 1, giving

$$E\left[e^{2\lambda Xt}\right]=(1-2t)^{-2} \Rightarrow E\left[e^{2\lambda n\overline{X}t}\right]=(1-2t)^{-2n}$$

Which is the MGF of the chi-square distribution, with $\nu = 4n$ degrees of freedom.

(c) The Information function is $-E\left[\frac{\partial^2\log L}{\partial\lambda^2}\right]=-E\left(-\frac{2n}{\lambda^2}\right)=\frac{2n}{\lambda^2}$.

The asymptotic variance of $\hat{\lambda}$ is $\left(-E\left[\frac{\partial^2\log L}{\partial\lambda^2}\right]\right)^{-1}=\frac{\lambda^2}{2n}$.

(d) (i) A 95% confidence interval for Y is $\left(Q_{0.025}(Y),Q_{0.975}(Y)\right)$, where Q_α is the α-quantile of the chi-square distribution with $4n = 40$ degrees of freedom.

Using CHISQ.INV in Excel, we find the CI for $Y=2n\lambda\overline{X}$ is (24.43, 59.34). Divide by $2n\overline{x}$ gives a 95% confidence interval for $\hat{\lambda}$ of (0.00341, 0.00829)

(ii) The MLE is asymptotically normally distributed. We have $\hat{\lambda} = \frac{2}{\overline{x}} = 0.005587$. The variance is approximately $\frac{\hat{\lambda}^2}{2n} = 1.5605 \times 10^{-6} = 0.00125^2$. So the approximate 95% CI is $0.005587 \pm 1.96 \times 0.00125 = (0.00314, 0.00804)$.

(e) The confidence intervals are quite close. The MLE CI is an asymptotic result, and we would expect less accuracy applying this to a sample of only 10 values. The chi-square confidence interval is an exact interval for *Y*, based on the actual sample size rather than on asymptotic results, so we would expect it to be more accurate.

Question 11

(a) (i) Let k_i denote the number of claims from the ith policy. The likelihood function is

$$L(\lambda) = \frac{\lambda^{\sum_{i=1}^{n} k_i} e^{-n\lambda}}{\prod_{i=1}^{n} k_i!} \Rightarrow \log L = \sum_{i=1}^{n} k_i \log \lambda - n\lambda - \sum_{i=1}^{n} \log k_i!$$

$$\Rightarrow \frac{\partial \log L}{\partial \lambda} = \frac{\sum_{i=1}^{n} k_i}{\lambda} - n \Rightarrow \hat{\lambda} = \frac{\sum_{i=1}^{n} k_i}{n} = \bar{k}$$

$$\Rightarrow \hat{\lambda} = \frac{1625 \times 0 + 307 \times 1 + 58 \times 2 + 9 \times 3 + 1 \times 4}{2000} = 0.227.$$

(ii) Using the Poisson probability function with the estimated value of λ, we have the following estimated frequency:

Number of claims (per policyholder)	0	1	2	3	4	5+
(Expected) Number of policyholders	1593.842	361.802	41.065	3.107	0.176	0.008

(iii) The chi-square test statistic for this model is

$$\chi^2 = \frac{(1625 - 1593.842)^2}{1593.842} + \frac{(307 - 361.802)^2}{361.802} + \frac{(58 - 41.065)^2}{41.065} + \frac{(10 - 3.291)^2}{3.291} = 29.571$$

(b) (i) The likelihood function is given by

$$L(\pi) = \pi^n (1-\pi)^{\sum_{i=1}^{n} k_i} \Rightarrow \log L = n \log \pi + \left(\sum_{i=1}^{n} k_i \right) \log(1-\pi)$$

$$\Rightarrow \frac{\partial \log L}{\partial \pi} = \frac{n}{\pi} - \frac{n\bar{k}}{1-\pi}.$$ Set equal to 0 for

$$\hat{\pi} = \frac{1}{\bar{k} + 1} = 0.8150$$

(ii) Using the estimated probability, we produce the following estimated frequency:

Number of claims (per policyholder)	0	1	2	3	4	5+
(Expected) Number of policyholders	1630.000	301.550	55.787	10.321	1.909	0.433

(iii) The chi-square test statistic for this model is 0.762.

(c) Since each model has one parameter to estimate, the corresponding degrees of freedom in the chi-square test are the same. Hence, because the chi-square test statistic based on the second model is much smaller, it will result in a much higher *p*-value, and is therefore a far better model.

Question 12

(a) The following procedure would be typical for insurance loss data model selection:
(i) Conduct preliminary investigation, such as histogram, summary statistics, of raw data;
(ii) Construct empirical distributions (eg use Kaplan-Meier for truncated/censored data);
(iii) Construct pictures such as *q-q* or *p-p* plots;
(iv) Conduct hypothesis tests: Kolmogorov-Smirnoff test, chi-square goodness-of-fit test;
(v) Calculate other criterion statistics such as SBC/BIC and AIC.
In addition, some other considerations to make include: (i) keep it simple if at all possible, and (ii) restrict the universe of potential models.

(b) The standard errors of the parameter estimates provide measure of the degree of accuracy of parameter estimates. Generally, it is preferred to have smaller standard errors. They can also be used for confidence intervals, the likely range of values of the parameter values. Standard errors can be evaluated based on the second derivative of the optimal value of the loglikelihood function, called the information matrix or the Hessian matrix.

(c) To test the null $H_0: \gamma = 1$, one can approximate the test statistic:

$$\frac{\widehat{\gamma}-1}{se(\widehat{\gamma})} = \frac{0.9956-1}{0.0638} = -0.06886$$

and if this is more than 2 (in absolute value, assuming two-sided test), then one can safely reject the null hypothesis. In this case, the test statistic is quite small so that we will not be able to reject the null hypothesis that the claims data may come from a Pareto.

(d) There are several hypothesis tests that can be considered to examine the goodness-of-fit of the data to the model being investigated. However, these tests generally fail to account for the number of parameters and the sample size. Alternatively, you may use the SBC statistic defined by $\ell - (r/2)\log(n)$ where ℓ is the log of the likelihood at the maximum, r is the number of parameters estimated, and n is the sample size. Generally we prefer a model with a larger SBC score. For the Pareto model, we have

$$\text{SBC}(\text{Pareto}) = -3515.941 - (2/2)\log(533) = -3522.22,$$

while for the Burr XII model, we have

$$\text{SBC}(\text{Burr}) = -3526.253 - (3/2)\log(533) = -3535.67.$$

Because the Pareto yields a larger SBC score, we would prefer this model over the Burr model.

(e) All the tests and criterion used indicate that the Pareto model is better than the Burr model for this data. The Pareto, with one fewer parameters, is also a more parsimonious model.

Question 13

(a) We are given that $Pr\left(\text{a policy incurs a claim by time } t_0\right) = 1 - e^{-\beta t_0}$ so that
$L\left(\beta;t_0\right) = \left(1-e^{-\beta t_0}\right)^M \left(e^{-\beta t_0}\right)^{n-M}$.

(b) We can maximize the loglikelihood given by
$\log L\left(\beta;t_0\right) = M\,\log\left(1-e^{-\beta t_0}\right) + (n-M)\log\left(e^{-\beta t_0}\right) = M\,\log\left(1-e^{-\beta t_0}\right) - (n-M)\beta t_0$.
Setting the derivative to zero

$$\frac{d\log L}{d\beta} = \frac{M\,t_0\,e^{-\beta t_0}}{1-e^{-\beta t_0}} - (n-M)t_0 = 0,$$

we have

$$\widehat{\beta} = -\frac{1}{t_0}\log\left(\frac{n-M}{n}\right).$$

(c) Use the asymptotic results of maximum likelihood estimates with standard error estimated using the information quantity. The 95% confidence interval can then be evaluated as
$\widehat{\beta} \pm 1.96\,se\left(\widehat{\beta}\right)$.

(d) Since the average time to a claim is $\mu = 1/\beta$, then its MLE is

$$\widehat{\mu} = 1/\widehat{\beta} = -\frac{t_0}{\log\left(\frac{n-M}{n}\right)}.$$

(e) Since μ is a function of β, i.e. $\mu = g(\beta)$, one can use the Delta method to estimate its variance or standard estimate:
$Var(\widehat{\mu}) = Var\left(\widehat{\beta}\right) \times \left[g'\left(\widehat{\beta}\right)\right]^2$.

Question 14

(a) X_1 is a mixture of two exponential distributions, i.e.

$X_1 = \beta\, Y_1 + (1-\beta)\, Y_2$ where $Y_1 \sim \exp(1000)$ and $Y_2 \sim \exp(500)$

$\mathrm{E}\left[X_1\right] = \beta\, \mathrm{E}\left[Y_1\right] + (1-\beta)\mathrm{E}\left[Y_2\right] = \beta(1000) + (1-\beta)500$

Set equal to 800 $\Rightarrow \beta = 0.6$

(b)

$$E\left[X_1 - 200 \middle| X_1 > 200\right] = \frac{1}{\Pr\left[X_1 > 200\right]} \int_{200}^{\infty} (x-200)\left(\beta f_{Y_1}(x) + (1-\beta) f_{Y_2}(x)\right)dx$$

$$\Pr\left[X_1 > 200\right] = \beta S_{Y_1}(200) + (1-\beta) S_{Y_2}(200)$$

where $S_Y(x) = 1 - F_Y(x)$ is the survival function of Y.

$$\Rightarrow \Pr\left[X_1 > 200\right] = 0.6\left(e^{-200/1000}\right) + 0.4\left(e^{-200/500}\right) = 0.49124 + 0.26813 = 0.75934$$

Also, from the memoryless property of the exponential distribution:

$$\int_{200}^{\infty} (x-200) f_Y(x)dx = \Pr\left[Y > 200\right] \times E\left[Y\right]$$

$$\Rightarrow E\left[X_1 - 200 \middle| X_1 > 200\right] = \frac{0.49124(1000) + 0.26813(500)}{0.75934} = 823.46.$$

(c) Let Q denote the 90% VaR of X_1. Then

$\Pr\left[X_1 > Q\right] = 0.6\Pr\left[Y_1 > Q\right] + 0.4\Pr[Y_2 > Q] = 0.1$

$\Rightarrow 0.6e^{-Q/1000} + 0.4e^{-Q/500} = 0.1.$

Let $Z = e^{-Q/1000}$, so that $Z^2 = e^{-Q/500}$. Then

$$Z = \frac{-0.6 \pm \sqrt{0.6^2 + 4(0.4)(0.1)}}{0.8} = 0.1514 \text{ (ignore } Z < 0 \text{ solution)}$$

$\Rightarrow Q = 1888.$

(d)

$$f_{X_2}(x) = \begin{cases} k(0.002)e^{-0.002x} & 0 < x \le 500 \\ c\dfrac{\alpha\theta^{\alpha}}{(\theta+x)^{\alpha+1}} & 500 < x \end{cases} \qquad \text{for constants } k \text{ and } c.$$

$$F_{X_2}(500) = k\left(1-e^{-1}\right) = F_{X_1}(500)0.6\left(1-e^{-0.5}\right)+0.4\left(1-e^{-1}\right) = 0.48893$$

$$\Rightarrow k = 0.77348$$

$$S_{X_2}(500) = c\left(\frac{2000}{2500}\right)^3 = 1-0.48893 = 0.51107 \Rightarrow c = 0.99818$$

(e) Let Q denote the 90% VaR of X_2. Then

$$\Pr\left[X_2 > Q\right] = 0.1 \Rightarrow c\left(\frac{2000}{2000+Q}\right)^3 = 0.1$$

$$\Rightarrow Q = 2306.3$$

(f) The tail of X_1 will behave very similarly to the exp(1000) distribution. The exponential distribution is thin tailed; it is in the MDA of the Gumbel distribution, and it has a constant hazard function.

The tail of X_2 will behave very similarly to the Pareto(3,2000) distribution. This is much fatter tailed. It lies in the MDA of the Frechet distribution, which is the fattest tailed of the EV distributions. It has decreasing hazard function.

Question 15

(a)

$$\Pr[Y>y]=\int_0^\infty \Pr[Y>y\,|\,\Lambda=\lambda]\; f_\Lambda(\lambda)\,d\lambda$$

$$=\int_0^\infty e^{-y\lambda}\,\frac{\lambda^{\alpha-1}e^{-\lambda/\theta}}{\theta^\alpha\Gamma(\alpha)}\,d\lambda$$

Let $\tau=\theta^{-1}$ -- this will just make the derivation less messy;

$$\Pr[Y>y]=\int_0^\infty \frac{\lambda^{\alpha-1}e^{-\lambda(\tau+y)}\tau^\alpha}{\Gamma(\alpha)}\,d\lambda=\int_0^\infty \frac{\lambda^{\alpha-1}e^{-\lambda(\tau+y)}\left(\tau+y\right)^\alpha}{\Gamma(\alpha)}\,d\lambda\left(\frac{\tau}{\tau+y}\right)^\alpha$$

$$\Pr[Y>y]=\left(\frac{\tau}{\tau+y}\right)^\alpha=\left(\frac{500}{500+y}\right)^4\Rightarrow F_Y(y)=1-\left(\frac{500}{500+y}\right)^4.$$

In the penultimate line, we multiply and divide by $(\tau+y)^\alpha$ so that the integrand is the pdf of a gamma distribution, which must integrate to 1.0.

(b) Let N denote the number of loss events, and N^* the number of payments.

Let $q=\Pr[Y_j>100]=\left(\frac{500}{600}\right)^4$

Let $I_j=\begin{cases}1 & \text{if } Y_j>100\\ 0 & \text{otherwise}\end{cases}$

The probability generating function of N^* is

$$P_{N^*}(z)=\sum_{n=0}^\infty\sum_{k=0}^\infty \Pr[N=n]\Pr\left[N^*=k\,\middle|\,N=n\right]z^k$$

$$=\sum_{n=0}^\infty \Pr[N=n]\left(\sum_{k=0}^n \Pr[I_1+I_2+\cdots+I_n=k]\,z^k\right)$$

$$=\sum_{n=0}^\infty \Pr[N=n]\left(P_I(z)\right)^n\quad=P_N\left(P_I(z)\right)$$

as the PGF of a sum of independent random variables is the product of their PGFs. The I_j indicator variables are Binomial(1, q) distributed, so

$$P_I(z)=(1+q(z-1)$$

$$P_N(z)=\left(1-\beta(z-1)\right)^{-r}$$

$$\Rightarrow P_{N^*}(z)=\left(1-\beta\left(1+q(z-1)-1\right)\right)^{-r}$$

$$=\left(1-\beta q(z-1)\right)^{-r}$$

Which shows that the number of payments also has a negative binomial distribution, with the same r parameter, and with $\beta^* = \beta q = 0.72338$.

(c) From the distribution function, we know that Y has a Pareto distribution with parameters $\alpha = 4, \theta = 500$.

$$Y^P = Y - 100 \mid Y > 100$$

$$\Pr\left[Y^P > y\right] = \Pr\left[Y > 100 + y \mid Y > 100\right] = \frac{\Pr\left[Y > 100 + y\right]}{\Pr\left[Y > 100\right]}$$

$$= \left(\frac{\theta}{\theta + 100 + y}\right)^{\alpha} \left(\frac{\theta + 100}{\theta}\right)^{\alpha}$$

$$= \left(\frac{600}{600 + y}\right)^4$$

(d) Let K denote the discretized random variable associated with Y^P, using a step size of 50, so that $\Pr\left[K = k\right]$ is an approximation for the probability that the claim size is 50k. Let f_j denote the probability function for K, then

$$f_0 = \Pr\left[K = 0\right] = \Pr\left[Y^P \le 25\right] = 1 - \left(\frac{600}{625}\right)^4 = 0.15065$$

$$f_1 = \Pr\left[K = 1\right] = \Pr\left[25 < Y^P \le 75\right] = \left(\frac{600}{625}\right)^4 - \left(\frac{600}{675}\right)^4 = 0.22505$$

$$f_2 = \Pr\left[K = 2\right] = \Pr\left[75 < Y^P \le 125\right] = \left(\frac{600}{675}\right)^4 - \left(\frac{600}{725}\right)^4 = 0.15521$$

The number of payments has a negative binomial distribution with parameters

$$r = 2, \beta^* = 0.72338$$

$$a = \frac{\beta^*}{1 + \beta^*} = 0.419745; \quad b = \frac{(r-1)\beta^*}{1 + \beta^*} = 0.419745$$

Let g_k denote the aggregate payment probability function (in units of 50).

$$g_0 = P_{N^*}(f_0) = \left(1 - \beta^*(f_0 - 1)\right)^{-2} = 0.38369$$

$$g_1 = \frac{(a + b)\, f_1 g_0}{1 - a f_0} = 0.07738$$

$$g_2 = \frac{\left(a + \frac{b}{2}\right) f_1 g_1 + \left(a + \frac{2b}{2}\right) f_2 g_0}{1 - a f_0} = 0.06507$$

$$\Rightarrow \Pr\left[S \le 100\right] \approx 0.38369 + 0.07738 + 0.06507 = 0.52614.$$

Question 16

(a) Given that the aggregate payment amounts are multiples of 2, the probability function for S_1 can be written as $\Pr(S_1 = 2k)$, $k = 0,1,2,\ldots$

Using the convolution formula, we have

$$\Pr(S_1 = 2k) = \sum_{n=0}^{\infty} p_n \cdot f_X^{*n}(2k).$$

Note that n convolutions of Bernoulli X_1 (i.e., sum of n i.i.d. Bernoulli random variables) is a Binomial distribution with parameters n and $p = 0.7$. Therefore,

$$\Pr(S_1 = 2k) = \sum_{n=0}^{\infty} p_n \cdot f_X^{*n}(2k) = \sum_{n=k}^{\infty} \frac{2^n}{n!} e^{-2} \cdot \binom{n}{k} 0.7^k 0.3^{n-k}.$$

The second equation starts with $n = k$ because there should be at least k number of claims to achieve an aggregate payment of $2k$.

(b) Let $M_j(t)$ $(j = 1,2)$ be the moment generating function (mgf) of the severity for S_1 and S_2, respectively: $M_1(t) = 0.3 + 0.7e^{2t}$ and $M_2(t) = 0.6 + 0.4e^{4t}$.

Then, the mgf for S_j is

$$M_{S_j}(t) = P_{N_j}\big(M_j(t)\big) = \exp\big(\lambda_j(M_j(t) - 1)\big)$$

and the mgf for S is given by

$$\begin{aligned} M_S(t) = \mathrm{E}[e^{tS}] &= M_{S_1}(t) \cdot M_{S_2}(t) \\ &= \exp\big(\lambda_1 M_1(t) + \lambda_2 M_2(t) - (\lambda_1 + \lambda_2)\big) \\ &= \exp\big(\lambda(M(t) - 1)\big), \end{aligned}$$

where $\lambda = \lambda_1 + \lambda_2 = 3$ and $M(t) = \frac{2}{3}M_1(t) + \frac{1}{3}M_2(t) = 0.4 + \frac{1.4}{3}e^{2t} + \frac{0.4}{3}e^{4t}$.

Therefore, $S = S_1 + S_2$ is also a compound Poisson process with $\lambda = 3$ and a secondary distribution on 0, 2, 4 with probability 0.4, 1.4/3 and 0.4/3, respectively.

(c) Let $g_k = \Pr(S = 2k)$ be the probability that the total claim payment amount of both lines be $2k$, for $k = 0,1,2,\ldots$ Also, denote the secondary distribution for S as $f_j = \Pr(X = 2j)$, for $j = 0,1,2$.

Then,

$$g_0 = P_N(f_0) = \exp\big(\lambda(f_0 - 1)\big) = \exp\big(3(0.4 - 1)\big) = 0.165299.$$

For compound Poisson, the recursive formula is

$$g_k = \frac{\lambda}{k}\sum_{j=1}^{k\wedge 2} j f_j g_{k-j}\,, \qquad k = 1,2,\dots$$

where j is capped at $k \wedge 2$ because maximum severity is 4 (or j be 2). Therefore,

$$g_1 = \lambda f_1 g_0 = 3\cdot\frac{1.4}{3}\cdot 0.165299 = 0.231418,$$

And

$$\begin{aligned} g_2 = \Pr(S=4) &= \frac{\lambda}{2}(f_1 g_1 + 2 f_2 g_0) \\ &= \frac{3}{2}\left(\frac{1.4}{3}\cdot 0.231418 + 2\cdot\frac{0.4}{3}\cdot 0.165299\right) \\ &= 0.195053. \end{aligned}$$

(d) The net stop-loss reinsurance premium satisfies

$$E[(S-d)_+] = E[S] - E[S\wedge d].$$

We determine each item on the right side of the equation. First,

$$E[S] = E[N]\cdot E[X] = 3\left(0.4\cdot 0 + \frac{1.4}{3}\cdot 2 + \frac{0.4}{3}\cdot 4\right) = 4.4.$$

Second,

$$\begin{aligned} E[S\wedge 4] &= 2\Pr(S=2) + 4\Pr(S\geq 4) \\ &= 2g_1 + 4(1-g_0-g_1) \\ &= 2\cdot 0.231418 + 4(1-0.165299-0.231418) = 2.875968. \end{aligned}$$

Finally, we have

$$E[(S-4)_+] = E[S] - E[S\wedge 4] = 4.4 - 2.875968 = 1.524.$$

Question 17

(a)

$$L(\theta) = f(120)\,f(640)\,f(700)\frac{f(820)}{S(100)}\frac{f(1220)}{S(100)}S(1500)$$

$$= \left(\frac{e^{-120/\theta}}{\theta}\right)\left(\frac{e^{-640/\theta}}{\theta}\right)\left(\frac{e^{-700/\theta}}{\theta}\right)\left(\frac{e^{-820/\theta}}{\theta e^{-100/\theta}}\right)\left(\frac{e^{-1220/\theta}}{\theta e^{-100/\theta}}\right)e^{-1500/\theta}$$

$$= \theta^{-5}\,e^{-4800/\theta}$$

(b)

$$\Rightarrow l(\theta) = \log L(\theta) = -5\log\theta - \frac{4800}{\theta}$$

$$\Rightarrow \frac{dl}{d\theta} = -\frac{5}{\theta} + \frac{4800}{\theta^2} \Rightarrow \hat{\theta} = \frac{4800}{5} = 960$$

(c)

$$V[\hat{\theta}] \approx -\left(\frac{d^2 l}{d\theta^2}\right)^{-1}_{\theta=\hat{\theta}}$$

$$\frac{d^2 l}{d\theta^2} = \frac{5}{\theta^2} - \frac{9600}{\theta^3} \Rightarrow V[\hat{\theta}] \approx \frac{\hat{\theta}^3}{9600-5\hat{\theta}} = 184,320$$

$$\Rightarrow SD[\hat{\theta}] \approx 429.3$$

(d)

$$Y = \max\left(X, 1000\right)$$

$$E[Y] = \theta\left(1 - e^{-1000/\theta}\right)$$ from the formula sheet

$$\approx 960(0.64713) = 621.25$$

(e)

$$g(\theta) = \theta\left(1 - e^{-1000/\theta}\right) \Rightarrow g\left(\hat{\theta}\right) = 621.25$$

$$V\left[g\left(\hat{\theta}\right)\right] \approx \left(g'\left(\hat{\theta}\right)\right)^2 V\left[\hat{\theta}\right]$$ by the delta method

$$g'\left(\theta\right) = 1 - e^{-1000/\theta} + \frac{1000}{\theta}e^{-1000/\theta} \Rightarrow g'\left(\hat{\theta}\right) = 1.0147$$

$$\Rightarrow V\left[g\left(\hat{\theta}\right)\right] \approx 1.0147^2\left(184,320\right) = 189,780 = 435.6^2$$

$\Rightarrow$ 80% CI is approximately $621.25 \pm 1.28(435.6)$

$$= (62.96, 1179.54)$$

Question 18

(a) The SBC = the maximum value of the Log Likelihood function less the log of the sample size times one half the number of parameters

The Neg Binomial has two parameters therefore

$$SBC = l(\theta) - \frac{2}{2} ln400 = -424.5 - 5.99 = 30.49$$

(b)

$$T = \begin{cases} 2ln(L_1) - 2ln(L_0) \\ 2(-426.2 + 427.8) \\ 3.2 \end{cases}$$

If H_1 is true then the test stat follows a Chi-Squared with 1 degree of freedom. We have $T = 3.2 < 3.841 = \chi^2_{0.95}$ and therefore we do not reject H_0.

(c)

	Poisson	*ZM Poisson*	*Geometric*	*ZM Geometric*	*Neg Binomial*	*ZM Neg Binomial*
χ^2 ***Statistic***	*6.07*	*3.29*	*14.53*	*1.26*	*5.15*	*1.14*
χ^2 ***Degrees of Freedom***	*2*	*1*	*3*	*1*	*1*	*1*
χ^2 ***P-value***	*2.5%-5%*	*5-10%*	*<0.5%*	*10-90%*	*1-2.5%*	*10-90%*

(d) Possible answers

- Graphing
- Experience with similar populations
- Likelihood ratio test between similar models
- Simplicity to program/use/explain

The Kolmogorov Smirnov test is meant for continuous distribution so not applicable

(e) Possible choices

- Geometric based on SBC, simplicity
- ZM Neg Binomial based on likelihood, Chi-squared test

Question 19

(a)

$$l(\mu,\sigma) = -\sum_{i=1}^{n} ln(x_j) - nln\sigma - nln\sqrt{2\pi} - \sum_{i=1}^{n} \frac{\left(ln(x_j)-\mu\right)^2}{2\sigma^2}$$

$$\frac{\partial l}{\partial \mu} = \begin{cases} -\frac{-2}{2\sigma^2}\sum_{i=1}^{n}[ln(x_j)-\mu] \\ \frac{1}{\sigma^2}\left[\sum_{j=1}^{n} ln(x_j) - n\mu\right] \end{cases} = 0. \therefore \hat{\mu} = \frac{1}{n}\sum_{j=1}^{n} ln(x_j)$$

(b)

$$\frac{\partial l}{\partial \sigma} = -\frac{n}{\sigma} + \frac{2}{2}\sigma^{-3}\sum_{i=1}^{n}\left(ln(x_j)-\mu\right)^2 = 0$$

$$\therefore n\sigma^2 = \sum_{i=1}^{n}\left(ln(x_j)-\mu\right)^2$$

(c)

$$\frac{\partial^2 l}{\partial \mu^2} = \frac{-n}{\sigma^2}$$

$$\frac{\partial^2 l}{\partial \mu \partial \sigma} = \frac{-2}{\sigma^3}\left[\sum_{j=1}^{n} ln(x_j) - n\mu\right] = 0$$

$$\frac{\partial^2 l}{\partial \sigma^2} = \begin{cases} n\sigma^{-2} - 3\sigma^{-4}\sum_{i=1}^{n}\left(ln(x_j)-\mu\right)^2 \\ n\sigma^{-2} - 3\sigma^{-4}(n\sigma^2) \\ -2n\sigma^{-2} \end{cases}$$

$$COV(\mu,\sigma) = -\begin{vmatrix} \frac{\partial^2 l}{\partial \mu^2} & \frac{\partial^2 l}{\partial \mu \partial \sigma} \\ \frac{\partial^2 l}{\partial \mu \partial \sigma} & \frac{\partial^2 l}{\partial \sigma^2} \end{vmatrix}^{-1} = -\begin{vmatrix} \frac{-n}{\sigma^2} & 0 \\ 0 & \frac{-2n}{\sigma^2} \end{vmatrix}^{-1} = \begin{vmatrix} \frac{\sigma^2}{n} & 0 \\ 0 & \frac{\sigma^2}{2n} \end{vmatrix}$$

(d)

$$g = exp\left[\mu + \frac{\sigma^2}{2}\right]$$

$$\partial g = \begin{bmatrix} \frac{\partial g}{\partial \mu} \\ \frac{\partial g}{\partial \sigma} \end{bmatrix} = \begin{bmatrix} exp\left[\mu + \frac{\sigma^2}{2}\right] \\ \sigma exp\left[\mu + \frac{\sigma^2}{2}\right] \end{bmatrix}$$

$$\widehat{Var(g)} = \begin{cases} \left| exp\left[\mu + \frac{\sigma^2}{2}\right] \quad \sigma exp\left[\mu + \frac{\sigma^2}{2}\right] \right| \begin{vmatrix} \frac{\sigma^2}{n} & 0 \\ 0 & \frac{\sigma^2}{2n} \end{vmatrix} \begin{bmatrix} exp\left[\mu + \frac{\sigma^2}{2}\right] \\ \sigma exp\left[\mu + \frac{\sigma^2}{2}\right] \end{bmatrix} \\ \frac{\sigma^2 exp[2\mu + \sigma^2]}{n} + \frac{\sigma^4 exp[2\mu + \sigma^2]}{2\,n} \end{cases}$$

Question 20

The cumulative claim run-off triangle is:

	Development Year (DY), j				
Accident Year (AY) i	**0**	1	2	3	4
0	1023	1155	1343	1386	1396
1	1358	1708	1905	1930	
2	1283	1566	1738		
3	1503	2011			
4	1536				
$\hat{f}_j$	1.2464	1.1258	1.0209	1.0072	
$\hat{\lambda}_j$	1.4428	1.1576	1.0283	1.0072	
$\hat{\beta}_j$	0.6931	?	0.9725	0.9928	1.000
$\hat{\gamma}_j$	0.6931	?	0.1086	0.0204	0.0072

(a)

$$\hat{f}_0 = \frac{1155+1708+1566+2011}{1023+1358+1283+1503} = 1.2464; \quad \hat{f}_1 = \frac{1343+1905+1738}{1155+1708+1566} = 1.1258$$

$$\hat{f}_2 = \frac{1386+1930}{1343+1905} = 1.0209; \quad \hat{f}_3 = \frac{1396}{1386} = 1.0072$$

$$\hat{\lambda}_3 = \hat{f}_3 = 1.0072; \hat{\lambda}_2 = \hat{f}_2 \times \hat{\lambda}_3 = 1.0283; \hat{\lambda}_1 = \hat{f}_1 \times \hat{\lambda}_2 = 1.1576; \hat{\lambda}_0 = \hat{f}_0 \times \hat{\lambda}_1 = 1.4428.$$

$$\hat{\beta}_1 = \frac{1}{\hat{\lambda}_1} = 0.8639; \quad \hat{\gamma}_1 = \hat{\beta}_1 - \hat{\beta}_0 = 0.1708.$$

(b) The ultimate projected claims cost for the i-th accident year is $\hat{C}_{i,4} = C_{i,4-i} \times \hat{\lambda}_i$, where $C_{i,4-i}$ is the latest available cumulative claims data for the i-th AY.

The estimated outstanding claims for AY i is $\hat{R}_i = \hat{C}_{i,4} - C_{i,4-i}$

AY, i	$\hat{C}_{i,4}$	$\hat{R}_i$

0	1396	0
1	1944	14
2	1787	49
3	2328	317
4	2216	680

The total of the outstanding claims is $\hat{R} = 1060$.

(c)

Assumption 1: $X_{i,j} \mid \theta_i$ are independent wrt DY, j.

Assumption 2: There exist γ_j such that $\mathrm{E}\left[X_{i,j} \mid \theta_i\right] = \gamma_j \mu(\theta_i)$ and $\mathrm{Var}\left[X_{i,j} \mid \theta_i\right] = \gamma_j v(\theta_i)$

Assumption 3: $\left(X_{i,j} \mid \theta_i\right)$, $\left(X_{l,j} \mid \theta_l\right)$ are independent and identically distributed for all $i \neq l$.

(d) Using the formula sheet, with $I = J = 4$,

(i) $s_2^2 = \frac{1}{2}\sum_{j=0}^{2} \hat{\gamma}_j \left(\frac{X_{2,j}}{\hat{\gamma}_j} - \hat{C}_{2,4}\right)^2$

$$= \left\{0.6931\left(\frac{1283}{0.6931} - 1787\right)^2 + 0.1708\left(\frac{283}{0.1708} - 1787\right)^2 + 0.1086\left(\frac{172}{0.1086} - 1787\right)^2\right\}$$

$= 5112$

(ii) $m_0 = \hat{\beta}_4 = 1.00$; $m_1 = \hat{\beta}_3 = 0.9929$; $m_2 = \hat{\beta}_2 = 0.9725$;

$m_3 = \hat{\beta}_1 = 0.8639$; $m_4 = \hat{\beta}_0 = 0.6931$.

$m = \sum_{i=0}^{4} m_i = 4.5224$; $\sum_{i=0}^{4} m_i^2 = 4.1583$.

$\bar{C} = \frac{1396 + 1930 + \cdots + 1536}{4.5224} = 1904.$

$$\hat{a} = \left\{\frac{1.00(1396-1904)^2 + 0.9929(1944-1904)^2 + \cdots 0.6931(2216-1904)^2 - 4(30582)}{4.5224 - 4.1583/4.5224}\right\}$$

$= \frac{373415}{3.6029} = 103643$

(iii) The estimator $\hat{v}$ is an estimate of $v\left(\theta_i\right)$, which is a measure of the variance of the estimated ultimate claims cost, based on the incremental claims information from AY i, DYj. This measures the uncertainty in the estimated ultimate claims based solely on the row (AY) information. The estimator $\hat{a}$ is a measure of the variance of $\mu\left(\theta_i\right)$, which measures the variability across different AYs. A small value for $\hat{v}$ indicates that there is little variation in payment patterns within each AY. A small value for $\hat{a}$ indicates that the payments patterns are similar across AYs, so the information from all the AYs is valuable in predicting claims in a single AY.

(iv) First, we calculate the Bühlmann-Straub estimate, then we iterate the credibility, using the m_i weights, and using the Bühlmann-Straub estimate as the 'prior' mean. That is,

$$Z_2 = \frac{m_2}{m_2 + \hat{v}/\hat{a}} = 0.7671$$

$$\hat{C}_{2,4}^{\text{BS}} = Z_2 \hat{C}_{2,4} + \left(1 - Z_2\right)\hat{\mu} = 1819$$

$$\hat{C}_{2,4}^{\text{BS2}} = m_2 \hat{C}_{2,4} + \left(1 - m_2\right)\hat{C}_{2,4}^{\text{BS}} = 1788$$

(e) The Bornhuetter Ferguson estimate can be written as $\hat{C}_i^{BF} = \hat{\beta}_{I-i}\hat{C}_{i,J} + \left(1 - \hat{\beta}_{I-i}\right)\mu_i$.

Typically, the μ_i estimate is determined using a loss ratio or other subjective method, but we can see that the Bühlmann-Straub estimate uses the same formula, but with μ_i set equal to the raw Bs estimator for AY *i*.

Question 21

(a) For a random sample of 1 person the expected number of accidents is

$E[N] = E\left[E[N \mid \theta]\right] = 0.6\times 1 + 0.3\times 2 + 0.1\times 5 = 1.7$

For 20 people the expected number is $20\times 1.7 = 34.$

(b) Working in 000's, and letting S denote the aggregate loss, we have:

$V[S] = E\left[V[S \mid \theta]\right] + V\left[E[S \mid \theta]\right]$

$V[S \mid \theta] = \theta E\left[Y^2\right] = \theta(1^2 + 2^2) = 5\theta$

$\Rightarrow E\left[V[S \mid \theta]\right] = 0.6\times 5 + 0.3\times 10 + 0.1\times 25 = 8.5$

$E[S \mid \theta] = \theta E[Y] = \theta$

$\Rightarrow V\left[E[S \mid \theta]\right] = \left(1(0.6) + 4(0.3) + 25(0.1)\right) - 1.7^2 = 1.41$

So, $V[S] = 8.5 + 1.41 = 9.91 \quad \Rightarrow SD[S] = 3.148.$

(c) Let $p(\theta)$ denote the prior distribution of θ, and let $\pi(\theta)$ denote the posterior distribution. Then $\pi(\theta) \propto L(\theta)p(\theta)$ where $L(\theta)$ is the likelihood function. Over a three year period, the distribution of the number of claims, given θ, is $\text{Poisson}(3\theta)$, so $L(\theta)$ is the probability of having 3 claims from the $\text{Poisson}(3\theta)$ distribution.

$$\pi(\theta = 1) \propto \frac{3^3 e^{-3}}{3!}(0.6) = 0.13442$$

$$\pi(\theta = 2) \propto \frac{6^3 e^{-6}}{3!}(0.3) = 0.02677$$

$$\pi(\theta = 5) \propto \frac{15^3 e^{-15}}{3!}(0.1) = 0.00002$$

And, since $\pi(\theta = 1) + \pi(\theta = 2) + \pi(\theta = 5) = 1$

we have $\pi(\theta = 1) = 0.833836; \quad \pi(\theta = 2) = 0.166057; \quad \pi(\theta = 5) = 0.000107.$

(d) $E[N] = 0.83384\times 1 + 0.16606\times 2 + 0.00011\times 5 = 1.167$

(e) $\mu(\theta) = E[N \mid \theta] = \theta$

$v(\theta) = V[N \mid \theta] = \theta$

$\mu = E[\mu(\theta)]$ and $v = E[v(\theta)],$ based on the prior distribution,

so $\mu = v = E[\theta] = 1.7$

$a = V[\mu(\theta)] = (4.3 - 1.7^2) = 1.41$

(f) $P = Z\overline{X} + (1-Z)\mu$

$\overline{X} = 3/3 = 1; \quad Z = \dfrac{n}{n + v/a} = 0.71332$

$\Rightarrow P = 1.201$

(g) With three claims in 3 years, there is still more than a 16% probability that Adam is only a "fair" driver, with double the claim frequency of a "good" driver. Underestimating this risk could be costly.

Question 22

This question is an Excel question. Please see the ASTAM Sample Excel Question file for details. The content of the Excel solution file for this question is provided below for reference.

(a)

	Cumulative Claims Paid										
AY	**DY, j**										
i	0	1	2	3	4	5	6	7	8	9	$\hat{R}_i$
0	10100	21056	26711	30594	32774	34094	35010	35153	35295	35320	0
1	10258	20337	26108	30312	32877	34053	34731	34932	35043	35067.82	24.82
2	11474	22742	29816	34581	36827	37826	38414	38569	38708.23	38735.65	166.65
3	11482	22675	29987	34063	36038	36841	37335	37507.25	37642.65	37669.31	334.31
4	11567	24416	31561	35351	37186	38383	39102.21	39282.62	39424.42	39452.35	1069.35
5	12755	25939	32694	36548	38618	39825.76	40572.00	40759.19	40906.33	40935.30	2317.30
6	13575	26229	32698	36707	39052.29	40273.63	41028.26	41217.56	41366.35	41395.65	4688.65
7	12736	24471	31704	36027.66	38329.55	39528.29	40268.95	40454.75	40600.78	40629.54	8925.54
8	12878	24721	31749.70	36079.60	38384.80	39585.27	40327.01	40513.06	40659.31	40688.11	15967.11
9	12621	25116.29	32257.39	36656.52	38998.58	40218.24	40971.84	41160.87	41309.46	41338.72	28717.72
$\hat{f}_i$	1.9900	1.2843	1.1364	1.0639	1.0313	1.0187	1.0046	1.0036	1.0007	**X**	**X**

Total Outstanding Claims = 62211.457

(b) (i) $\hat{\beta}_5 = 1/(1.0187 \times 1.0046 \times 1.0036 \times 1.0007) = 0.972895233$

(ii) This parameter is the estimated proportion of the ultimate claims that are paid by year 5; in this case, just over 97%.

(c)

$f_{i,j}$	j						
i	0	1	2	3	4	5	6
0	2.0848	1.2686	1.1454	1.0713	1.0403	1.0269	1.0041
1	1.9826	1.2838	1.1610	1.0846	1.0358	1.0199	1.0058
2	1.9820	1.3111	1.1598	1.0649	1.0271	1.0155	1.0040
3	1.9748	1.3225	1.1359	1.0580	1.0223	1.0134	
4	2.1108	1.2926	1.1201	1.0519	1.0322		
5	2.0336	1.2604	1.1179	1.0566			
6	1.9322	1.2466	1.1226				
7	1.9214	1.2956					

(d)

DY, j	0-1	1-2	2-3	3-4	4-5	5-6
Correlation	-0.0812	0.4268	0.8191	0.5444	0.9606	-0.1063

(e)

DY, j	0-1	1-2	2-3	3-4	4-5	5-6
Test Statistic	-0.1995	1.0552	2.8555	1.1241	4.8858	-0.1070
df	6	5	4	3	2	1
p-value	0.8485	0.3396	0.0461	0.3428	0.0394	0.9322

(f) The t-test for the Pearson correlation is suitable when the underlying random variables have the same variance. That is not part of the chain ladder assumptions, and indeed, is unlikely to be true.