Actuarial Society of India

Examinations
November 2006

CT6 –Statistical Methods

Indicative Solution

Q.1) As a first step, convert the given samples from the exponential distribution to samples from the uniform distribution, via the distribution function transformation.

If X has the exponential distribution with mean 5, then $e^{-X/5}$ has the uniform distribution over (0,1). Thus, the numbers $e^{-10.6101/5}$, $e^{-2.7768/5}$, $e^{-11.8926/5}$, $e^{-0.1976/5}$, $e^{-6.6885/5}$ and $e^{-6.4656/5}$ are samples from the uniform distribution.

Now use Box-Muller transformation on successive pairs to obtain standard normal samples; multiply by 2 and add 2 to get requisite mean and variance :

$$
\begin{aligned}
&2 + 2 * (-2\log(e^{-10.6101/5}))^{1/2}\cos(2 * \pi * e^{-2.7768/5}),\\
&2 + 2 * (-2\log(e^{-10.6101/5}))^{1/2}\sin(2 * \pi * e^{-2.7768/5}),\\
&2 + 2 * (-2\log(e^{-11.8926/5}))^{1/2}\cos(2 * \pi * e^{-0.1976/5}),\\
&2 + 2 * (-2\log(e^{-11.8926/5}))^{1/2}\sin(2 * \pi * e^{-0.1976/5}),\\
&2 + 2 * (-2\log(e^{-6.6885/5}))^{1/2}\cos(2 * \pi * e^{-6.4656/5}),\\
&2 + 2 * (-2\log(e^{-6.6885/5}))^{1/2}\sin(2 * \pi * e^{-6.4656/5}).
\end{aligned}
$$

These expressions lead to the samples: -1.68437, 0.15567, 6.23348, 0.94843, 1.50017, 5.23292. [5]

Q.2) (i) Expected value of the loss, $A = p_1X_1 + p_2X_2 + p_3X_3$. The premium collected by the direct insurer is $A(1+\theta)$.

Expected value of reinsurer's share of loss is $B = p_3(X_3 - X_2)$. The premium collected by the reinsurer is $B(1+\xi)$.

The overall loss matrix for the direct insurer is as under.

	X_1	X_2	X_3
d_1	0	0	0
d_2	$X_1 - A(1+\theta)$	$X_2 - A(1+\theta)$	$X_3 - A(1+\theta)$
d_3	$X_1 - A(1+\theta)$ $+B(1+\xi)$	$X_2 - A(1+\theta)$ $+B(1+\xi)$	$X_2 - A(1+\theta)$ $+B(1+\xi)$

(ii) The average overall loss for decision d_1 is 0.

The average overall loss for decision d_2 is

$$
\begin{aligned}
&p_1(X_1 - A(1+\theta)) + p_2(X_2 - A(1+\theta)) + p_3(X_3 - A(1+\theta))\\
&= \; A - A(1+\theta) \; = \; -A\theta.
\end{aligned}
$$

The average overall loss for decision d_3 is

$$
\begin{aligned}
&p_1(X_1 - A(1+\theta) + B(1+\xi)) + p_2(X_2 - A(1+\theta) + B(1+\xi))\\
&\quad +p_3(X_2 - A(1+\theta) + B(1+\xi))\\
&= \; p_1X_1 + p_2X_2 + p_3X_3 - p_3(X_3 - X_2) - A(1+\theta) + B(1+\xi)\\
&= \; A - B - A(1+\theta) + B(1+\xi) \; = \; -A\theta + B\xi.
\end{aligned}
$$

It is clear that the Bayes strategy, which minimizes the average overall loss, is d_2.

(iii) The minimum losses for decisions d_1, d_2 and d_3 are 0, $X_1 - A(1+\theta)$ and $X_1 - A(1+\theta) + B(1+\xi)$, respectively. Note that

$$\begin{aligned} X_1 - A(1+\theta) &< A - A(1+\theta) = -A\theta < 0, \\ X_1 - A(1+\theta) &< X_1 - A(1+\theta) + B(1+\xi). \end{aligned}$$

Thus, the minimum loss is minimized by strategy d_2.

(iv) For the specified values of the losses and probabilities, we have $A = 17$ and $B = 9$. The maximum losses for decisions d_1, d_2 and d_3 are 0, $X_3 - A(1+\theta) = 83 - 17\theta = 74.5$ and $X_2 - A(1+\theta) + B(1+\xi) = 2 - 17\theta + 9\xi = -1.1$. Thus, the minimax strategy is d_3.

(v) Continuing from part (iv), the maximum loss for decision d_3 is $-6.5 + 9\xi$, which is negative if and only if $\xi < 13/18$. Thus, the minimax strategy is d_3 for $1/2 < \xi < 13/18$ and d_1 for $\xi \geq 13/18$. [8]

Q.3) (i) The likelihood function is

$$L(\mu, \sigma^2) = \prod_{i=1}^{10} \frac{\exp\left[-\frac{1}{2\sigma^2}(\log x_i - \mu)^2\right]}{x_i(2\pi\sigma^2)^{1/2}}.$$

The log-likelihood is

$$\ell(\mu, \sigma^2) = -\frac{1}{2}\sum_{i=1}^{10}\left(\frac{\log x_i - \mu}{\sigma}\right)^2 - 10\log\sigma - 10\log(2\pi)^{1/2} - \sum_{i=1}^{10}\log x_i.$$

Hence,

$$\begin{aligned} \frac{\partial \ell}{\partial \mu} &= \frac{1}{\sigma}\sum_{i=1}^{10}\left(\frac{\log x_i - \mu}{\sigma}\right), \\ \frac{\partial \ell}{\partial \sigma} &= \frac{1}{\sigma}\sum_{i=1}^{10}\left(\frac{\log x_i - \mu}{\sigma}\right)^2 - \frac{10}{\sigma}. \end{aligned}$$

By equating the first expression to zero, we get

$$\hat{\mu} = \frac{1}{10}\sum_{i=1}^{10}\log x_i,$$

and by equating the second expression to zero, we get

$$\hat{\sigma}^2 = \frac{1}{10}\sum_{i=1}^{10}(\log x_i - \hat{\mu})^2 = \frac{1}{10}\sum_{i=1}^{10}(\log x_i)^2 - \hat{\mu}^2.$$

From the data, we have $\sum_{i=1}^{10}\log x_i = 61.9695$ and $\sum_{i=1}^{10}(\log x_i)^2 = 403.1326$. It follows that $\hat{\mu} = 6.197$ and $\hat{\sigma}^2 = 1.911$, i.e., $\hat{\sigma} = 1.382$.

(ii) For a Pareto distribution, we know that

$$E(X) = \frac{\lambda}{\alpha - 1}, \qquad Var(X) = \frac{\alpha\lambda^2}{(\alpha-1)^2(\alpha-2)}.$$

the other hand, the sample moments are

$$\begin{aligned} \bar{X} &= \frac{1}{10}\sum_{i=1}^{10} x_i = 1,094.1, \\ \overline{X^2} &= \frac{1}{10}\sum_{i=1}^{10} x_i^2 = 3,076,167.9. \end{aligned}$$

Thus, the sample variance is $3,076,167.9 - 1,094.1^2 = 1,879,113$.
Equating the moment expressions to the corresponding sample moments, we have (from the ratio of variance and mean-square)

$$\frac{\hat{\alpha}}{\hat{\alpha}-2} = \frac{1,879,113}{1,094.1^2}; \text{ i.e., } \hat{\alpha} = \frac{2 \times 1,879,113/1,094.1^2}{1,879,113/1,094.1^2 - 1} = 5.51013,$$

and (from the first moment equation)

$$\hat{\lambda} = 1,094.1 \times (\hat{\alpha} - 1) = 4,934.4$$

(iii) For log-normal model,

$$P(X > 3000) = 1 - \Phi\left(\frac{\log 3000 - 6.197}{1.382}\right) = 1 - \Phi(1.309) = 0.09527.$$

For Pareto,

$$P(X > 3000) = \left(\frac{4934.5}{4934.5 + 3000}\right)^{5.51013} = 0.073011. \qquad [10]$$

Q.4) (i) The surplus process is

$$U(t) = U + Ct - S(t) = 10 + 6t - S(t),$$

where $S(t)$ is the accumulated claim till time t.
Note that the function $S(t)$ has jumps (of size 2 or 10, depending on the size of claim) at integer values of t, and stays constant in between integer values of t.
Size of claim arising at the end of year n can be written as $2 + 8X_n$, where

$$X_n = \begin{cases} 0 & \text{with probability } \frac{3}{4}, \\ 1 & \text{with probability } \frac{1}{4}. \end{cases}$$

Therefore,

$$S(n) = 2n + 8\sum_{j=1}^{n} X_j.$$

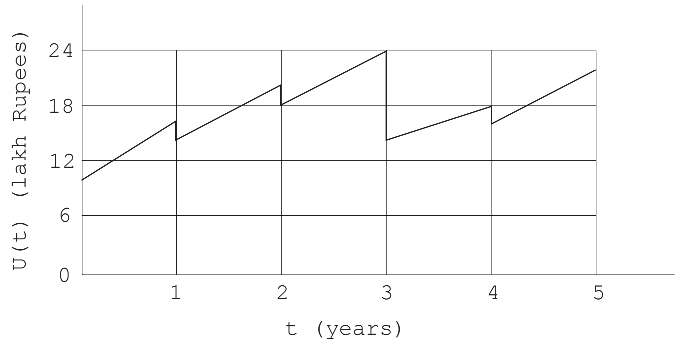
Thus,

$$U(t) = 10 + 6t - 2n - 8\sum_{j=1}^{n} X_j,$$

where n is the integer part of t (i.e., greatest integer less than or equal to t). Specifically for integer time n,

$$U(n) = 10 + 4n - 8\sum_{j=1}^{n} X_j.$$

(ii) The sketch is as under



(iii) Probability of ruin at the end of the first year is

$$P(U(1) < 0) = P(10 + 4 - 8X_1 < 0) = P(X_1 > 14/8) = 0.$$

(iv) Probability of ruin at the end of the second year is

$$\begin{aligned} P(U(2) < 0) &= P(10 + 8 - 8(X_1 + X_2) < 0) \\ &= P(X_1 + X_2 > 18/8) \\ &= 0. \end{aligned}$$

(v) At the end of the fourth year, we have $U(4) = 26 - 8(X_1 + X_2 + X_3 + X_4)$. This expression can be negative only if $X_1 = X_2 = X_3 = X_4 = 1$. However, this means that $U(3) = 22 - 8(X_1 + X_2 + X_3) < 0$, that is, ruin has already occurred at the end of the third year. Therefore, the probability that the *first* ruin occurs at the end of the fourth year is actually 0. [10]

Q.5) (i) $\alpha = e^{\mu+\sigma^2/2} = e^{\mu+1/2}$.

(ii)

$$\begin{aligned} E(\alpha) &= E\left(e^{\mu+1/2}\right) \\ &= e^{1/2}\int_{-\infty}^{\infty} e^{\mu}(2\pi(4))^{-1/2}e^{-(\mu-10)^2/(2\cdot 4)}d\mu \\ &= e^{10+1/2}\int_{-\infty}^{\infty} e^{(\mu-10)}(8\pi)^{-1/2}e^{-(\mu-10)^2/8}d\mu \end{aligned}$$

$$
\begin{aligned}
&= e^{10+1/2}\int_{-\infty}^{\infty} e^{2u}(2\pi)^{-1/2}e^{-u^2/2}du \\
&= e^{10+1/2}\int_{-\infty}^{\infty} (2\pi)^{-1/2}e^{-(u^2-4u)/2}du \\
&= e^{10+1/2+2^2/2}\int_{-\infty}^{\infty} (2\pi)^{-1/2}e^{-(u-2)^2/2}du \\
&= e^{10+1/2+2} \quad = \quad e^{12.5}.
\end{aligned}
$$

(iii) We can write the mean squared error as

$$
\begin{aligned}
E[(\hat{\alpha}-\alpha)^2] &= E[E\{(\hat{\alpha}-\alpha)^2|\alpha\}] \\
&= E[E\{(z\bar{X}+(1-z)E(\alpha)-z\alpha-(1-z)\alpha)^2|\alpha\}] \\
&= E[E\{(z(\bar{X}-\alpha)+(1-z)(E(\alpha)-\alpha))^2|\alpha\}] \\
&= E[z^2E\{(\bar{X}-\alpha)^2|\alpha\}+(1-z)^2(E(\alpha)-\alpha)^2] \\
&= z^2E[Var(\bar{X}|\alpha)]+(1-z)^2Var(\alpha)
\end{aligned}
$$

(iv) Let $A = E[Var(\bar{X}|\alpha)]$ and $B = Var(\alpha)$. The function $z^2A+(1-z)^2B$ is to be minimized with respect to z over the interval $[0,1]$. Since $A > 0$ and $B > 0$, the quadratic function has a unique minimum. Differentiating the function with respect to z and setting the derivative equal to zero, we have $2zA-2(1-z)B = 0$, which leads to the solution $(1-z)/z = A/B$, or,

$$
z = \frac{B}{B+A} = \frac{Var(\alpha)}{Var(\alpha)+E[Var(\bar{X}|\alpha)]},
$$

which is clearly between 0 and 1.

(v) Following similar steps to part (ii), we get

$$
\begin{aligned}
E(\alpha^2) &= E\left(e^{2\mu+1}\right) \\
&= e^1\int_{-\infty}^{\infty} e^{2\mu}(2\pi(4))^{-1/2}e^{-(\mu-10)^2/(2\cdot4)}d\mu \\
&= e^{20+1}\int_{-\infty}^{\infty} e^{2(\mu-10)}(8\pi)^{-1/2}e^{-(\mu-10)^2/(2\cdot4)}d\mu \\
&= e^{20+1}\int_{-\infty}^{\infty} e^{4u}(2\pi)^{-1/2}e^{-u^2/2}du \\
&= e^{20+1}\int_{-\infty}^{\infty} (2\pi)^{-1/2}e^{-(u^2-8u)/2}du \\
&= e^{20+1+8}\int_{-\infty}^{\infty} (2\pi)^{-1/2}e^{-(u-4)^2/2}du \\
&= e^{20+1+8} \quad = \quad e^{29}.
\end{aligned}
$$

Therefore,

$$
Var(\alpha) = E(\alpha^2) - [E(\alpha)]^2 = e^{29} - e^{25}.
$$

(vi) It is easy to see that $Var(\bar{X}|\alpha) = Var(X_1|\alpha)/n$. Further, $Y = \log(X_1)$ has the normal distribution with mean μ and variance 1. Therefore,

$$\begin{aligned}
E(X_1^2|\alpha) &= E(e^{2Y}|\alpha) = \int_{-\infty}^{\infty} e^{2y}(2\pi)^{-1/2}e^{-(y-\mu)^2/2}dy \\
&= \int_{-\infty}^{\infty}(2\pi)^{-1/2}e^{-(y^2-2y\mu+\mu^2-4y)/2}dy \\
&= \int_{-\infty}^{\infty}(2\pi)^{-1/2}e^{-(y^2-2y(\mu+2)+(\mu+2)^2-4\mu-4)/2}dy \\
&= e^{2\mu+2}\int_{-\infty}^{\infty}(2\pi)^{-1/2}e^{-(y^2-2y(\mu+2)+(\mu+2)^2)/2}dy \\
&= e^{2\mu+2}\int_{-\infty}^{\infty}(2\pi)^{-1/2}e^{-(y-(\mu+2))^2)/2}dy \\
&= e^{2\mu+2}.
\end{aligned}$$

Hence,

$$Var(X_1|\alpha) = E(X_1^2|\alpha) - E(X_1|\alpha)^2 = e^{2\mu+2} - e^{2\mu+1} = \alpha^2(e-1).$$

It follows from the calculations of part (v) that

$$E[Var(\bar{X}|\alpha)] = \frac{E[Var(X_1|\alpha)]}{n} = \frac{E(\alpha^2)(e-1)}{n} = \frac{e^{30}-e^{29}}{n}.$$

(vii) Substituting the results of parts (v) and (vi) in that of part (iv), we have

$$z = \frac{Var(\alpha)}{E[Var(\bar{X}|\alpha)] + Var(\alpha)} = \frac{e^{29}-e^{25}}{(e^{29}-e^{25}) + (e^{30}-e^{29})/n}.$$

Substituting this value of z and the result of part (ii) in the expression for the credibility premium $\hat{\alpha}$ given in the question, we get the following expression for $\hat{\alpha}$

$$\hat{\alpha} = z\bar{X} + (1-z)E(\alpha) = \frac{(e^{29}-e^{25})\bar{X} + e^{12.5}(e^{30}-e^{29})/n}{(e^{29}-e^{25}) + (e^{30}-e^{29})/n}. \qquad [15]$$

Q.6) (i) The transition matrix is

$$\begin{pmatrix} q & 1-q & 0 \\ q & 0 & 1-q \\ q^2 & q(1-q) & 1-q \end{pmatrix}.$$

(ii) At equilibrium, we have

$$\begin{aligned}
q(\pi_1+\pi_2) + q^2\pi_3 &= \pi_1, && (1) \\
(1-q)\pi_1 + q(1-q)\pi_3 &= \pi_2, && (2) \\
(1-q)(\pi_2+\pi_3) &= \pi_3. && (3)
\end{aligned}$$

From (3), we have $(1-q)\pi_2 = q\pi_3$, i.e., $(1-q)^2\pi_2 = q(1-q)\pi_3$.
Substituting the left hand side of the last equation in (2), we get $(1-q)\pi_1 + (1-q)^2\pi_2 = \pi_2$.

$$\begin{aligned}(\pi_1,\ \pi_2,\ \pi_3) &\propto q(1-q)(\pi_1,\ \pi_2,\ \pi_3)\\ &\propto \Big(q[1-(1-q)^2]\pi_2,\ q(1-q)\pi_2,\ (1-q)^2\pi_2\Big)\\ &\propto \Big(q^2(2-q),\ q(1-q),\ (1-q)^2\Big).\end{aligned}$$

Thus,

$$(\pi_1,\ \pi_2,\ \pi_3) = \Big(kq^2(2-q),\ kq(1-q),\ k(1-q)^2\Big),$$

for a positive number k which ensures $\pi_1+\pi_2+\pi_3 = 1$. Solving the latter equation, we have

$$1 = \Big(kq^2(2-q) + kq(1-q) + k(1-q)^2\Big) = k(1-q+2q^2-q^3).$$

It follows that $k = 1/(1-q+2q^2-q^3)$.

(iii) The expected premium for high risk policy holders is

$$350k[q^2(2-q) + 0.65q(1-q) + 0.5(1-q)^2] = \text{Rs. } 183.76.$$

Comparison of the expected premiums of the two groups show that bad risks only pay a little more than good risks. The NCD system does not discriminate sufficiently between high- and low-risk policies. [12]

Q.7) (i) CUMULATIVE NUMBER OF REPORTED CLAIMS

	Development Year				
Accident Year	*0*	*1*	*2*	*3*	*Ultimate*
2002	41	46	48	49	50
2003	45	51	53		
2004	50	56			
2005	54				

Chain ladder development factors:

$$\begin{aligned}f_{01} &= \frac{46+51+56}{41+45+50} = \frac{153}{136} = 1.125,\\ f_{12} &= \frac{48+53}{46+51} = \frac{101}{97} = 1.0412,\\ f_{23} &= \frac{49}{48} = 1.0208,\\ f_{34} &= \frac{50}{49} = 1.0204.\end{aligned}$$

CUMULATIVE NUMBER OF REPORTED CLAIMS
(Forecasts in bold)

Accident Year	Development Year 0	1	2	3	Ultimate
2002	41	46	48	49	50
2003	45	51	53	**54.10**	**55.21**
2004	50	56	**58.31**	**59.52**	**60.74**
2005	54	**60.75**	**63.26**	**64.57**	**65.89**

(ii) AVERAGE COST PER CLAIM

Accident Year	Development Year 0	1	2	3	Ultimate
2002	8.3414	9.3261	9.5416	9.6122	9.8000
2003	10.6889	13.5098	13.2264		
2004	11.6800	14.2857			
2005	12.3148				

AVERAGE COST PER CLAIM
(with grossing up factors and ultimate forecasts)

Accident Year	Development Year 0	1	2	3	Ultimate
2002	8.3414	9.3261	9.5416	9.6122	9.8000
	85.12%	95.16%	97.36%	98.08%	100.0%
2003	10.6889	13.5098	13.2264		**13.5850**
	78.68%	99.45%			
2004	11.6800	14.2857			**14.6814**
	79.56%				
2005	12.3148				**15.1810**
Average	81.12%	97.31%	97.36%	98.08%	100.0%

(iii) ULTIMATE PROJECTIONS

Accident Year	No. of Claims	Cost per Claim	Projected Loss
2002	50.00	9.8000	490.0
2003	55.21	13.5850	750.0
2004	60.74	14.6814	891.7
2005	65.89	15.1810	1000.3
Total			**3132.0**

Claims paid to date : Rs. 1821.3.
Reserve required : $3132 - 1821.3 = 1310.7$, i.e., Rs. 1,310,700. [11]

Q.8) (i) The given density is gamma with parameters $\alpha = 3$ and $\beta = 3/\mu$. Therefore, the mean is $\alpha/\beta = \mu$.

(ii) The log-density can be written as

$$\log\frac{27}{2} - 3\log\mu + 2\log y - 3\frac{y}{\mu} = \frac{y\cdot\frac{1}{\mu} - \log\frac{1}{\mu}}{-\frac{1}{3}} + \log\frac{27}{2} + 2\log y.$$

The first term is of the form $(y\theta - b(\theta))/a(\phi)$, where $\theta = 1/\mu$, $b(\theta) = \log(\theta)$ and $a(\phi) = -1/3$. Thus, this an exponential family with natural parameter $1/\mu$.

(iii) The canonical link function is the reciprocal function. Thus, the model is $1/\mu = \alpha + \beta x$. Given data (x_i, y_i), $i = 1, 2, \ldots, 20$, the log-likelihood for the parameters is

$$\sum_{i=1}^{20}\left(\log\frac{27}{2} - 3\log\mu + 2\log y_i - 3\frac{y_i}{\mu}\right)\Bigg|_{\mu=1/(\alpha+\beta x_i)}.$$

Let $\mu_0 = 1/\alpha$ and $\mu_1 = 1/(\alpha+\beta)$. Then the likelihood function simplifies to

$$\sum_{\substack{i=1\\x_i=0}}^{20}\left(\log\frac{27}{2} - 3\log\mu_0 + 2\log y_i - 3\frac{y_i}{\mu_0}\right) + \sum_{\substack{i=1\\x_i=1}}^{20}\left(\log\frac{27}{2} - 3\log\mu_1 + 2\log y_i - 3\frac{y_i}{\mu_1}\right).$$

The first sum depends only on μ_0, while the second, only on μ_1. The derivative of the likelihood with respect to μ_0 is

$$-3\frac{n_0}{\mu_0} + 3\frac{1}{\mu_0^2}\sum_{\substack{i=1\\x_i=0}}^{20} y_i,$$

where n_0 is the number of cases with $x_i = 0$. The likelihood equation leads to the maximum likelihood estimator

$$\hat{\mu}_0 = \frac{1}{n_0}\sum_{\substack{i=1\\x_i=0}}^{20} y_i.$$

The second derivative of the log-likelihood with respect to μ_0, evaluated at $\mu_0 = \hat{\mu}_0$ is

$$3\frac{n_0}{\hat{\mu}_0^2} - 6\frac{1}{\hat{\mu}_0^3}\sum_{\substack{i=1\\x_i=0}}^{20} y_i = 3\frac{n_0}{\hat{\mu}_0^2} - 6\frac{n_0}{\hat{\mu}_0^2} = -3\frac{n_0}{\hat{\mu}_0^2} < 0.$$

Thus, $\hat{\mu}_0$ indeed corresponds to the unique maximum of the likelihood function. Likewise, the MLE of μ_1 is

$$\hat{\mu}_1 = \frac{1}{n_1}\sum_{\substack{i=1\\x_i=1}}^{20} y_i,$$

where n_1 is the number of cases with $x_i = 1$. Thus, we have

$$\frac{1}{n_0}\sum_{\substack{i=1\\x_i=0}}^{20} y_i = \hat{\mu}_0 = \frac{1}{\hat{\alpha}}, \quad \frac{1}{n_1}\sum_{\substack{i=1\\x_i=1}}^{20} y_i = \hat{\mu}_1 = \frac{1}{\hat{\alpha}+\hat{\beta}}.$$

After eliminating $\hat{\alpha}$ from the two equations, we get

$$\hat{\beta} = \frac{1}{\hat{\mu}_1} - \frac{1}{\hat{\mu}_0} = n_1\left(\sum_{\substack{i=1\\x_i=1}}^{20} y_i\right)^{-1} - n_0\left(\sum_{\substack{i=1\\x_i=0}}^{20} y_i\right)^{-1}.$$ [9]

Q.9) (i) We have, the autocovariance at lags 0, 1 and 2 as under:

$$\begin{aligned}\gamma(0) &= Var(X_t) = Var(e_t + \theta e_{t-1}) = \sigma^2(1+\theta^2),\\ \gamma(1) &= Cov(X_t, X_{t-1}) = Cov(e_t + \theta e_{t-1}, e_{t-1} + \theta e_{t-2}) = \theta\sigma^2.\end{aligned}$$

Likewise, $\gamma(k)$ for $|k| > 1$ is 0, and $\gamma(-1) = \theta\sigma^2$.
The autocorrelation function is

$$\rho(k) = \gamma(k)/\gamma(0) = \begin{cases} 1 & \text{if } k = 0,\\ \frac{\theta}{1+\theta^2} & \text{if } |k| = 1,\\ 0 & \text{if } |k| > 1.\end{cases}$$

(ii) It follows from part (i) that, the ACF should be non-zero only for $k = \pm 1$. The sample ACFs should follow this pattern. From the table, it is clear that this pattern is there for the column corresponding to $m = 2$ only. Therefore, the most reasonable choice for d is 2.
By matching the sample ACF $r(1)$ of column $m = 2$ with the value $\theta/(1+\theta^2)$ obtained from part (i), we have the equation

$$\frac{\theta}{(1+\theta^2)} = -.476.$$

Solving this equation, we get $\theta = -1.372$ or $\theta = -0.729$.
For invertibility, we choose $\theta = -0.729$. [8]

Q.10) (i) $M_S(t) = M_N(\log M_X(t))$.

$$M_X(t) = \int_0^\infty e^{tx}\theta e^{-\theta x}dx = \left[-\frac{\theta}{\theta - t}e^{-(\theta-t)x}\right]_0^\infty = \frac{\theta}{\theta - t}.$$

On the other hand,

$$M_N(t) = \sum_{j=0}^{\infty}\frac{e^{nt}e^{-\lambda}\lambda^n}{n!} = e^{-\lambda}\sum_{j=0}^{\infty}\frac{(e^t\lambda)^n}{n!} = e^{\lambda(e^t-1)}.$$

It follows that

$$M_S(t) = e^{\lambda[\theta/(\theta-t)-1]} = e^{\lambda t/(\theta - t)}.$$

(ii) For $Ga(\alpha, \nu)$, the mean is α/ν the second moment is $\alpha(\alpha+1)/\nu^2$, and the variance is α/ν^2. For the prior distribution of λ, we have

$$\frac{\alpha(\alpha+1)}{\nu^2} = \frac{3}{2}, \quad \frac{\alpha}{\nu^2} = \frac{1}{2}.$$

After solving these equations, we get $\alpha = 2$, $\nu = 2$. Therefore, the prior mean is $\alpha/\nu = 1$.

By substituting $\lambda = 1$ and $\theta = 0.005$, we have from part (i) the MGF of aggregate claim as $e^{t/(0.005-t)}$.

(iii) The likelihood function for λ is

$$\prod_{i=1}^{8} \left(\frac{e^{-\lambda} \lambda^{n_i}}{n_i!} \right) \propto e^{-8\lambda} \lambda^{\sum_{i=1}^{8} n_i} = e^{-8\lambda} \lambda^5.$$

From part (ii), the prior distribution for λ is $Ga(2, 2)$. Therefore, the posterior distribution for λ is proportional to

$$e^{-8\lambda} \lambda^5 \times \lambda^{2-1} e^{-2\lambda} = \lambda^{7-1} e^{-10\lambda},$$

which is immediately recognized as $Ga(7, 10)$. The Bayes estimate of λ is the posterior mean, which is 0.7.

(iv) Solving the moment equations for the prior distribution of θ, we have

$$\frac{\alpha}{\nu} = 0.005, \quad \frac{\sqrt{\alpha}}{\nu} = 0.001, \quad \text{i.e., } \nu = 5000, \ \alpha = 25.$$

Therefore, the prior distribution of θ is $Ga(25, 5000)$.

The likelihood function for θ is

$$\prod_{i=1}^{5} (\theta e^{-\theta x_i}) = \theta^5 e^{-\theta \sum_{i=1}^{5} x_i} = \theta^5 e^{-1129.71\theta}.$$

Therefore, the posterior distribution of θ is proportional to

$$\theta^5 e^{-1129.71\theta} \times \theta^{25-1} e^{-5000\theta} = \theta^{30-1} e^{-6129.71\theta},$$

which is $Ga(30, 6129.71)$. The Bayes estimate of θ is the posterior mean, $30/6129.71 = 0.00489$.

(v) By substituting $\lambda = 0.7$ and $\theta = 0.00489$ in the result of part (i), we get the MGF of aggregate claim as $e^{0.7t/(0.00489-t)}$.

(vi) From the result of part (i), we have

$$\begin{aligned} M_S(t) &= e^{\lambda t/(\theta - t)}. \\ \text{Hence, } M_S'(t) &= \left[\frac{\lambda}{\theta - t} + \frac{\lambda t}{(\theta - t)^2} \right] M_S(t), \end{aligned}$$

$$
\begin{aligned}
M_S''(t) &= \left[\frac{\lambda}{\theta - t} + \frac{\lambda t}{(\theta - t)^2}\right]^2 M_S(t) \\
&\quad + \left[\frac{\lambda}{(\theta - t)^2} + \frac{\lambda}{(\theta - t)^2} + \frac{2\lambda t}{(\theta - t)^3}\right] M_S(t).
\end{aligned}
$$

$$
\begin{aligned}
\text{Therefore, } E(S) &= M_S'(0) = \frac{\lambda}{\theta}, \\
E(S^2) &= M_S''(0) = \frac{\lambda^2}{\theta^2} + \frac{2\lambda}{\theta^2}, \\
Var(S) &= E(S^2) - [E(S)]^2 = \frac{2\lambda}{\theta^2}.
\end{aligned}
$$

Substituting the prior means and Bayes estimates of the parameters from parts (ii) and (iv), we have $Var(S) = 80000$ and 58548, respectively. Thus, a considerable reduction in the variance of S has resulted from claim information of the last eight years. [12]