IAQS
INSTITUTE OF ACTUARIAL
& QUANTITATIVE STUDIES

Subject:

Chapter:

Category:

1.

(i) (a) Let the parameters of the Lognormal distribution be μ and σ.

Then we must solve

$$e^{\mu+\frac{\sigma^2}{2}} = 230 \qquad \text{(A)}$$

$$e^{2\mu+\sigma^2}(e^{\sigma^2}-1) = 110^2 \qquad \text{(B)}$$

$$\text{(B)} \div \text{(A)}^2 \Rightarrow e^{\sigma^2} - 1 = \frac{110^2}{230^2}$$

so $$e^{\sigma^2} = 1 + \frac{110^2}{230^2} = 1.22873$$

so $$\sigma^2 = \log 1.22873 = 0.205984$$

so $\sigma = 0.45385$

Substituting into (A) gives

$$e^{\mu+\frac{0.205984}{2}} = 230$$

$$\mu = \log(230) - \frac{0.205984}{2}$$

$$= 5.3351$$

(b) This time we have

$$e^{\mu+0.6745\sigma} = 510 \qquad \text{(A)}$$

$$e^{\mu-0.6745\sigma} = 80 \qquad \text{(B)}$$

$\log A + \log B \Rightarrow 2\mu = \log 510 + \log 80$

so $\mu = 5.30822$

and substituting into (A)

$$5.30822 + 0.6745\sigma = \log 510$$

$$\sigma = \frac{\log 510 - 5.30822}{0.6745} = 1.37315$$

(ii) Calculating the upper and lower quartiles from the parameter in (i)(a) gives

$$UQ = e^{5.3351+0.6745\times 0.45385} = 282 \qquad \text{cf } 510$$

$$LQ = e^{5.3351-0.6745\times 0.45385} = 153 \qquad \text{cf } 80$$

This is not a good fit, suggesting the underlying claims have greater weight in the tails than a Lognormal distribution.

2.

(i) Let the parameters be c and γ as per the tables.

Then we have:

$$1-e^{-c\times 20^{\gamma}}=0.1 \text{ so } \quad e^{-c\times 20^{\gamma}}=0.9 \quad \text{and so } c\times 20^{\gamma}=-\log 0.9 \text{ (A)}$$

And similarly $c\times 95^{\gamma}=-\log 0.1$ (B)

(A) divided by (B) gives $\left(\dfrac{20}{95}\right)^{\gamma}=\dfrac{\log 0.9}{\log 0.1}=0.0457575$

So $\gamma=\dfrac{\log 0.0457575}{\log\left(\dfrac{20}{95}\right)}=1.9795337$

And substituting into (A) we have $c=-\dfrac{\log 0.9}{20^{1.9795337}}=0.000280056$

(ii) The 99.5$^{\text{th}}$ percentile loss is given by

$$1-e^{-0.00280056x^{1.9795337}}=0.995$$

So that $-0.000280056x^{1.9795337}=\log 0.005$

$$\log x=\frac{\log\left(\dfrac{\log 0.005}{-0.000280056}\right)}{1.9795337}=4.97486366$$

So $x=e^{4.97486366}=144.73$

3.

(i) $M_S(t) = E(e^{tS})$

$= E(E(e^{t(X_1+X_2+\cdots+X_N)} | N)$

$= E(M_X(t)^N)$ since the X_i are independent and identically distributed

$= E(e^{N \log M_X(t)})$

$= M_N(\log M_X(t))$

$= \exp(\lambda(\exp(\log(M_X(t)-1))))$

$= \exp(\lambda(M_X(t)-1))$

(ii) $M'_S(t) = M_S(t) \times \lambda M'_X(t)$

$E(S) = M'_S(0) = M_S(0) \times \lambda \times M'_X(0)$

$= 1 \times \lambda \times \mu$

$= \lambda\mu$

$M''_S(t) = M'_S(t) \times \lambda \times M'_X(t) + M_S(t) \times \lambda \times M''_X(t)$

$E(S^2) = M''_S(0) = M'_S(0) \times \lambda \times M'_X(0) + M_S(0) \times \lambda \times M''_X(0)$

$= \lambda\mu \times \lambda \times \mu + 1 \times \lambda \times (\sigma^2 + \mu^2)$

$= \lambda^2\mu^2 + \lambda\mu^2 + \lambda\sigma^2$

And so

$$\mathrm{Var}(S) = E(S^2) - E(S)^2$$
$$= \lambda^2\mu^2 + \lambda\mu^2 + \lambda\sigma^2 - \lambda^2\mu^2$$
$$= \lambda(\mu^2 + \sigma^2)$$

(iii) First, we must calculate the mean and variance of a single claim, say Y. Let us denote by X the underlying loss. Then

$$\begin{aligned}
E(Y) &= \int_0^{200} 0.01xe^{-0.01x}dx + \int_{200}^{\infty} (x+50)\times 0.01\times e^{-0.01x}dx \\
&= \int_0^{\infty} 0.01xe^{-0.01x}dx + 50\int_{200}^{\infty} 0.01e^{-0.01x}dx \\
&= E(X) + 50\times P(X > 200) \\
&= 100 + 50\times e^{-200\times 0.01} \\
&= 100 + 6.76676 \\
&= 106.76676
\end{aligned}$$

$$\begin{aligned}
E(Y^2) &= \int_0^{200} 0.01x^2e^{-0.01x}dx + \int_{200}^{\infty} (x+50)^2\times 0.01e^{-0.01x}dx \\
&= \int_0^{\infty} 0.01x^2e^{-0.01x}dx + \int_{200}^{\infty} xe^{-0.01x}dx + 50^2\int_{200}^{\infty} 0.01e^{-0.01x}dx \\
&= E(X^2) + \left[-100xe^{-0.01x}\right]_{200}^{\infty} + \int_{200}^{\infty} 100e^{-0.01x}d + 50^2 P(X > 200) \\
&= 100^2 + 100^2 + 20,000e^{-2} + \left[-100^2e^{-0.01x}\right]_{200}^{\infty} + 2,500e^{-2} \\
&= 20,000 + 20,000e^{-2} + 10,000e^{-2} + 2,500e^{-2} \\
&= 20,000 + 32,500e^{-2} \\
&= 24,398.39671
\end{aligned}$$

And finally, using the results from part (ii)

$$E(S) = 500E(X) = 500 \times 106.76676$$
$$= 53{,}383.38$$

and

$$\text{Var}(S) = 500E(X^2) = 500 \times 24{,}398.39671 = 12{,}199{,}198.36$$

4.

(i) (a) $E(N) = E[E(N|\mu)]$
$= E[\mu] = 2/8 = 0.25$

(b) $\text{var}(N) = E[\text{var}(N|\mu)] + \text{var}[E(N|\mu)]$
$= E[\mu] + \text{var}[\mu]$
$= 2/8 + 2/8^2 = 0.28125$

(ii) Let Y be aggregate claims from one policy.

Individual claim is gamma with $\alpha = 2$ and $\lambda = 0.001$.

$$E(Y) = E(X)E(N) = 2000 \times 0.25 = 500.$$

$$Var(Y) = E(N)Var(X) + Var(N)E(X)^2$$

$$= 0.25 \times 2000000 + \tfrac{9}{32} \times 2000^2 = 1{,}625{,}000.$$

So the mean and variance of total claims are 500,000 and 1,625,000,000 respectively.

(iii) Our approximate distribution for S is $S \sim N(500{,}000\ ,\ 1625000000)$.

$$P(S > 550000) = P\left(Z > \frac{550000 - 500000}{\sqrt{1625000000}}\right) = P(Z > 1.24035) = 0.1074\,.$$

(iv) The prob three years in a row is $0.1074^3 = 0.00124$.

The probability of this happening is very low. It is more likely that the insurance company's belief about the distribution of claims amounts is incorrect.

The normal approximation tails off quickly and so underestimates the probability of extreme events

5.

(i) Let the loss amount be X. Then

$E(X) = 0.8 \times 100 + 0.2 \times 115 = 103$

$E(X^2) = 0.8 \times (100^2 + 400) + 0.2\ (115^2 + 900) = 11{,}145$

$\text{Var}(X) = E(X^2) - E(X)^2 = 11{,}145 - 103^2 = 536$

(ii) No, the loss distribution is not Normal. To see this, note that (for example) the pdf of the combined distribution will have local maxima at both 100 and 115. [Consider the case where the variances are very small to see this]

(iii) $\Pr(X > 130) = 0.8 \times \Pr(N(100, 20^2) > 130) + 0.2 \times \Pr(N(115{,}30^2) > 130)$

$= 0.8 \times \Pr\left(N(0,1) > \dfrac{130-100}{20}\right) + 0.2 \times \Pr\left(N(0,1) > \dfrac{130-115}{30}\right) \Pr(X >$ $130) = 0.8 \times \Pr(N(100, 20^2) > 130) + 0.2 \times \Pr(N(115, 30^2) > 130)$

$= 0.8 \times \Pr(N(0,1) > 1.5) + 0.2 \times \Pr(N(0,1) > 0.5)$
$= 0.8 \times (1 - 0.93319) + 0.2 \times (1 - 0.69146)$

$= 0.115156$

(iv) The relevant proportion is given by:

$$\frac{0.2 \times (1 - 0.69146)}{0.115156} = 53.6\%$$

6.

(i) Let the individual loss amounts have distribution X. Then

$$E(X) = \int_0^{100} 0.01333xe^{-0.01333x}dx + 100 \times P(X > 100)$$

$$= \left[-xe^{-0.01333x}\right]_0^{100} + \int_0^{100} e^{-0.01333x}dx + 100\int_{100}^{\infty} 0.01333e^{-0.01333x}dx$$

$$= -100e^{-1.333} + \left[-75e^{-0.01333x}\right]_0^{100} + 100\left[-e^{-0.01333x}\right]_{100}^{\infty}$$

$$= -100e^{-1.333} - 75e^{-1.333} + 75 + 100e^{-1.333}$$

$$= 55.2302$$

Hence $E(S) = 50 \times 55.2302 = 2761.5$

$$E(X^2)=\int_0^{100} 0.01333x^2e^{-0.01333x}dx+100^2P(X>100)$$

$$=\left[-x^2e^{-0.01333x}\right]_0^{100}+\int_0^{100} 2xe^{-0.01333x}dx+100^2e^{-1.333}$$

$$=-100^2e^{-1.333}+\left[-\frac{2x}{0.01333}e^{-0.01333x}\right]_0^{100}+\int_0^{100}\frac{2}{0.01333}e^{-0.01333x}dx+100^2e^{-1.333}$$

$$=-\frac{200}{0.01333}e^{-1.333}+\left[-\frac{2}{0.01333^2}e^{-0.01333x}\right]_0^{100}$$

$$=-\frac{200}{0.01333}e^{-1.333}-\frac{2}{0.01333^2}e^{-1.333}+\frac{2}{0.01333^2}$$

$$=4330.6$$

and so

$$Var(S)=50\times 4330.6=216529=(465.33)^2$$

(ii) (a) The normal distribution is $N(2761.5, 465.33^2)$

(b) The Log-Normal distribution has parameters μ and σ with

$$E(S) = e^{\mu+\sigma^2/2}$$

$$Var(S) = e^{2\mu+\sigma^2}(e^{\sigma^2}-1) = E(S)^2 \times (e^{\sigma^2}-1)$$

So substituting gives

$$216529 = 2761.5^2 \times (e^{\sigma^2}-1)$$

$$e^{\sigma^2} = \frac{216529}{2761.5^2} + 1 = 1.028394$$

$$\sigma^2 = \log(1.028394) = 0.027998$$

$$\sigma = 0.167327$$

And now we can substitute for σ to give

$$2761.5 = e^{\mu+0.027998/2}$$

$$\mu = \log(2761.5) - 0.027998/2 = 7.90953$$

(iii) Using the Normal distribution:

$$P(N(2761.5, 465.33^2) > 3000) = P\left(N(0,1) > \frac{3000 - 2761.5}{465.33}\right)$$

$$= P(N(0,1) > 0.51) = 1 - 0.69497 = 0.30503$$

From tables.

Using the log-normal distribution,

$$P(\log N(7.90953, 0.167327^2) > 3000) = P(N(7.90953, 0.167327^2) > \log(3000))$$

$$= P(N(0,1) > \frac{\log 3000 - 7.90953}{0.167327}).$$

$$= P(N(0,1) > 0.58) = 1 - 0.71904 = 0.28096$$

7.

Let the annual number of claims be denoted by N. Then

$$P(N = k) = \int_0^\infty P(N = k|\mu)\, f(\mu)\, d\mu$$

$$= \int_0^\infty e^{-\mu} \frac{\mu^k}{k!} \lambda e^{-\lambda\mu} d\mu$$

$$= \frac{\lambda}{k!} \int_0^\infty \mu^k e^{-(1+\lambda)\mu} d\mu$$

$$= \frac{\lambda}{k!} \times \frac{\Gamma(k+1)}{(1+\lambda)^{k+1}} \int_0^\infty \frac{(1+\lambda)^{k+1}}{\Gamma(k+1)} \mu^k e^{-(1+\lambda)\mu} d\mu$$

$$= \frac{\lambda}{(1+\lambda)^{k+1}} \times 1$$

Where the final integral is 1 since the integrand is the pdf of a Gamma distribution.

So

$$P(N=k) = \frac{\lambda}{(1+\lambda)^{k+1}} = \frac{\lambda}{1+\lambda} \times \frac{1}{(1+\lambda)^k}, \text{ for k = 0, 1, 2, …}$$

Which means that N has a geometric distribution with parameter $p = \frac{\lambda}{1+\lambda}$. This is equivalent to a Type II negative binomial with k=1

8.

Using truncated moments of lognormal distribution:

$$E[Z] = \int_{900}^{1700} (C-900) f(c)dc + \int_{1700}^{\infty} (800) f(c)dc$$

$$= \int_{900}^{1700} cf(c)dc - \int_{900}^{1700} (900) f(c)dc + \int_{1700}^{\infty} (800) f(c)dc$$

$$= \int_{900}^{1700} cf(c)dc - 900P(900<C<1700) + 800\ [P(C) > 1700]$$

Using truncated moments of lognormal distribution:

E[Z] = $e^{\mu + \ {}^2}$ [Φ{(ln 1700 – μ) / σ)– σ - Φ {(ln 900 – μ) / σ)– σ]

- 900 [Φ{(ln 1700 – μ) / σ)– σ - Φ {(ln 900 – μ) / σ)– σ]

+ 900 [1- Φ{(ln 1700 – μ) / σ]

= e^8 [Φ (-1.28) – Φ (-1.599)] – 900[Φ (0.72) – Φ (0.4)] + 800[1- Φ (0.72)]
= e^8 (0.10027-0.05491) – 900(0.76424-0.65542) +800(1-0.76424)
= 135.216-97.938+188.608
= 225.886

9.

We need to find the parameters of the Gamma distribution, say α and λ
Then

$$\frac{E(X)}{Var(X)} = \frac{\alpha/_\lambda}{\alpha/_{\lambda^2}} = \lambda = \frac{50}{25} = 2$$

And hence $\alpha = E(X) * \lambda = 50 * 2 = 100$
The posterior distribution is given by:

$$f(\theta_1|x) \propto f(x|\theta_1) * f(\theta_1)$$
$$\propto \left(\prod_{j=1}^{5} e^{-\theta_1} \theta_1^{n_{1j}}\right) * \theta_1^{\alpha-1} e^{-\lambda\theta_1}$$
$$\propto e^{-(\lambda+5)\theta_1} \theta_1^{\alpha+\sum_{j=1}^{5} n_{1j}-1}$$

Which is the pdf of a gamma distribution with parameters

$$\alpha + \sum_{j=1}^{5} n_{1j} = 100 + 232 = 332$$

And $\lambda + 5 = 7$
Under quadratic loss the Bayes estimate is the mean of the posterior distribution. So we have an estimate of 332/7 = 47.43

ii)
We have

$$\bar{n}_1 = \frac{232}{5}$$
$$\bar{n}_2 = \frac{260}{5}$$
$$\bar{n}_3 = \frac{145}{5}$$

This gives $\bar{n} = \frac{46.4+52+29}{3} = 42.4667$

$$\sum_{j=1}^{5} (n_{1j} - \bar{n}_1)^2 = \sum_{j=1}^{5} n_{1j} - 2\sum_{j=1}^{5} n_{1j} * \bar{n}_1 + 5 * \bar{n}_1^2$$

= 11,434 – 2*232 * 46.4 + 5*46.4^2
= 669.2

Similarly,

$$\sum_{j=1}^{5} (n_{2j} - \bar{n}_2)^2 = \sum_{j=1}^{5} n_{2j} - 2\sum_{j=1}^{5} n_{2j} * \bar{n}_2 + 5 * \bar{n}_2^2$$

= 14028 – 2* 260 * 52.0 + 5*52^2
= 508

$$\sum_{j=1}^{5}(n_{3j}-\bar{n}_3)^2 = \sum_{j=1}^{5} n_{3j} - 2\sum_{j=1}^{5} n_{3j} * \bar{n}_3 + 5 * \bar{n}_3^2$$

$= 4399 - 2 * 145 * 29.0 + 5 * 29^2$

$= 194$

So

$E(s^2(\theta)) = \frac{1}{3} * \frac{1}{4} * (669.2 + 508 + 194) = 114.2667$

$Var(m(\theta)) = \frac{1}{2} * ((46.2-42.4667)^2 + (52-42.4667)^2 + (29-42.4667)^2) - \frac{1}{5} * 114.2667$

$= 121$

So

$Z = \frac{5}{5+\frac{114.2667}{121}} = 0.8411$

So expected claims for next year are:

Cat 1 $0.1589 \times 42.4667 + 0.8411 \times 46.4 = 45.78$

Cat 2 $0.1589 \times 42.4667 + 0.8411 \times 52 = 50.49$

Cat 3 $0.1589 \times 42.4667 + 0.8411 \times 29 = 31.14$

(iii) The main differences are that:

· The approach under (i) makes use of prior information about the distribution of θ1 whereas the approach in (ii) does not.

· The approach under (i) uses only the information from the first category to produce a posterior estimate, whereas the approach under (ii) assumes that information from the other categories can give some information about category 1.

· The approach under (i) makes precise distributional assumptions about the number of claims (i.e. that they are Poisson distributed) whereas the approach under (ii) makes no such assumptions.

(iv) The insurance policies were newly introduced 5 years ago, and it is therefore likely that the volume of policies written has increased (or at least not been constant) over time. The assumption that the number of claims has a Poisson distribution with a fixed mean is therefore unlikely to be accurate, as one would expect the mean number of claims to be proportional to the number of policies. Let P_{ij} be the number of policies in force for risk i in year j.

Then the models can be amended as follows: The approach in (i) can be taken assuming that that the mean number of claims in the Poisson distribution is $P_{ij}\theta_i$. The approach in (ii) can be generalised by using EBCT Model 2 which explicitly incorporates an adjustment for the volume of risk

10.

i)

$$L = \prod_{i=1}^{10}[0.5 \propto X 0.4^{\propto} X x_i^{0.5} X (0.4 * x_i^{0.5})^{-(\propto+1)}] \; X \; [P[X > 234]]^7$$

$$= 0.5^{10} \; X \propto^{10} \; X 0.4^{10\propto} X \prod_{i=1}^{10}[x_i^{0.5}] \; X \prod_{i=1}^{10}[0.4 + x_i^{0.5}]^{-(\propto+1)} X \frac{0.4}{0.4+2340.5}^{7\propto}$$

$$\propto \propto^{10} \; X 0.4^{10\propto} X 0.40.0255^{7\propto} X \prod_{i=1}^{10}[0.4 + x_i^{0.5}]^{-(\propto+1)}$$

⇨ $\ln L \propto 10\ln\propto + 10\propto \ln(0.4) - (\propto+1)\sum_{i=1}^{10} \ln(0.4 + x_i^{0.5}) + 7 \propto \ln(0.0255)$

$= 10\ln\propto - 0.9163 X 10\propto - 25.68 \propto - 26.47 \propto$

$= 10\ln\propto - 61.31\propto$

⇨ $\frac{d\ln L}{d\propto} = \frac{10}{\propto} - 61.31 = 0$

⇨ $\propto = 0.1631$

$$\frac{d^2\ln L}{d\propto^2} = \frac{-10}{\propto^2} < 0$$

ii) *Median = 217*

⇨ $1-\left(\frac{0.4}{0.4+217^{0.5}}\right)^{\propto} = 0.5$

⇨ $\left(\frac{0.4}{0.4+217^{0.5}}\right)^{\propto} = 0.5$

⇨ $(0.02643)^{\propto} = 0.5$

$\propto = 0.1907$

(iii) The method of percentiles using median assumes is a more robust way to estimate the parameters as it assumes that the population has equivalent median to that observed in the sample. While on the other hand, the Maximum Likelihood Estimator is a more efficient way as it makes more stricter assumptions of full density. It estimates the most likely underlying parameters of the distributions. This inherent basis of the two methods leads to difference value of '$\propto$' in the above parts.

11.

i) Sample mean = (16.4 + 17.3 + 16.7) / 3 = 16.8

Prior mean = A/B = 15/1 = 15 (formula from Tables)

Credibility factor Z = 3/(3+1) = 0.75

Credibility estimate = Z * 16.8 + (1-Z) * 15 = 16.35

(ii) The variance of the gamma distribution is mean / B. Reducing the B parameter while keeping the mean constant increases the variance, reflecting greater uncertainty.

(iii) Revised credibility factor Z = 3/3.2 = 0.94 (approx.)
Revised credibility estimate = 0.94 * 16.8 + 0.06* 15 = 16.69

12.
(i) The completed table is

Loss Table	Reinsurer's loss			Insurer's Loss		
Claims	1	2	3	1	2	3
0	0	-10	-20	-60	-50	-40
150	0	-10	17.5	90	100	72.5
200	0	40	30	140	100	110
250	0	90	42.5	190	100	147.5

(ii) If no claim– arrangement 1 gives the best result
If claims is 150– arrangement 3 gives the best result
If claim is 200 or 250 – arrangement 2 gives the best result
So the strategy for Reinsurance depends the claims and not dominated by any one.

(iii) The maximum losses for insurer are
1. 190
2. 100
3. 147.5
So lowest loss is for arrangement 2 so the arrangement 2 is the minimax solution.

13.
(i) Mean and Variances of policies

Policy Number (i)	$\bar{x}_i$	s_i^2
1	4,061	6,252
2	4,366	6,359
3	4,274	6,216
4	4,484	6,121

E[m(θ)] = Overall mean = 4296.25
E[s²(θ)] = Mean of the variances = 6237

Var[m(θ)] $= \frac{1}{3}\sum_{i=1}^{4}(\bar{x}_i - \bar{x})^2 - \frac{1}{3}\left[\frac{1}{4}\sum_{i=1}^{4}\frac{1}{2}\sum_{j=1}^{3}(x_{ij} - \bar{x}_i)\right]$

$= \frac{1}{3}\left[(4061 - 4296.25)^2 + (4366 - 4296.25)^2 + (4274 - 4296.25)^2 + (4484 - 4296.25)^2\right] - \frac{1}{3} \times 6237$

= 29,905

The estimated credibility factor is $Z = \frac{3}{3+\frac{6237}{29905}}$ = 0.935

Thus, the credibility premium for Policy Number 4 is
= 0.935 x 4,484 + 0.065 x 4296.25 = 4,471.8

(ii) EBCT Model 2 can be used with risk volumes P(i,j) for each policy number for each year. P(i,1) = 1 , P(i,2)=1/1.2 and P(i,3)=1/1.2ˆ2
(Alternate solution: Adjust year-wise figures to bring it year 3 level. Multiply 1.2ˆ2 to year 1 and 1.2 to year 2. And then apply the EBCT Model 1 to adjusted on-level figures.)
Data Required: Year wise data will be required since only totals are given.

(iii) Policy number 1 has a lowest mean. Removing it will reduce the variance of means, var[m(θ)]. The variance of policy number is similar to other policies and close to overall mean of variance.
Thus, $E(s\ 2\ (\theta)]$ will remain similar and won't change much. Hence, proportionately smaller var[m(θ)] will lead to reduction of credibility factor.

14.

i) $L(p) = constant * p^x*(1-p)^(n-x)$
$logL(p) = constant + x logp + (n-x)log(1-p)$
Taking derivative w.r.t. p
$logL'(p) = x/p - (n-x)/(1-p)$
Equating to 0
$x(1-p)-(n-x)p = 0$
$x - xp - np + xp = 0$
$p = x/n$

OR (if instead of "n", 5000 is substituted):

$L(p) = constant * p^x*(1-p)^(5000-x)$
$logL(p) = constant + x logp + (5000-x)log(1-p)$
Taking derivative w.r.t. p
$logL'(p) = x/p - (5000-x)/(1-p)$
Equating to 0
$x(1-p)-(5000-x)p = 0$
$x - xp - 5000p + xp = 0$
$p = x/5000$

ii) $f(p) = 1/(1-0)$
Let posterior distribution of p be denoted by P(p)

$P(p) \alpha L(p) * f(p)$
$P(p) \alpha p^x*(1-p)^(n-x) * 1$
$P(p) \alpha p^(x+1-1)*(1-p)^(n-x+1-1)$

Therefore, the posterior distribution is beta distribution with parameters x+1, n-x+1

OR if instead of "n", 5000 is substituted, posterior distribution would be beta distribution with parameters x+1, 5001-x+

(iii) p = 200/500 = 0.4

(iv) Under quadratic loss, the Bayesian estimator is the expectation of the posterior distribution. In this case, p = (200+1)/ (200+1+500-200+1) = 201/502 = 0.4004

(v) The two estimates are almost equal, this is because the impact of prior distribution is very limited and the Bayesian estimator is mainly determined by the actual data

vi) Posterior mean can be written in credibility form as:

p = (x+1)/(n+2)
p = x/(n+2) + 1/(n+2)
= x/n * n/(n+2) + 2/(n+2) * (1/2)
= E(X) * Z + E(p) * (1-Z)
Where E(X) = x/n and E(p) = ½ and Z = n/(n+2) = 500/502