Faculty of Actuaries **Institute of Actuaries**

REPORT OF THE BOARD OF EXAMINERS ON THE EXAMINATIONS HELD IN

April 2002

Subject 106 — Actuarial Mathematics 2

Introduction

The attached subject report has been written by the Principal Examiner with the aim of helping candidates. The questions and comments are based around Core Reading as the interpretation of the syllabus to which the examiners are working. They have however given credit for any alternative approach or interpretation which they consider to be reasonable.

K Forman
Chairman of the Board of Examiners

11 June 2002

COMMENT

Question 1
This question was generally well answered though some candidates struggled to determine the variance.

Question 2
This question was well answered though some candidates did not state the correct conclusion in part (ii) despite making the correct calculations.

Question 3
This question was generally well answered though some candidates struggled with the algebra.

Question 4
Many candidates did not consider the conditional probability in part (i) giving an answer of £128 rather than £250. Candidates were only penalised once for this error.

Question 5
Overall this question was answered well though many candidates were unable to state all the assumptions required in part (ii).

Question 6
Many candidates answered this question well though some struggled with completing the transition matrix in part (ii) but went on to give good answers for part (iii). Part (iii) used the answer to part (ii) but candidates were given full credit if they had used the correct method.

Question 7
This question had a wide variety of marks ranging from very good to very poor.

Question 8
Some candidates found this question difficult with many omitting it altogether although many candidates answered this question well. Candidates should note that the examiners will continue to set questions on this topic.

Question 9
Overall this question was answered reasonably well though many candidates struggled with part (ii)(c).

Question 10
This question had a wide variety of answers ranging from very good to very poor. Candidates had most trouble with parts (iii) and (iv) although many strong candidates scored well in these parts.

1 Let N be the number of claims in a year, so $N \sim \text{Bin}(m, 0.2)$, and let X be a typical claim.

Let S be the total amount claimed in a year, so S has a compound Binomial distribution.

The expected value of S is

$$E(S) = E(N)\,E(X) = m \times 0.2 \times 400 = 80m.$$

The variance of S is

$$\begin{aligned} V(S) &= E(N)V(X) + V(N)\,(E(X))^2 \\ &= m \times 0.2 \times 110 + m \times 0.2 \times 0.8 \times 400^2 \\ &= 25{,}622m. \end{aligned}$$

2

	25%	25%	50%		
	θ_1	θ_2	θ_3	max	Expected loss
d_1	14	12	13	**14**	13
d_2	13	15	14	15	14
d_3	11	15	5	15	**9**

(i) The minimax solution is d_1.

(ii) The Bayes criterion solution is d_3.

3 (i)

Claim sizes have a Gamma(2,1) distribution and hence have mean 2 and moment generating function

$$M(r) = (1 - r)^{-2}, \; r < 1.$$

The adjustment coefficient $R > 0$ satisfies

$$(1 - R)^{-2} - 1 - 2(1 + \theta)\,R = 0,$$

(or equivalent). Hence we have

$$1 - (1 - R)^2 - 2(1 + \theta)\,R(1 - R)^2 = 0,$$

and, using $R > 0$, this simplifies to

$$2(1 + \theta)R^2 - (3 + 4\theta)\,R + 2\theta = 0.$$

(ii)

When $\theta = 0.4$ this is

$$2.8R^2 - 4.6R + 0.8 = 0,$$

and this has solutions 0.1977 and 1.4452.

Only the first of these is in the range of definition of the moment generating function, so the adjustment coefficient is $R = 0.198$.

The probability of ultimate ruin satisfies $\psi(50) \le e^{-50R}$ by Lundberg's inequality, i.e. $\psi(50) \le 5.09 \times 10^{-5}$.

4 (i) $$E[X - 100 \mid X > 100] = \frac{\int_{100}^{\infty} (x-100) f(x)dx}{P(X > 100)}$$

$$\int_{100}^{\infty} x \frac{3\lambda^3}{(\lambda + x)^4} dx = \left[-x \left(\frac{\lambda}{\lambda + x} \right)^3 \right]_{100}^{\infty} + \int_{100}^{\infty} \frac{\lambda^3}{(\lambda + x)^3} dx$$

$$= 100 \left(\frac{400}{500} \right)^3 + \left[-\frac{\lambda^3}{2(\lambda + x)^2} \right]_{100}^{\infty}$$

$$= 100 \times \left(\frac{4}{5} \right)^3 + \frac{400^3}{2 \times 500^2}$$

$$100 \int_{100}^{\infty} f(x)dx = 100 \left(\frac{400}{500} \right)^3$$

$$\therefore E[X - 100 \mid X > 100] = \frac{\frac{400^3}{2 \times 500^2}}{\left(\frac{4}{5} \right)^3} = \frac{128}{\left(\frac{4}{5} \right)^3} = 250$$

(ii) $E[X] = \dfrac{400}{2} = 200$

The expected claim size increases due to the heavy tail of the distribution.

5 (i) *This solution has been completed using simple averages of the grossing-up factors, but use of the basic chain ladder method to project average claims amounts and numbers of reported claims would also be acceptable.*

Cumulative cost of incurred claims:

	Development year		
Accident year	*0*	*1*	*2*
1999	2,317	3,754	4,336
2000	3,287	5,079	
2001	4,816		

Project number of reported claims:

	Development year			
Accident year	*0*	*1*	*2*	*ultimate*
1999	132	197	207	207
	63.77%	95.17%	**100.00%**	
2000	183	258		271
	67.50%	**95.17%**		
2001	261			398
	65.64%			

Project average incurred cost per claim:

	Development year			
Accident year	*0*	*1*	*2*	*ultimate*
1999	17.553	19.056	20.947	20.947
	83.80%	90.97%	**100.00%**	
2000	17.962	19.686		21.640
	83.01%	**90.97%**		
2001	18.452			22.123
	83.41%			

Projected total claims:

1999: 207	×	20.947	=	4,336
2000: 271	×	21.640	=	5,864
2001: 398	×	22.123	=	8,805
Total:				19,005

Less claims paid −10,237

O/s claims reserve 8,768

(ii) Assumptions are:

For each accident year, the number of claims reported in each development year is a constant proportion of the total number of claims arising from that accident year.
For each accident year, the average claim amount in each development year is a constant proportion of the ultimate average claim amount for that accident year.
Weighted average past inflation is appropriate estimate of future inflation.

6 (i)

Discount	*If claim*	*If don't claim*	*Difference*
0%	900, 675, 495	675, 495, 360	540

So the smallest loss for which a claim will be made at the 0% level is 540.

(ii) $P(\text{Claim}) = P(\text{Claim}|\text{Accident})\,.\,P(\text{Accident})$
$= P(X > x) * 0.2$

where X is the loss, which has a lognormal distribution, and x is the minimum loss for which a claim will be made. Now we know that

$$E(X) = \exp\{\mu + \tfrac{1}{2}\sigma^2\} = 1{,}188$$
$$V(X) = \exp\{2(\mu + \tfrac{1}{2}\sigma^2)\}.(\exp\{\sigma^2\} - 1) = (495)^2$$

hence

$$(\exp\{\sigma^2\} - 1) = (495)^2 / (1{,}188)^2$$
$$\exp\{\sigma^2\} = 1.1736$$
$$\sigma^2 = 0.16$$

So

$$\sigma = 0.4 \text{ and } \mu = 7.$$

$$P(X > x) = 1 - P(X < x) = 1 - \Phi\left(\frac{\ln x - \mu}{\sigma}\right) = 1 - \Phi\left(\frac{\ln x - 7}{0.4}\right)$$

$P(\text{Claim} \mid \text{Accident})$:

$$P(X > 540) = 1 - \Phi(-1.771) = \Phi(1.771) = 0.9617$$

$$P(\text{Claim}) = 0.9617 * 0.2 = 0.1923$$

So the transition matrix is:

$$\begin{pmatrix} 0.192 & 0.808 & 0 & 0 \\ 0.147 & 0 & 0.853 & 0 \\ 0.120 & 0 & 0 & 0.880 \\ 0 & 0.197 & 0 & 0.803 \end{pmatrix}$$

(iii) The steady state distribution is therefore the solution to:

$$\begin{aligned} 0.192\pi_0 + 0.147\pi_1 + 0.120\pi_2 &= \pi_0 \\ 0.808\pi_0 + 0.197\pi_3 &= \pi_1 \\ 0.853\pi_1 &= \pi_2 \\ 0.880\pi_2 + 0.803\pi_3 &= \pi_3 \\ \pi_0 + \pi_1 + \pi_2 + \pi_3 &= 1 \end{aligned}$$

Expressing π_0, π_1 and π_3 in terms of π_2:

$$\pi_3 = \frac{0.880}{0.197}\,\pi_2 = 4.4670\,\pi_2$$

$$\pi_1 = \frac{1}{0.853}\,\pi_2 = 1.1723\,\pi_2$$

$$\pi_0 = \frac{\frac{0.147}{0.853} + 0.120}{0.808}\,\pi_2 = 0.3618\,\pi_2$$

And so:

$$0.3618\ \pi_2 + 1.1723\,\pi_2 + \pi_2 + 4.4670\,\pi_2 = 1$$

which gives

$$\pi_2 = 0.1428$$

Proportions at each discount level in stable state:

0%:	5.2%
25%:	16.7%
45%:	14.3%
60%:	63.8%

(Check: 0.052 + 0.167 + 0.143 + 0.638 = 1.000)

7 (i) Let S_i be aggregate claims from the ith policy. Then

$$E[S] = E\left[\sum S_i\right] = \sum E[S_i] \quad \text{and } V[S] = V\left[\sum S_i\right] = \sum V[S_i]$$

$$E[S_i] = E[E[S_i \mid \lambda_i]] = E[\lambda_i E[X]] = E[\lambda]E[X]$$

$$V[S_i] = V[E[S_i \mid \lambda_i]] + E[V[S_i \mid \lambda_i]]$$

$$= V[\lambda_i E[X]] = E[\lambda_i E[X^2]]$$

$$= V[\lambda](E[X])^2 + E[\lambda]E[X^2]$$

$$X \sim \Gamma\left(2, \frac{1}{100}\right)$$

$\lambda \sim \Gamma(2, 10)$

$E[X] = 200$

$E[\lambda] = 0.2$

$\therefore\ E[S] = 100 \times 0.2 \times 200 = 4{,}000$

$V[\lambda] = 0.02$

$V[X] = 20{,}000$

$\therefore E[X^2] = 20{,}000 + 200^2 = 60{,}000$

$V[S] = 100\left(V[\lambda](E[X])^2 + E[\lambda]E[X^2]\right)$

$= 100 \times 200^2 \times 0.02 + 100 \times 0.2 \times 60{,}000$

$= 1{,}280{,}000$

(ii) $P(S < U + (1+\theta)\,E[S]) = 0.95$

$$P\left(\frac{S - E[S]}{\sqrt{V[S]}} < \frac{U + (1+\theta)E[S] - E[S]}{\sqrt{V[S]}}\right) = 0.95$$

$$\frac{U + \theta E[S]}{\sqrt{V[S]}} = 1.645$$

$$\theta = \frac{1.645\sqrt{V[S]} - U}{E[S]} = \frac{1861 - 2000}{4000} = -0.035$$

The company does not need to add a premium loading since the reserve is sufficient for the first year to provide the necessary security. It could even sell at (very slightly) less than expected cost (3.5% less than expected cost).

(iii) λ fixed would reduce variability and therefore the value of θ would decrease. This would provide further scope for competitive pricing in the first year.

8 (i) Writing y for $(y_1, \ldots, y_m)$, the likelihood is

$$f(y \mid \theta_1, \ldots, \theta_m) = \prod_{i=1}^{m} \binom{n_i}{y_i} \theta_i^{y_i} (1-\theta_i)^{n_i - y_i},$$

so the log-likelihood is

$$l(\theta_1, \ldots, \theta_m) = \sum_{i=1}^{m} y_i \log\theta_i + \sum_{i=1}^{m} (n_i - y_i)\log(1-\theta_i) + \sum_{i=1}^{m} \log\binom{n_i}{y_i}.$$

Differentiating with respect to θ_i and setting to zero gives

$$\frac{y_i}{\theta_i} - \frac{n_i - y_i}{1-\theta_i} = 0,$$

and so

$$y_i(1-\theta_i) = (n_i - y_i)\,\theta_i,$$

which gives $\hat{\theta}_i = y_i/n_i$.

(ii) (a) First note that $1 - \theta_i = (1 + e^{\alpha + \beta x_i})^{-1}$. Using the expression for the log-likelihood derived in (i) above, the log-likelihood can be written

$$l = \sum_{i=1}^{m} y_i \log\theta_i + \sum_{i=1}^{m}(n_i - y_i)\log(1-\theta_i) + c$$

$$= \sum_{i=1}^{m} y_i \log\left(\frac{\theta_i}{1-\theta_i}\right) - \sum_{i=1}^{m} n_i \log(1 + e^{\alpha+\beta x_i}) + c$$

$$= \sum_{i=1}^{m} y_i(\alpha + \beta x_i) - \sum_{i=1}^{m} n_i \log(1 + e^{\alpha+\beta x_i}) + c\,.$$

$$= \alpha\sum_{i=1}^{m} y_i + \beta\sum_{i=1}^{m} x_i y_i - \sum_{i=1}^{m} n_i \log(1+\exp(\alpha+\beta x_i)) + c$$

(b) Differentiating with respect to α and β in turn, and setting to zero gives the equations satisfied by $\hat{\alpha}$ and $\hat{\beta}$, i.e.

$$\sum_{i=1}^{m} y_i - \sum_{i=1}^{m} n_i \frac{e^{\hat{\alpha}+\hat{\beta}x_i}}{1+e^{\hat{\alpha}+\hat{\beta}x_i}} = 0$$

$$\sum_{i=1}^{m} x_i y_i - \sum_{i=1}^{m} n_i \frac{x_i e^{\hat{\alpha}+\hat{\beta}x_i}}{1+e^{\hat{\alpha}+\hat{\beta}x_i}} = 0\,.$$

(iii) The required deviance D is twice the difference between the maximised log-likelihoods for the models in (i) and (ii).

Substituting the maximum likelihood estimators into the log-likelihood in (i), and e_i/n_i for θ_i in the model in (ii), we obtain

$$D = 2\left(\sum_{i=1}^{m} y_i \log\frac{y_i}{n_i} + \sum_{i=1}^{m}(n_i - y_i)\log\left(1-\frac{y_i}{n_i}\right) - \sum_{i=1}^{m} y_i \log\frac{e_i}{n_i} - \sum_{i=1}^{m}(n_i - y_i)\log\left(1-\frac{e_i}{n_i}\right)\right)$$

$$= 2\left(\sum_{i=1}^{m} y_i \log\frac{y_i}{e_i} + \sum_{i=1}^{m}(n_i - y_i)\log\frac{n_i - y_i}{n_i - e_i}\right).$$

(iv) (a) Model 2 has one extra parameter fitted, so the degrees of freedom is 4.

(b) The first model does not fit all that well (deviance = 13.33 compared to χ_5^2 distribution), but the second model is better (deviance 1.67 compared to a χ_4^2 distribution).

The drop in deviance resulting from including regression on age is $13.33 - 1.67 = 11.66$, which is significant when compared to a χ^2 distribution on $5 - 4 = 1$ degree of freedom, implying that the age term should be included in the model.

(c) In Model 2, $\hat{\beta}/\text{s.e.}(\hat{\beta}) = -3.45$ and is significant confirming that β should be in the model. The estimate of β is negative, so logit(θ_i) decreases as age increases.

Comparing odds,

$$\frac{\theta_{i+1}}{1-\theta_{i+1}} = e^{\alpha+\beta x_{i+1}}$$

$$= e^{\alpha+\beta(x_i+1)}$$

$$= \frac{\theta_i}{1-\theta_i}e^{\beta}.$$

Substituting the estimate of β, the odds $\theta_i/(1-\theta_i)$ are multiplied by $e^{\hat{\beta}} = e^{-0.2492} = 0.78$ to get the odds for age $x_{i+1.}$

9 (i) (a) The prior density for λ is $f(\lambda) = e^{-\lambda/\mu}/\mu$, and, writing x for $(x_1, \ldots, x_n)$,

$$f(x\,|\,\lambda) = \frac{e^{-n\lambda}\lambda^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i\,!}$$

so that the posterior density is

$$f(\lambda\,|\,x) \propto f(\lambda)\times f(x\,|\,\lambda)$$

$$\propto e^{-\lambda/\mu}e^{-n\lambda}\lambda^{\sum_{i=1}^{n}x_i}$$

$$= \lambda^{\sum_{i=1}^{n}x_i}\exp\left(-\lambda\left(n+\frac{1}{\mu}\right)\right),$$

i.e. gamma$(1 + \sum_{i=1}^{n} x_i, n + (1/\mu))$.

(b) The Bayesian estimate under quadratic loss is the posterior mean,

$$\frac{\sum_{i=1}^{n} x_i + 1}{n + \frac{1}{\mu}} = \frac{n}{n + \frac{1}{\mu}}\,\overline{x} + \frac{\frac{1}{\mu}}{n + \frac{1}{\mu}}\,\mu,$$

which is in the form of a credibility estimate with credibility factor $n/(n + (1/\mu))$.

(c) With the given values, the estimate of λ is

$$\frac{550 + 1}{10 + (1/50)} = 54.99.$$

(ii) (a) First note that the Pareto(α, ν) distribution has $\nu/(\alpha - 1) = 2$ and $\alpha\nu^2/((\alpha - 1)^2(\alpha - 2)) = 12$. Solving these gives $\alpha = 3$ and $\nu = 4$.

Let X be a randomly chosen claim, so X has density

$$f(x) = 0.6e^{-x} + 0.4\ \frac{3 \times 4^3}{(4 + x)^4}.$$

Then,

$$P(X > M) = 0.6e^{-M} + 0.4\left(\frac{4}{4 + M}\right)^3.$$

(b) Let T be the total amount of time spent answering the messages from one day.

$$E(T) = \lambda(0.6 \times 1 + 0.4 \times 2) = 54.99 \times 1.4 = 76.99.$$

(c) Let Y be the time to answer an e-mail message under the new strategy, so $Y = \min(X, 1.5)$ and

$$E(Y) = E(\min(X, 1.5)) = \int_0^{1.5} xf(x)dx + 1.5P(X > 1.5)$$

$$= E(X) - \int_0^{\infty} xf(x + 1.5)dx.$$

The reduction in expectation per message is

$$\int_0^\infty xf(x+1.5)dx = \int_0^\infty x\left(0.6e^{-x-1.5} + 0.4\times\frac{3\times 4^3}{(4+1.5+x)^4}\right)dx$$

$$= 0.6e^{-1.5}\int_0^\infty xe^{-x}dx + \frac{0.4\times 4^3}{5.5^3}\int_0^\infty x\frac{3\times 5.5^3}{(5.5+x)^4}dx$$

$$= 0.6e^{-1.5} + \frac{0.4\times 4^3}{5.5^3}\times\frac{5.5}{2}$$

$$= 0.5570$$

Hence the total reduction for one day's messages is
$54.99 \times 0.5570 = 30.63$ minutes.

(Or any correct calculation method.)

10 (i) The likelihood function is:

$$L(\mu,\sigma^2) = \frac{\prod_{i=1}^{10}\exp\left\{-\frac{1}{2}\left(\frac{\ln x_i - \mu}{\sigma}\right)^2\right\}}{x_i\sqrt{2\pi\sigma^2}}$$

and l = log likelihood is:

$$l(\mu,\sigma^2) = -\frac{1}{2}\sum_{i=1}^{10}\left(\frac{\ln x_i - \mu}{\sigma}\right)^2 - 10\ln\sigma - 10\ln\sqrt{2\pi} - \sum_{i=1}^{10}\ln x_i$$

So:

$$\frac{\partial l}{\partial\mu} = \frac{1}{\sigma}\sum_{i=1}^{10}\left(\frac{\ln x_i - \mu}{\sigma}\right)$$

and

$$\frac{\partial l}{\partial\sigma} = \frac{1}{\sigma}\sum_{i=1}^{10}\left(\frac{\ln x_i - \mu}{\sigma}\right)^2 - \frac{10}{\sigma}$$

Equating both to zero gives:

$$\sum_{i=1}^{10}\left(\frac{\ln x_i - \hat{\mu}}{\hat{\sigma}}\right) = 0 \text{ which gives } \hat{\mu} = \frac{1}{10}\sum_{i=1}^{10}\ln x_i$$

and

$$\sum_{i=1}^{10}\left(\frac{\ln x_i - \hat{\mu}}{\hat{\sigma}}\right)^2 = 10$$

$$\hat{\sigma}^2 = \sum_{i=1}^{10}\left(\frac{\ln x_i^2 - 2\hat{\mu}\ln x_i + \hat{\mu}^2}{10}\right)$$

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{10}\ln x_i^2 - 10\hat{\mu}^2}{10} = \frac{\sum_{i=1}^{10}\ln x_i^2}{10} - \hat{\mu}^2$$

Now from the data we have:

$$\sum_{i=1}^{10}\ln x_i = 61.9695 \text{ and } \sum_{i=1}^{10}\ln x_i^2 = 403.1326 \text{ so}$$

$$\hat{\mu} = 6.197 \text{ and } \hat{\sigma}^2 = 1.911 \quad \text{i.e. } \hat{\sigma} = 1.382$$

(ii) Under a Pareto distribution we have:

$$E(X) = \frac{\lambda}{(\alpha - 1)} \text{ and } V(X) = \frac{\alpha\lambda^2}{(\alpha-1)^2(\alpha-2)}$$

Equating to sample moments, which are:

$$\text{Mean} = \frac{1}{10}\sum_{i=1}^{10}x_i = 1{,}094.1$$

and

$$\text{Variance} = \frac{\sum x_i^2}{10} - 1{,}094.1^2 = \frac{30{,}761{,}679}{10} - 1{,}094.1^2 = 1{,}879{,}113$$

gives:

$$\frac{1{,}879{,}113}{1{,}094.1^2} = \frac{\hat{\alpha}}{\hat{\alpha} - 2}$$

$$\hat{\alpha} = \frac{2\left(\dfrac{1{,}879{,}113}{1{,}094.1^2}\right)}{\dfrac{1{,}879{,}113}{1{,}094.1^2} - 1} = 5.51013$$

and

$$\hat{\lambda} = 1{,}094.1 \times (\hat{\alpha} - 1) = 4{,}934.5$$

[*Some candidates divided by 9 rather than 10 in the variance and were not penalised. This gives* $\hat{\alpha} = 4.68745$ *and* $\hat{\lambda} = 4{,}034.44$. *Following through to part (iv), the answer for the Pareto then becomes 0.07383.]*

(iii) 25^{th} percentile is ½ × (111 + 201) = 156 and
75^{th} percentile is ½ × (843 + 1,330) = 1,086.5

Then we need to solve:

$$1 - \exp(-c \times 156^{\gamma}) = 0.25 \text{ and } 1 - \exp(-c \times 1{,}086.5^{\gamma}) = 0.75$$

Simplifying and taking logs gives:

$$-\hat{c}156^{\hat{\gamma}} = \ln(0.75) \text{ and } -\hat{c}1{,}086.5^{\hat{\gamma}} = \ln(0.25)$$

Dividing out:

$$\left(\frac{1{,}086.5}{156}\right)^{\hat{\gamma}} = \frac{\ln 0.25}{\ln 0.75}$$

Taking logs:

$$\hat{\gamma}\ln\left(\frac{1{,}086.5}{156}\right) = \ln\left(\frac{\ln 0.25}{\ln 0.75}\right)$$

So $\hat{\gamma} = 0.81022$ and $\hat{c} = 0.00481$

(iv) $P(X > 3{,}000) = 1 - F(3{,}000)$:

Lognormal: $1 - \Phi\left(\dfrac{\ln 3{,}000 - 6.197}{1.382}\right) = 1 - \Phi(1.309) = 0.09527$

Pareto: $\left(\dfrac{4,934.5}{4,934.5+3,000}\right)^{5.51013} = 0.073011$

Weibull: $\exp\{-0.00481 \times 3,000^{0.81022}\} = 0.047542$