EXAMINATION

September 2005

Subject CT6 — Statistical Methods
Core Technical

EXAMINERS' REPORT

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.

M Flaherty
Chairman of the Board of Examiners

15 November 2005

© Faculty of Actuaries
© Institute of Actuaries

EXAMINERS' COMMENTS

Comments on solutions presented to individual questions for this September 2005 paper are given below.

Q1 The general standard of answers was poor. Very few candidates correctly derived the condition for ruin to occur in terms of time until the first claim.

Q2 Reasonably well answered.

Q3 Well answered by the majority of candidates.

Q4 Very few candidates generated enough different points to score high marks on this question.

Q5 Most candidates were able to answer part (i). Answers to part (ii) were more varied, with many candidates struggling with the necessary integration, or evaluating the wrong integral.

Q6 Reasonably well answered, but a large number of candidates failed to cope with the possible additional fixed claim amounts. A common error was to treat the separate elements as entirely independent – which is incorrect since the fixed claim can only occur when a variable claim has already been made.

Q7 Answers were generally of a high standard. However, a surprisingly large number of candidates ignored the instruction in part (iv) of the question and instead derived the proportions in each state in a stationary population.

Q8 In general this was reasonably well answered, although many candidates failed to give any justification for their classification in part (i). In part (iv) a number of candidates claimed that the autocorrelation function was zero for k greater than 2.

Q9 The inflation adjustment needed to past claims data produced a number of common alternative approaches – credit was given for these providing they were sensible, consistent and clearly explained. Most candidates were able to derive the necessary development factors. However, a large number of candidates failed to adjust the outstanding claims for future inflation (ie the answer was given in 2004 prices) and so failed to obtain full marks. Many candidates did not show any working for the calculations in part (ii) making it difficult to give any credit for partially correct solutions.

Q10 Many candidates simply wrote down the result in part (i) (which is given in the tables) rather than deriving it as instructed. The rest of the question was very well answered in general.

1 Ruin will occur if the time of the first claim t is such that

$$U + 1.2\lambda dt < d$$

i.e. if

$$t < \frac{d-U}{1.2\lambda d}.$$

The time until the first claim follows an exponential distribution with parameter λ. So the probability of ruin is given by

$$\int_0^{\frac{d-U}{1.2\lambda d}} \lambda e^{-\lambda x} dx = \left[-e^{-\lambda x}\right]_0^{\frac{d-U}{1.2\lambda d}}$$

$$= 1 - e^{-\lambda\left(\frac{d-U}{1.2\lambda d}\right)}$$

$$= 1 - e^{-\frac{1}{1.2}\left(1-\frac{U}{d}\right)}.$$

2 The deviance is $2(l_f - l_c)$, where l_f and l_c are the log-likelihoods of the full and current model, respectively.

$$f(y) = \frac{\mu^y e^{-\mu}}{y!}$$

$$l = \sum_{i=1}^{n}\left[y_i \log \mu_i - \mu_i - \log y_i!\right]$$

$$l_f = \sum_{i=1}^{n}\left[y_i \log y_i - y_i - \log y_i!\right]$$

$$l_c = \sum_{i=1}^{n}\left[y_i \log \hat{\mu}_i - \hat{\mu}_i - \log y_i!\right]$$

Hence the deviance is

$$2\left[\sum_{i=1}^{n} y_i \log \frac{y_i}{\hat{\mu}_i} - \sum_{i=1}^{n}(y_i - \hat{\mu}_i)\right]$$

3 (i) **Revenue (£)**

	Low	*Medium*	*High*
Basic	1,470,000	2,100,000	2,415,000
Deluxe	1,764,000	2,520,000	2,898,000
Supreme	2,205,000	3,150,000	3,622,500

Costs (£)

	Low	*Medium*	*High*
Basic	1,400,000	1,400,000	1,400,000
Deluxe	1,700,000	1,700,000	1,700,000
Supreme	2,300,000	2,300,000	2,300,000

Profit (£)

	Low	*Medium*	*High*	*Min*	*Max*	*Expected Profit*
Basic	70,000	700,000	1,015,000	70,000	1,015,000	589,750
Deluxe	64,000	820,000	1,198,000	64,000	1,198,000	687,700
Supreme	–95,000	850,000	1,322,500	–95,000	1,322,500	684,625

(ii) Minimax: Decision is "Basic".

(iii) Highest expected profit. Decision is "Deluxe".

4 (i) Higher risk of default.
Cost of MIG insurance.
Other insurers.
Cost of underwriting.
Profitability of normal product.
Adjust for underlying economic conditions.
Exceptional defaults.

(ii) Higher SVR.
Lower LTV.
Higher MIG.
Penalty Payments.
Increase charges.
Compulsory insurance.
Maximum loan amount.
Charges/SVR variable according to wish.

Other sensible suggestions were given credit.

5 (i) We must find D such that

$$\int_0^D f(x)dx = 0.25$$

this means that

$$0.25 = \int_0^D \frac{\alpha\lambda^\alpha}{(\lambda+x)^{\alpha+1}}dx$$

$$= \left[-\frac{\lambda^\alpha}{(\lambda+x)^\alpha}\right]_0^D$$

$$= 1-\frac{\lambda^\alpha}{(\lambda+D)^\alpha}$$

$$= 1-\left(\frac{300}{300+D}\right)^{2.5}$$

So

$$\frac{300}{300+D} = (1-0.25)^{\frac{1}{2.5}} = 0.8913$$

and hence

$$300 + D = \frac{300}{0.8913}$$

so that

$$D = \frac{300}{0.8913}-300 = 36.59.$$

(ii) The average net claim is given by $E[X - D \mid X > D]$

$$\int_D^\infty (x-D)f(x)dx = \int_D^\infty \frac{x\alpha\lambda^\alpha}{(\lambda+x)^{\alpha+1}}dx - D\int_D^\infty \frac{\alpha\lambda^\alpha}{(\lambda+x)^{\alpha+1}}dx$$

$$= \left[-\frac{x\lambda^\alpha}{(\lambda+x)^\alpha}\right]_D^\infty + \int_D^\infty \frac{\lambda^\alpha}{(\lambda+x)^\alpha}dx - D\left[-\frac{\lambda^\alpha}{(\lambda+x)^\alpha}\right]_D^\infty$$

$$= 0 + \frac{D\lambda^{\alpha}}{(\lambda+D)^{\alpha}} + \left[-\frac{\lambda^{\alpha}}{(\alpha-1)(\lambda+x)^{\alpha-1}} \right]_D^{\infty} + 0 - \frac{D\lambda^{\alpha}}{(\lambda+D)^{\alpha}}$$

$$= \frac{\lambda^{\alpha}}{(\alpha-1)(\lambda+D)^{\alpha-1}}$$

$$= \frac{300^{2.5}}{1.5 \times (300+36.59)^{1.5}}$$

$= 168.29.$

Hence $E[X - D \mid X > D] = \dfrac{168.29}{0.75} = 224.39$

6 We first calculate $E(100X)$ and $\mathrm{Var}(100X)$. Taking the expectation first, we have

$$\begin{aligned}
E(100X) &= 100E(X) \\
&= 100 \times \frac{3}{32} \times \int_1^5 6x^2 - x^3 - 5x dx \\
&= \frac{300}{32}\left[2x^3 - \frac{x^4}{4} - \frac{5x^2}{2}\right]_1^5 \\
&= \frac{300}{32} \times (31.25 + 0.75) \\
&= 300.
\end{aligned}$$

Now for the variance

$$\begin{aligned}
E((100X)^2) &= 100^2 \times E(X^2) \\
&= 100^2 \times \frac{3}{32} \times \int_1^5 6x^3 - x^4 - 5x^2 dx \\
&= \frac{30,000}{32}\left[\frac{6x^4}{4} - \frac{x^5}{5} - \frac{5x^3}{3}\right]_1^5
\end{aligned}$$

$$= \frac{30,000}{32} \times (104.166 + 0.366)$$

$$= 98,000$$

so that

$$\begin{aligned} \mathrm{Var}X &= E(X^2) - E(x)^2 \\ &= 98,000 - 300^2 \\ &= 8000 \\ &= (89.44)^2. \end{aligned}$$

Now let Y denote the cost (if any) of the associated repair. Then Y is independent of X and

$$E(Y) = 0.3 \times 200 = 60$$

and

$$\mathrm{Var}(Y) = 0.3 \times 200^2 - 60^2 = 8,400 = (91.65)^2.$$

So the mean individual claim amount Z is

$$E(X + Y) = E(X) + E(Y) = 300 + 60 = 360$$

and the variance of an individual claim is

$$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) = 7,998.75 + 8,400.00 = 16,398.75$$

The mean and variance of the aggregate claims S are given by the formulae

$$E(S) = E(N)E(Z)$$

and

$$\mathrm{Var}(S) = E(N)\ \mathrm{Var}(Z) + \mathrm{Var}(N)E(Z)^2$$

where N is the total number of claims. In our case

$$E(S) = 50 \times 360 = 18,000$$

and

$$\mathrm{Var}(S) = 50 \times 16,398.75 + 50 \times 360^2 = 7,299,937.5 = (2,701.84)^2$$

7 (i)

Discount level	*Premiums if no claim*	*Premiums if claim*	*Difference*
0%	480, 360, 300, 240	600, 480, 360, 300	360
20%	360, 300, 240, 240	600, 480, 360, 300	600
40%	300, 240, 240, 240	600, 480, 360, 300	720
50%/60%	240, 240, 240, 240	480, 360, 300, 240	420

(ii)

0% $P(\text{Cost} > 360) = \exp(-360/1{,}750) = 0.814$
20% $P(\text{Cost} > 600) = \exp(-600/1{,}750) = 0.710$
40% $P(\text{Cost} > 720) = \exp(-720/1{,}750) = 0.663$
50%/60% $P(\text{Cost} > 420) = \exp(-420/1{,}750) = 0.787$

(iii) The amount for which the 50%/60% discount drivers will claim is much lower than the 20% and 40% categories. It seems illogical for this to be the case as this leads to a higher probability of a claim in the event of an accident. Suggest that structure is altered to make claims for these categories less likely.

(iv) The transition matrix is

$$\begin{bmatrix} 0.0814 & 0.9186 & 0 & 0 & 0 \\ 0.0710 & 0 & 0.9290 & 0 & 0 \\ 0.0663 & 0 & 0 & 0.9337 & 0 \\ 0 & 0.0787 & 0 & 0 & 0.9213 \\ 0 & 0.0787 & 0 & 0 & 0.9213 \end{bmatrix}$$

This year the proportions at each level are

(0.2, 0.2, 0.2, 0.2, 0.2)

Next year, the expected proportions are

$0.2 \times (0.0814 + 0.0710 + 0.0663) = 0.04374$
$0.2 \times (0.9186 + 0.0787 + 0.0787) = 0.2152$
$0.2 \times 0.9290 = 0.1858$
$0.2 \times 0.9337 = 0.18674$
$0.2 \times (0.9213 + 0.9213) = 0.36852$

8 (i) $(1 - 0.4B - 0.2B^2)\, Y_t = Z_t + 0.025$

Characteristic equation

$$1 - 0.4z - 0.2z^2 = 0$$

has no root at $z = 1$, so $d = 0$.

No functional dependence on Z_{t-1}, Z_{t-2}, etc., so $q = 0$.

Hence this is an ARIMA(2, 0, 0).

(ii) Roots of characteristic equation are $-1 \pm \sqrt{6}$, which are outside (−1, +1), so $\{Y_t\}$ is stationary.

(iii) Mean is stationary over time

$$(1 - 0.4 - 0.2)E[Y_t] = 0.025$$

$$\therefore\ E[Y_t] = \frac{0.025}{0.4} = 0.0625.$$

(iv) $Y_t - 0.0625 = 0.4(Y_{t-1} - 0.0625) + 0.2(Y_{t-2} - 0.0625) + Z_t$

$$\rho_k = E[(Y_t - 0.0625)(Y_{t-k} - 0.0625)] = 0.4\rho_{k-1} + 0.2\rho_{k-2}$$

Put $k = 1$, and note that $\rho_0 = 1$ and $\rho_{-1} = \rho_1$

$$\therefore\ \rho_1 = 0.4 + 0.2\rho_1 \quad \therefore\ \rho_1 = \frac{0.4}{0.8} = 0.5$$

$$\rho_2 = 0.4\rho_1 + 0.2 = 0.4$$

$$\rho_3 = 0.4\rho_2 + 0.2\rho_1 = 0.26$$

and so on.

(v) $(1 - 0.4B - 0.2B^2)(Y_t - 0.0625) = z_t$

$$Y_t - 0.0625 = (1 - 0.4B - 0.2B^2)^{-1} Z_t$$

So we need to invert $(1 - 0.4B - 0.2B^2)$
and multiply by Z_t to obtain the equivalent moving average process.

9 (Figures in £000s)

Inflation factors for each development year

AY	*0*	*1*	*2*	*3*
2001	1.02499	1.00390	0.99200	1.00000
2002	1.00390	0.99200	1.00000	
2003	0.99200	1.00000		
2004	1.00000			

Other inflation adjustments were given credit providing they were sensible, consistent and some explanation was given.

Inflation adjusted claim payments in mid 2004 prices

AY	*0*	*1*	*2*	*3*
2001	1,287.39	948.69	625.95	378
2002	1,444.61	1,063.42	723	
2003	1,530.66	1,133		
2004	1,480			

Inflation adjusted cumulative claim payments in mid 2004 prices

AY	*0*	*1*	*2*	*3*
2001	1,287.39	2,236.08	2,862.03	3,240.03
2002	1,444.61	2,508.03	3,231.03	
2003	1,530.66	2,663.66		
2004	1,480			

Column sum		7,407.77	6,093.06	3,240.03
Column sum minus last entry	4,262.66	4,744.11	2,862.03	
Development factor	1.73783	1.28434	1.13207	

(i)

Outstanding amounts arising from 2004 policies

Accumulated	1,480.0	2,572.0	3,303.3	3,739.6
Disaccumulated		1,092.0	731.3	436.3
Inflation		1.025	1.050625	1.076891
Inflation adj by year		1,119.3	768.3	469.8
Total		2,357.4		

For example, 2,572.0 = 1,480 × 1.7378

(ii) Ultimate amount for 2004 policies 5,250 × 0.75 = 3,937.50

Outstanding amounts for 2004 policies

	1,149.8	770.0	459.4
Infl adj by year	1.025	1.050625	1.076781
Inflation adj by year	1,178.5	809.0	494.7
Total	2,482.2		

For example,

1,149.8 = 3,937.5 × (1 / (1.28434 * 1.13207) – 1 / (1.73783 * 1.28434 * 1.13207))

10 (i) $x_1, \ldots, x_n$ are the observed claims:

$$f(\theta) \propto e^{-\frac{(\theta-\mu)^2}{2\sigma_2^2}} \propto e^{-\frac{1}{2\sigma_2^2}(\theta^2-2\theta\mu)}$$

$$p(\underline{x}|\theta) \propto \prod_{i=1}^{x} e^{-\frac{(x_1-\theta)^2}{2\sigma_1^2}}$$

$$= e^{-\frac{1}{2\sigma_1^2}\sum_{i=1}^{n}(x_i-\theta)^2}$$

$$\propto e^{-\frac{1}{2\sigma_1^2}\sum_{i=1}^{n}(\theta^2-2\theta x_i)}$$

$$= e^{-\frac{1}{2\sigma_1^2}\left(n\theta^2-2\theta\sum_{i=1}^{n}x_i\right)}$$

$$p(\theta|\underline{x}) \propto p(\underline{x}|\theta)\, p(\theta)$$

$$\propto e^{-\frac{1}{2\sigma_2^2}(\theta^2-2\theta\mu)-\frac{1}{2\sigma_1^2}\left(n\theta^2-2\theta\sum_{i=1}^{n}x_i\right)}$$

$$= e^{-\left\{\left(\frac{1}{2\sigma_2^2}+\frac{n}{2\sigma_1^2}\right)\theta^2-\left(\frac{\mu}{2\sigma_2^2}+\frac{\sum_{i=1}^{n}x_i}{2\sigma_1^2}\right)2\theta\right\}}$$

$$= e^{-\left(\frac{\sigma_1^2+n\sigma_2^2}{2\sigma_1^2\sigma_2^2}\right)\theta^2-2\theta\left(\frac{\mu\sigma_1^2+\left(\sum_{i=1}^{n}x_i\right)\sigma_2^2}{2\sigma_1^2\sigma_2^2}\right)}$$

$$\propto e^{-\frac{\sigma_1^2+n\sigma_2^2}{2\sigma_1^2\sigma_2^2}\left(\theta-\left(\frac{\mu\sigma_1^2}{\sigma_1^2+n\sigma_2^2}+\frac{\sigma_2^2\sum_{i=1}^{n}x_i}{\sigma_1^2+n\sigma_2^2}\right)\right)^2}$$

$$\Rightarrow \theta|\underline{x} \sim N\left(\frac{\mu\sigma_1^2+\left(\sum_{i=1}^{n}x_i\right)\sigma_2^2}{\sigma_1^2+n\sigma_2^2}, \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+n\sigma_2^2}\right)$$

(ii) Point estimator under quadratic loss is the posterior mean:

$$E(\theta|\underline{x}) = \frac{\sigma_1^2}{\sigma_1^2+n\sigma_2^2}\mu+\frac{\sigma_2^2}{\sigma_1^2+n\sigma_2^2}\left(\sum_{i=1}^{n}x_i\right)$$

(iii) $E(\theta|\underline{x}) = (1 - Z)\ \mu + Z\left(\frac{\sum_{i=1}^{n}x_i}{n}\right)$

which is in the form of a credibility estimate, and

$$Z = \frac{n\sigma_2^2}{\sigma_1^2+n\sigma_2^2}$$

is the credibility factor.

(iv)

	Company A	*Company B*
n	5	5
σ_1^2	500	350
σ_2^2	800	600
$Z = \dfrac{\sigma_2^2}{\sigma_2^2+\frac{\sigma_1^2}{n}}$	0.8889	0.8955
$\overline{x}$	439	356
μ	400	300

Credibility Premium $= Z\bar{x} + (1-Z)\mu$	434.7	350.1
Premium (CP + 25%)	543.3	437.7

(v) σ_1^2 increases — Reduces credibility factor and hence credibility premium moves towards prior mean.

σ_2^2 increases — Increases credibility factor and hence credibility premium moves towards sample mean.

END OF EXAMINERS' REPORT