

**Subject CT6 — Statistical Methods  
Core Technical**

**EXAMINERS' REPORT**

**April 2009**

**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.

R D Muckart  
Chairman of the Board of Examiners

June 2009

## Comments

Comments on solutions presented to individual questions for this April 2009 paper are given below.

- Q1** *This was generally well answered with various examples for liability insurance mentioned.*
- Q2** *This is a standard theoretical question. Many candidates scored full marks here.*
- Q3** *Many candidates misspecified the parameters of the Beta distribution in (i)*
- Q4** *Some candidates ignored the premium of 60 so all their figures were out by this amount. This however did not affect the answers in (ii) and (iii).*
- Q5** *Part (i) was generally not well answered despite the fact that it is a standard result about conditional expectations. Alternative correct solutions were also presented. The following parts of the question were straightforward.*
- Q6** *This question was well answered with many candidates scoring full marks. The final figures could differ from those shown in the model solution due to differences in rounding in the intermediate steps. Full credit was given for these provided a consistent approach was taken to rounding.*
- Q7** *Many candidates were able to give a correct general idea, but very few were able to specify a complete algorithm so that not many scored full marks.*
- Q8** *This was a straightforward question where many candidates scored well.*
- Q9** *Many marks were dropped in the last stages of the solution.*
- Q10** *This was one of the least well answered questions and many marks were dropped particularly in parts (ii) and (iii).*
- Q11** *Many candidates scored well here despite the lengthy calculations required. Some however dropped marks at the calculation of  $P(\text{claim/accident})$  at the 50% discount level because they did not distinguish between the cases where the accident was due to a criminal offences and those where it was not.*

- 1** The essential characteristic of liability insurance is to provide indemnity where the insured, owing to some form of negligence, is legally liable to pay compensation to a third party.

Examples

- Employer's liability
- Motor 3<sup>rd</sup> party liability
- Public liability
- Product liability
- Professional indemnity

- 2** (i) We need to express the distribution function of the Gamma distribution in the form:

$$f_Y(y; \theta, \varphi) = \exp \left[ \frac{(y\theta - b(\theta))}{a(\varphi)} + c(y, \varphi) \right]$$

Suppose  $Y$  has a Gamma distribution with parameters  $\alpha$  and  $\lambda$ . Then

$$f_Y(y) = \frac{\lambda^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\lambda y}$$

And substituting  $\lambda = \frac{\alpha}{\mu}$  we can write the density as

$$f_Y(y; \theta, \varphi) = \frac{\alpha^\alpha}{\mu^\alpha \Gamma(\alpha)} y^{\alpha-1} e^{-\frac{y\alpha}{\mu}}$$

$$f_Y(y; \theta, \varphi) = \exp \left[ \left( -\frac{y}{\mu} - \log \mu \right) \alpha + (\alpha - 1) \log y + \alpha \log \alpha - \log \Gamma(\alpha) \right]$$

Which is in the right form with  $\theta = -\frac{1}{\mu}$ ;  $\varphi = \alpha$ ;  $a(\varphi) = \frac{1}{\varphi}$ ;  $b(\theta) = -\log(-\theta)$  and  $c(y, \varphi) = (\varphi - 1) \log y + \varphi \log \varphi - \log \Gamma(\varphi)$ .

Thus the natural parameter is  $\frac{1}{\mu}$ , ignoring the minus sign, and the scale parameter is  $\alpha$ .

- (ii) The corresponding link function is  $\frac{1}{\mu}$ .

3 (i)  $f(\theta|x) \propto f(x|\theta)f(\theta)$

$$\propto \theta^9(1-\theta)^{27} \theta^{\alpha-1}(1-\theta)^{\beta-1}$$

$$\propto \theta^9(1-\theta)^{27} \theta^1(1-\theta)^3$$

$$\propto \theta^{10}(1-\theta)^{30}$$

Which is the pdf of a Beta(11,31) distribution.

- (ii) Under all or nothing loss, the Bayes estimate is the value that maximises the pdf of the posterior.

$$f(\theta) = C \times \theta^{10}(1-\theta)^{30}$$

$$f'(\theta) = C \left( 10\theta^9(1-\theta)^{30} + \theta^{10} \times 30(1-\theta)^{29} \times -1 \right)$$

$$= C\theta^9(1-\theta)^{29} (10(1-\theta) - 30\theta)$$

And  $f'(\theta) = 0$  when

$$10(1-\theta) - 30\theta = 0$$

$$40\theta = 10$$

$$\theta = 1/4$$

We can check this is a maximum by observing that  $f(0.25) > 0$  whilst  $f(0) = f(1) = 0$ . Since the maximum on  $[0,1]$  must occur either at the end-points of the interval, or at turning point we can see that we do have a maximum.

- 4 (i) The loss table is as follows:

Claims	Impact of reinsurance			Insurer's Loss			
	A	B	C	A	B	C	
0	0	-10	-20	-60	-50	-40	
100	0	-10	5	40	50	35	
150	0	-10	17.5	90	100	72.5	
200	0	40	30	140	100	110	

- (ii) If there are no claims, A gives the best result.  
 If claims are 100 C gives the best result.  
 If claims are 200 then B gives the best result.

So each strategy can be best under certain circumstances, and so no approach is dominated.

(iii) The maximum losses are:

A	140
B	100
C	110

The lowest is for B, so approach B is the minimax solution.

- 5** (i) Let  $f(s)$  denote the marginal probability density for  $S$  and let  $f(s|\lambda)$  denote the conditional probability density for  $S|\lambda$ . Then

$$\begin{aligned} E[E(S|\lambda)] &= \sum_{i=1}^3 p(\lambda_i) \int_0^{\infty} sf(s|\lambda) ds \\ &= \int_0^{\infty} s \sum_{i=1}^3 p(\lambda_i) f(s|\lambda) ds \end{aligned}$$

But  $\sum_{i=1}^3 p(\lambda_i) f(s|\lambda) = f(s)$  by definition, so

$$E[E(S|\lambda)] = \int_0^{\infty} sf(s) ds = E(S)$$

(ii) Using the results for compound distributions, we have:

$$E(S|\lambda) = E(N|\lambda)E(X|\lambda) = E(N|\lambda)E(X) = 4\lambda$$

$$\begin{aligned} \text{Var}(S|\lambda) &= E(N|\lambda)\text{Var}(X) + \text{Var}(N|\lambda)E(X)^2 \\ &= \lambda \times 16 + \lambda \times 4^2 \\ &= 32\lambda \end{aligned}$$

(iii)  $E(S) = E[E(S|\lambda)] = E(4\lambda) = 4E(\lambda) = 12$

- (iv) First note that  $E(\lambda) = 3$  and

$$\text{Var}(\lambda) = 0.2 \times 2^2 + 0.6 \times 3^2 + 0.2 \times 4^2 - 9 = 0.4$$

$$\begin{aligned} \text{Var}(S) &= \text{Var}[E(S|\lambda)] + E[\text{Var}(S|\lambda)] \\ &= \text{Var}(4\lambda) + E(32\lambda) \\ &= 16 \times \text{Var}(\lambda) + 32E(\lambda) \\ &= 16 \times 0.4 + 32 \times 3 \\ &= 102.4 \end{aligned}$$

- 6** The accumulated claims are:

	<i>Development Year</i>			
<i>Accident year</i>	<i>0</i>	<i>1</i>	<i>2</i>	<i>Ult</i>
2006	442	593	643	<b>643</b>
2007	623	734		<b>796</b>
2008	681			<b>927</b>

The accumulated claim costs are:

	<i>Development Year</i>			
<i>Accident year</i>	<i>0</i>	<i>1</i>	<i>2</i>	
2006	6321	8222	8923	
2007	7012	9249		
2008	7278			

The average costs per claim are:

	<i>Development Year</i>			
<i>Accident year</i>	<i>0</i>	<i>1</i>	<i>2</i>	<i>Ult</i>
2006	14.301	13.865	13.877	<b>13.877</b>
2007	11.255	12.601		<b>12.612</b>
2008	10.687			<b>11.115</b>

The grossing-up factors for the claim numbers:

<i>Accident year</i>	<i>0</i>	<i>1</i>	<i>2</i>
2006	0.687	0.922	1
2007	0.783	0.922	
2008	0.735		

The grossing-up factors for the average cost per settled claim:

<i>Accident year</i>	<i>0</i>	<i>1</i>	<i>2</i>
2006	1.031	0.999	1
2007	0.892	0.999	
2008	0.961		

Projected loss:  $643 \times 13.877 + 796 \times 12.612 + 927 \times 11.115 = 29265.67$

Claims paid to date:  $8923 + 9249 + 7278 = 25450$

Outstanding Claims are therefore: 3815.7

The calculations are based on rounding the intermediate calculations to the accuracy shown. Different rounding will give slightly different answers, and is acceptable. Carrying all the calculations through in full accuracy gives a solution of 3808.0.

**7** (i) We must find  $C$  where

$$C = \max \frac{f(x)}{h(x)} = \max \frac{6x(1-x)}{2(1-x)} = \max 3x = 3$$

We need to be able to generate a random variable from the distribution with density  $h(x)$ . We can do this as follows.

First note that the cdf of  $h(x)$  is

$$H(x) = \int_0^x g(t) dt = \int_0^x 2(1-t) dt = \left[ 2t - t^2 \right]_0^x = 2x - x^2$$

Given a random sample  $z$  from  $U(0,1)$  we can use the inverse transform method to sample from  $h$  by setting:

$$2x - x^2 = z$$

$$x^2 - 2x + z = 0$$

$$x = \frac{2 \pm \sqrt{4 - 4z}}{2} = 1 \pm \sqrt{1 - z}$$

And we can see that the solution we want is  $x = 1 - \sqrt{1 - z}$

The algorithm to generate a sample  $y$  from  $f$  is then:

- Sample  $z$  from  $U(0,1)$
- Generate  $x$  from  $h(x)$  via  $x = 1 - \sqrt{1 - z}$
- Generate  $u$  from  $U(0,1)$
- If  $u < \frac{f(x)}{3h(x)} = \frac{6x(1-x)}{6(1-x)} = x$  then generate  $y = x$  otherwise begin again.

- (ii) On average, we expect to use  $C = 3$  realisations from  $h$  to generate one sample from  $f$ .
- (iii) In this case, we must find the maximum value of  $f(x)$ .

$$f'(x) = 6 - 12x$$

And  $f'(x) = 0$  when  $x = 0.5$

Since  $f(0) = f(1) = 0$  and  $f(0.5) = 6/4 = 1.5$  we can see that this is the maximum.

Since  $C$  is lower for  $g(x) = 1$ , using the constant function would be more efficient.

- 8**
- (i)  $U(t)$  represents the insurers surplus at time  $t$ . It represents the initial surplus plus cumulative premiums received less claims incurred:

$$U(t) = U + ct - S(t)$$

Where  $U$  is the initial surplus and  $c$  is the annual premium income (assumed payable continuously).

- (ii) (a)  $\psi(U, t) = P(U(s) < 0 \text{ for some } 0 \leq s \leq t \text{ given that } U(0) = U)$
- (b)  $\psi_h(U, t) = P(U(s) < 0 \text{ for some } s, s = h, 2h, \dots, t - h, t \text{ given that } U(0) = U)$

- (iii) (a) We can say that  $\psi(10, 2) > \psi(10, 1) > \psi(20, 1)$

The first inequality holds because the longer the period considered when checking, the more likely that ruin will occur. The second inequality holds because a larger initial surplus reduces the probability of ruin (higher claims are required to cause ruin).

- (b) We can reach no definite conclusion here. The first term has a longer term suggesting a higher probability, but a higher initial surplus suggesting a lower probability. We can't say anything about the size of the two effects, so we can't reach a definite conclusion.

- (c) We can say that  $\psi_{0.5}(10, 2) \leq \psi_{0.25}(10, 2)$ .

This is because the second term checks for ruin at the same times as the first term, as well as at some additional times. So the probability of ruin when checking over the larger set of times must be higher.



**9** (i)  $E(X_i) = \alpha + (1+k)(1-\alpha)$   
 $= 1+k(1-\alpha)$

$$\begin{aligned} \text{Var}(X_i) &= E(X_i^2) - E(X_i)^2 \\ &= \alpha + (1+k)^2(1-\alpha) - (1+k(1-\alpha))^2 \\ &= \alpha + (1-\alpha) + 2k(1-\alpha) + k^2(1-\alpha) - 1 - 2k(1-\alpha) - k^2(1-\alpha)^2 \\ &= k^2(1-\alpha)(1-(1-\alpha)) \\ &= k^2\alpha(1-\alpha) \end{aligned}$$

- (ii) Let  $Y$  denote the aggregate claims in a year. Then  $Y$  has a compound Poisson distribution, and using the standard results from the tables:

$$E(Y) = 500 \times E(X_i) = 500 + 500k(1-\alpha) = 500 + 400k = 633.60$$

And

$$\begin{aligned} \text{Var}(Y) &= 500 \times E(X_i^2) \\ &= 500 \left( \alpha + (1+k)^2(1-\alpha) \right) \\ &= 500 \left( \alpha + (1-\alpha) + 2k(1-\alpha) + k^2(1-\alpha) \right) \\ &= 500 + (1000k + 500k^2)(1-\alpha) \\ &= 500 + 800k + 400k^2 \\ &= 811.82 \end{aligned}$$

Using a normal approximation, we find

$$\begin{aligned} P(Y > 700) &= P\left(Z > \frac{700 - 633.6}{\sqrt{811.82}}\right) \\ &= P(Z > 2.33) \\ &= 0.01 \end{aligned}$$

- 10** (i) Since  $Y_t$  is stationary, we know there is a constant  $\mu_Y$  such that  $E(Y_t) = \mu_Y$  for all values of  $t$ .

But then  $E(X_t) = E(a + bt + Y_t) = a + bt + \mu_Y$  which depends on  $t$  since  $b$  is non-zero.

Hence  $X_t$  is not stationary.

- (ii) First note that

$$\begin{aligned}
 E(\Delta X_t) &= E(X_t - X_{t-1}) \\
 &= E(X_t) - E(X_{t-1}) \\
 &= E(a + bt + Y_t) - E(a + b(t-1) + Y_{t-1}) \\
 &= a + bt + \mu_Y - a - b(t-1) - \mu_Y \\
 &= b
 \end{aligned}$$

i.e. the mean is a constant independent of  $t$ .

Secondly,

$$\begin{aligned}
 \text{Cov}(\Delta X_t, \Delta X_{t-s}) &= \text{Cov}(X_t - X_{t-1}, X_{t-s} - X_{t-s-1}) \\
 &= \text{Cov}(b + Y_t - Y_{t-1}, b + Y_{t-s} - Y_{t-s-1}) \\
 &= \text{Cov}(Y_t - Y_{t-1}, Y_{t-s} - Y_{t-s-1}) \\
 &= \text{Cov}(Y_t, Y_{t-s}) - \text{Cov}(Y_{t-1}, Y_{t-s}) - \text{Cov}(Y_t, Y_{t-s-1}) + \text{Cov}(Y_{t-1}, Y_{t-s-1}) \\
 &= \gamma_Y(s) - \gamma_Y(s-1) - \gamma_Y(s+1) + \gamma_Y(s) \\
 &= 2\gamma_Y(s) - \gamma_Y(s-1) - \gamma_Y(s+1)
 \end{aligned}$$

Since the autocovariance depends only on the lag  $s$ , and the mean is constant, the new series is stationary.

- (iii) Suppose that  $Y_t = e_t + \beta e_{t-1}$  where  $e_t$  is a white noise process with variance  $\sigma^2$ .

Then

$$\begin{aligned}
 \Delta X_t &= a + bt + e_t + \beta e_{t-1} - a - b(t-1) - e_{t-1} - \beta e_{t-2} \\
 &= b + e_t + (\beta - 1)e_{t-1} - \beta e_{t-2} \\
 &= b + (1 + (\beta - 1)B - \beta B^2)e_t \\
 &= b + (1 - B)(1 + \beta B)e_t
 \end{aligned}$$

So the lag polynomial has a unit root, and hence the time series is not invertible.

Now

$$\begin{aligned}
 \text{Var}(Y_t) &= \text{Var}(e_t + \beta e_{t-1}) \\
 &= \text{Var}(e_t) + \beta^2 \text{Var}(e_{t-1}) \\
 &= (1 + \beta^2)\sigma^2
 \end{aligned}$$

$$\text{So } \gamma_Y(0) = (1 + \beta^2)\sigma^2$$

$$\text{Also } \gamma_Y(1) = \gamma_Y(-1) = \text{Cov}(e_t + \beta e_{t-1}, e_{t-1} + \beta e_{t-2}) = \beta \sigma^2$$

So using the result from part (ii)

$$\begin{aligned} \text{Var}(\Delta X_t) &= 2\gamma_Y(0) - \gamma_Y(1) - \gamma_Y(-1) \\ &= 2(1 + \beta^2)\sigma^2 - 2\beta\sigma^2 \\ &= 2(1 - \beta + \beta^2)\sigma^2 \end{aligned}$$

And finally

$$\begin{aligned} \text{Var}(\Delta X_t) - \text{Var}(Y_t) &= (2 - 2\beta + 2\beta^2)\sigma^2 - (1 + \beta^2)\sigma^2 \\ &= (1 - 2\beta + \beta^2)\sigma^2 \\ &= (1 - \beta)^2\sigma^2 > 0 \end{aligned}$$

- 11** (i) Premiums at the three levels are £1,000, £750 and £500.

**0% level**

Claims involving criminal offences make no difference here.

Claim:  $1000 + 750 = 1750$

No Claim:  $750 + 500 = 1250$

No claim if the accident cost is less than 500.

**25% level**

Claims involving criminal offences make no difference here.

Claim:  $1000 + 750 = 1750$

No Claim:  $500 + 500 = 1000$

No claim if the accident cost is less than 750.

**50% level**

Criminal offence claims make a difference here so we distinguish two scenarios:

*No criminal offence*

Claim:  $750 + 500 = 1250$

Not claim:  $500 + 500 = 1000$

No claim if the accident cost is less than 250.

*Criminal offence*

Claim:  $1000 + 750 = 1750$

Not claim:  $500 + 500 = 1000$

No claim if the accident cost is less than 750.

- (ii) Let  $X$  be the repair cost for each accident then  $X \sim \text{Exp}\left(\frac{1}{400}\right)$

$$\text{i.e. } P(X > x) = e^{-\frac{x}{400}}$$

$$\text{For 0\% level: } P(\text{claim}|\text{accident}) = P(X > 500) = e^{-\frac{5}{4}} = 0.2865$$

$$\text{For 25\% level: } P(\text{claim}|\text{accident}) = P(X > 750) = e^{-\frac{75}{40}} = 0.1534$$

$$\begin{aligned} \text{For 50\% level: } P(\text{claim}|\text{accident}) &= P(\text{claim}|\text{criminal offence}) \\ &\quad + P(\text{claim}|\text{no criminal offence}) \end{aligned}$$

$$= P(X > 750) \times 0.1 + P(X > 250) \times 0.9$$

$$= 0.1e^{-\frac{75}{40}} + 0.9e^{-\frac{25}{40}} = 0.4971$$

- (iii) Note that  $P(\text{Accident}) = 0.1$  and  $P(\text{criminal offence}|\text{accident}) = 0.1$ .

Hence at 0% level:

$$\begin{aligned} P_{11} &= P(\text{current level}) = P(\text{make a claim}) \\ &= P(\text{claim}|\text{accident}) P(\text{accident}) = 0.02865 \end{aligned}$$

$$P_{12} = P(\text{move to 25\% level}) = P(\text{not claim}) = 1 - 0.02865 = 0.97135$$

$$P_{13} = P(\text{move to 50\% level}) = 0$$

At 25% level:

$$\begin{aligned} P_{21} &= P(\text{move down to 0\% level}) = P(\text{make a claim}) \\ &= P(\text{claim}|\text{accident}) P(\text{accident}) = 0.015335 \end{aligned}$$

$$P_{22} = 0 \text{ since it can never remain at this level for more than a year at a time}$$

$$P_{23} = 1 - P_{21} = 1 - 0.015335 = 0.984665$$

At the 50% levels:

$$\begin{aligned}P_{31} &= P(\text{make a claim and criminal offence involved}) \\&= P(\text{accident}) P(\text{criminal offence}|\text{accident}) P(\text{claim}|\text{criminal offence}) \\&= 0.1 P(X > 750) = 0.01 \times e^{-\frac{75}{40}} = 0.0015335\end{aligned}$$

$$\begin{aligned}P_{32} &= P(\text{make a claim and no criminal offence involved}) P(\text{accident}) \\&\quad P(\text{no criminal offence}|\text{accident}) P(\text{claim}|\text{no criminal offence}) \\&= 0.1 \times 0.9 \times P(X > 250) = 0.09 \times e^{-\frac{25}{40}} = 0.0481735\end{aligned}$$

$$\begin{aligned}P_{33} &= P(\text{no claim}) = 1 - P(\text{claim}) \\&= 1 - 0.01 \times e^{-\frac{75}{40}} - 0.09 \times e^{-\frac{25}{40}} \\&= 0.950293 \text{ (rounding errors possible here)}\end{aligned}$$

Therefore the transition matrix is now

$$P = \begin{pmatrix} 0.02865 & 0.97135 & 0 \\ 0.015335 & 0 & 0.984665 \\ 0.0015335 & 0.0481735 & 0.950293 \end{pmatrix}$$

For the stationary distribution we need to find  $\pi = (\pi_0, \pi_1, \pi_2)$  such that  $\pi P = \pi$ :

$$\begin{aligned}\pi_0 0.02865 + \pi_1 0.015335 + \pi_2 0.0015335 &= \pi_0 \\ \pi_0 0.97135 + 0 + \pi_2 0.481735 &= \pi_1 \\ 0 + \pi_1 0.984665 + \pi_2 0.950293 &= \pi_2 \\ \pi_0 + \pi_1 + \pi_2 &= 1\end{aligned}$$

Further calculations show that  $\pi_0 = 0.002256424$ ,  $\pi_1 = 0.047946812$  and  $\pi_2 = 0.949796764$ .

**END OF EXAMINERS' REPORT**