INSTITUTE AND FACULTY OF ACTUARIES

EXAMINERS' REPORT

September 2010 examinations

Subject CT6 — Statistical Methods
Core Technical

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.

T J Birse
Chairman of the Board of Examiners

December 2010

1 We know that, approximately, $\theta - \hat{\theta} \approx N(0, \tau^2/n)$ where τ^2 can be approximated by 0.15.

Then $P\left[-1.96 \le \dfrac{\theta - \hat{\theta}}{\sqrt{0.15/n}} \le 1.96\right] = 0.95$

And we require $1.96 \times \sqrt{\dfrac{0.15}{n}} \le 0.01$

That is $n \ge \dfrac{1.96^2 \times 0.15}{0.01^2} = 5762.4$ i.e. n must be at least 5763

This question was generally poorly answered.

2 The adjustment coefficient satisfies the equation:

$$\lambda + \lambda\mu(1+\theta)R = \lambda M_X(R)$$

Where X is exponentially distributed with mean μ so that $M_X(t) = \dfrac{1/\mu}{1/\mu - t} = \dfrac{1}{1-\mu t}$

So we have $1 + \mu(1+\theta)R = \dfrac{1}{1-\mu R}$

and so $1 - \mu R + R\mu(1+\theta)(1-\mu R) = 1$

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

Dividing through by μR gives

$$\mu R(1+\theta) = \theta$$

So $R = \dfrac{\theta}{\mu(1+\theta)}$.

This question was well answered by most candidates.

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

Then we have:

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

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

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

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

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

(ii) The 99.5th percentile loss is given by

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

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

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

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

Most candidates scored well on this question.

4 Let the prior distribution of μ have a Gamma distribution with parameters α and λ as per the tables.

Then $\frac{\alpha}{\lambda}=50$ and $\frac{\alpha}{\lambda^2}=15^2$

Then dividing the first by the second $\lambda=\frac{50}{15^2}=0.22222$

And so $\alpha=50\times 0.22222=11.111111$

The posterior distribution of μ is then given by

$$f(\mu|x) \propto f(x|\mu)\, f(\mu)$$

$$\propto e^{-10\mu} \times \mu^{630} \times \mu^{10.11111} e^{-0.22222\mu}$$
$$\propto \mu^{640.11111} e^{-10.22222\mu}$$

Which is the pdf of a Gamma distribution with parameters $\alpha' = 641.11111$ and $\lambda' = 10.22222$

Now under all or nothing loss, the Bayesian estimate is given by the mode of the posterior distribution. So we must find the maximum of

$$f(x) = x^{640.11111} e^{-10.2222x} \text{ (we may ignore constants here)}$$

Differentiating:

$$f'(x) = e^{-10.22222x}\left(-10.2222x^{640.11111} + 640.1111x^{639.11111}\right)$$
$$= x^{639.1111} e^{-10.22222x} (-10.2222x + 640.11111)$$

And setting this equal to zero we get

$$x = \frac{640.111111}{10.22222} = 62.62$$

Alternatively, credit was given for differentiating the log of the posterior (which is simpler). This question was well answered by most candidates.

5 (i) The overall mean is given by $\bar{X} = \dfrac{127.9 + 88.9 + 149.7}{3} = 122.167$

$$E\left(s^2(\theta)\right) = \frac{1}{3}\sum_{i=1}^{3}\left(\frac{1}{6}\sum_{j=1}^{7}(X_{ij} - \bar{X}_i)^2\right) = \frac{335.1 + 65.1 + 33.9}{3} = 144.7$$

$$Var(m(\theta)) = \frac{1}{2}\sum_{i=1}^{3}\left(\bar{X}_i - \bar{X}\right)^2 - \frac{1}{7}E(S^2(\theta))$$
$$= \frac{(127.9 - 122.1)^2 + (88.9 - 122.1)^2 + (149.7 - 122.1)^2}{2} - \frac{144.7}{7}$$
$$= 928.14$$

So the credibility factor is $Z = \dfrac{7}{7 + {}^{144.7}\!/_{928.14}} = 0.978213$

And the credibility premia for the risks are:

For risk 1 : $0.978213 \times 127.9 + (1 - 0.978213) \times 122.167 = 127.8$
For risk 2 : $0.978213 \times 88.9 + (1 - 0.978213) \times 122.167 = 89.6$
For risk 3 : $0.978213 \times 149.7 + (1 - 0.978213) \times 122.167 = 149.1$

(ii) The data show that the variation within risks is relatively low (the S_i^2 are low, especially for the 2nd and 3rd risks) but there seems to be quite a high variation between the average claims on the risks.

With the S_i^2 being low, this variation cannot be explained just by variability in the claims, and must be due to variability in the underlying parameter.

This means that we can put relatively little weight on the information provided by the data set as a whole, and must put more on the data from the individual risks, leading to a relatively high credibility factor.

Most candidates scored well on part (i). Only the better candidates were able to give a clear explanation in part (ii).

6 (i) We must write *f(x)* in the form:

$$f(x) = \exp\left[\frac{x\theta - b(\theta)}{a(\phi)} + c(x,\phi)\right]$$

For some parameters θ, ϕ and functions *a,b* and *c*.

$$f(x) = \frac{\alpha^\alpha}{\mu^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x\alpha}{\mu}}$$

$$= \exp\left[\left(-\frac{x}{\mu} - \log\mu\right)\alpha + (\alpha - 1)\log x + \alpha\log\alpha - \log\Gamma(\alpha)\right]$$

Which is of the required form with:

$$\theta = -\frac{1}{\mu}$$

$$\phi = \alpha$$

$$a(\phi) = \frac{1}{\phi}$$

$$b(\theta) = -\log(-\theta) = \log\mu$$

$$c(x,\phi) = (\phi-1)\log x + \phi\log\phi - \log\Gamma(\phi)$$

(ii) The mean and variance for members of the exponential family are given by $b'(\theta)$ and $a(\phi)b''(\theta)$.

In this case $b'(\theta) = -\dfrac{1}{\theta} = \mu$

$b''(\theta) = \theta^{-2} = \mu^2$ so the variance is μ^2/α as required.

Generally well answered, though many candidates did not score full marks on part (i) because they failed to specify all the parameters involved.

7 (i) First note that the probability of a claim exceeding 100 is $e^{-100\lambda}$.

The likelihood function for the given data is:

$$L = C \times \lambda^{85} e^{-85\times 42\times\lambda} \times (e^{-100\times\lambda})^{39}$$

Where C is some constant. Taking logarithms gives

$$l = \log L = C' + 85\log\lambda - 85\times 42\times\lambda - 100\times 39\times\lambda$$

Differentiating with respect to λ gives

$$\frac{\partial l}{\partial\lambda} = \frac{85}{\lambda} - 85\times 42 - 100\times 39$$

Setting this expression equal to zero we get:

$$\hat{\lambda} = \frac{85}{85\times 42 + 100\times 39} = 0.011379$$

And this gives a maximum since $\dfrac{\partial^2 l}{\partial\lambda^2} = -\dfrac{85}{\lambda^2} < 0$

(ii) We must first calculate the mean amount paid by the insurer per claim. This is

$$\int_0^{100} x\lambda e^{-\lambda x}dx + 100P(X > 100) = \left[-xe^{-\lambda x}\right]_0^{100} + \int_0^{100} e^{-\lambda x}dx + 100e^{-100\lambda}$$

$$= -100e^{-100\lambda} + \left[-\frac{1}{\lambda}e^{-\lambda x}\right]_0^{100} + 100e^{-100\lambda}$$

$$= \frac{1}{\lambda}\left(1 - e^{-100\lambda}\right)$$

So we must show the given value of λ results in the actual average paid by the insurer. This is $\dfrac{85 \times 42 + 39 \times 100}{85 + 39} = 60.24$

Substituting for λ in the expression derived above, we get

$$\frac{1}{0.011164}\left(1 - e^{-100 \times 0.0111654}\right) = \frac{0.6725435}{0.011164} = 60.24 \text{ as required.}$$

Stronger candidates scored well on this question, whereas the weaker candidates struggled with the calculations required in part (ii).

8 (i) The adjustment coefficient satisfies the equation

$$\lambda + \lambda(1+\theta)E(X_1)R = \lambda M_{X_1}(R)$$

That is $1 + (1+\theta)E(X_1)R = \sum_{j=1}^{M} e^{Rj} p_j$

Applying the inequality given in the question we have

$$1 + (1+\theta)E(X_1)R \le \sum_{j=1}^{M} p_j\left(\frac{j}{M}e^{RM} + 1 - \frac{j}{M}\right)$$

So $1 + (1+\theta)E(X_1)R \le \dfrac{e^{RM}}{M}\sum_{j=1}^{M} jp_j + 1 - \dfrac{1}{M}\sum_{j=1}^{M} jp_j = \dfrac{e^{RM}E(X_1)}{M} + 1 - \dfrac{E(X_1)}{M}$

So $(1+\theta)E(X_1)R \le \dfrac{E(X_1)}{M}(e^{RM} - 1)$

and so

$$(1+\theta)R \le \frac{1}{M}\left(1+RM+\frac{R^2M^2}{2!}+\frac{R^3M^3}{3!}+\cdots-1\right)=R\left(1+\frac{RM}{2!}+\frac{R^2M^2}{3!}+\cdots\right)$$

$$(1+\theta)R < R\left(1+RM+\frac{R^2M^2}{2!}+\cdots\right)=R\times e^{RM}$$

Taking logs, we have

$$\log(1+\theta) < RM$$

And so $R > \dfrac{\log(1+\theta)}{M}$ as required.

To get the other inequality, we go back to

$$1+(1+\theta)E(X_1)R=\sum_{j=1}^{M} e^{Rj}p_j$$

And so $1+(1+\theta)E(X_1)R > \displaystyle\sum_{j=1}^{M} p_j(1+Rj+\frac{R^2j^2}{2}) = 1+RE(X_1)+\frac{R^2}{2}E(X_i^2)$

So we have $(1+\theta)m_1R > m_1R+\dfrac{R^2m_2}{2}$

i.e. $\theta m_1 > \dfrac{Rm_2}{2}$

i.e. $R < \dfrac{2\theta m_1}{m_2}$ as required.

(ii) (a) In this case we have:

$$M = 3$$

And $E(X_1) = 2.5$ and $E(X_1^2) = (4+9)/2 = 6.5$

So the inequality in the question gives:

$$\frac{1}{3}\log 1.3 < R < \frac{2\times 0.3\times 2.5}{6.5}$$

That is $0.08745 < R < 0.23077$

(b) By Lundberg;s inequality $\psi(U) \le e^{-RU} \le e^{-0.08745U}$.

This question was not well answered, with relatively few candidates scoring more than 5 marks.

9 (i) The development ratio for development year 2 to development year 3 is given by 1862.3/1820 = 1.023242

Therefore W = 1762 × 1.023242 = 1803.0

Because there is no claims development beyond development year 3 X = 1803.0 also.

The development factor from development year 1 to ultimate is given by 2122.5/1805 = 1.1759003

So the ratio from development year 1 to development year 2 is given by 1.1759003/1.023242 = 1.149190785

But under the definition of the chain ladder approach, this is calculated as:

$$1.149190785 = \frac{1762+1820}{Y+1485} = \frac{3582}{Y+1485}$$

So $Y = \dfrac{3582}{1.149190785} - 1485 = 1632.0$

(ii) We require the development ratio from year 0 to year 1; this is given by:

$$\frac{1485+1632+1805}{1001+1250+1302} = \frac{4922}{3553} = 1.385308$$

The development factor to ultimate is therefore

$$1.385308 \times 1.149190785 \times 1.023242 = 1.628984285$$

And so $Z = 2278.8 - 2500 \times 0.9 \times \left(1 - \dfrac{1}{1.628984285}\right) = 1410.0$

(iii) The outstanding claims reserve is

$$1862.3 + 2122.5 + 2278.8 - 1820 - 1805 - 1410 = 1228.6$$

This slightly unusual question was nevertheless generally well answered, showing that candidates understood the principles underlying the calculations. Many candidates scored full marks here.

10 (i) We require:

- The risk of flood damage is a constant p for each building.
- There can only be one claim per policy per year.
- The risk of flood damage is independent from building to building.

(ii) Let the individual claim amounts net of re-insurance be X. Then

$$E(\alpha X) = \alpha E(X) = 400\alpha$$

And $Var(\alpha X) = \alpha^2 Var(X) = (50\alpha)^2$

So if Y represents the aggregate annual claims net of re-insurance, then we have:

$$E(Y) = 10,000 \times 0.03 \times 400\alpha = 120,000\alpha$$

and

$$Var(Y) = 10,000 \times 0.03 \times (50\alpha)^2 + 10,000 \times 0.03 \times 0.97 \times (400\alpha)^2 = 47,310,000\alpha^2$$
$$= (6,878.23\alpha)^2$$

We require α to be chosen so that

$$P(Y > 120,000) = 0.01$$

i.e. $P(N(0,1) > \dfrac{120,000 - 120,000\alpha}{6,878.23\alpha}) = 0.01$

$$\frac{120,000 - 120,000\alpha}{6,878.23\alpha} = 2.3263$$

i.e. $\alpha = \dfrac{120,000}{120,000 + 2.3263 \times 6,878.23} = 0.8823476; \alpha = 88.2\%$ *to* $3sf$

(iii) The mean claim amount for the re-insurer is $(1 - 0.882) \times 400 = 47.20$

The annual premiums for reinsurance are $10,000 \times 0.03 \times 47.20 \times 1.15 = 16,284$

(iv) We must show that using a retention of 358.50 to calculate the premium for the individual excess of loss arrangement gives the same result as the proportional reinsurance arrangement in part (ii).

We first calculate the mean claim amount paid by re-insurer. This is equal to

$$\int_{358.50}^{\infty} (x - 358.50) f(x) dx$$

$$= 400\left[1 - \Phi(\frac{358.50-400}{50})\right] - 50\left[0 - \varphi(\frac{358.50-400}{50})\right] - 358.50 \times \left(1 - \Phi\left(\frac{358.50-400}{50}\right)\right)$$

This gives

$$400\left[1 - \Phi(-0.83)\right] + 50\varphi(-0.83) - 358.50 \times (1 - \Phi(-0.83))$$

$$= 400 \times 0.79673 + 50 \times \frac{1}{\sqrt{2\pi}} e^{-\frac{0.83^2}{2}} - 358.50 \times 0.79673$$

$$= 47.20$$

Then the aggregate premium charged will be $10,000 \times 0.03 \times 47.20 \times 1.15$ = 16,284 which is the same as under the first arrangement as required.

Carrying forward more than 3 significant figures from the result in (ii) gives a slightly different value in (iii). To full accuracy, the solution in (iii) becomes 16,236 resulting in a minor discrepancy between the answers in (iii) and (iv). This appears not to have concerned candidates who were generally happy to observe that the results in (iii) and (iv) were approximately equal. The examiners gave credit for either approach.

This question was a good differentiator – the better prepared candidates were able to score well whilst weaker candidates struggled.

11 (i) Let B be the backward shift operator. Then the time series has the form:

$$(1 - 2\alpha B + \alpha^2 B^2) Y_t = e_t$$

$$(1 - \alpha B)^2 Y_t = e_t$$

And the roots of the characteristic equation will have modulus greater than 1 ands so the series will be stationary provided that $|\alpha| < 1$.

(ii) Firstly, note that $Cov(Y_t, e_t) = Cov(e_t, e_t) = \sigma^2$

So, taking the covariance of the defining equation with Y_t we get:

$$\gamma_0 = 2\alpha\gamma_1 - \alpha^2\gamma_2 + \sigma^2 \quad \text{(A)}$$

Taking the covariance with Y_{t-1} we get

$$\gamma_1 = 2\alpha\gamma_0 - \alpha^2\gamma_1$$

i.e. $(1+\alpha^2)\gamma_1 = 2\alpha\gamma_0 \quad$ (B)

Finally, taking the covariance with Y_{t-2} gives:

$$\gamma_2 = 2\alpha\gamma_1 - \alpha^2\gamma_0 \quad \text{(C)}$$

In general, for $k \geq 2$ we have $\gamma_k = 2\alpha\gamma_{k-1} - \alpha^2\gamma_{k-2}$

Substituting the expression for γ_2 in (C) into (A) gives:

$$\gamma_0 = 2\alpha\gamma_1 - \alpha^2(2\alpha\gamma_1 - \alpha^2\gamma_0) + \sigma^2$$

So that

$$(1-\alpha^4)\gamma_0 = 2\alpha(1-\alpha^2)\gamma_1 + \sigma^2$$

And now substituting the expression for γ_1 in (B) we get

$$(1-\alpha^4)\gamma_0 = 2\alpha(1-\alpha^2) \times \frac{2\alpha\gamma_0}{(1+\alpha^2)} + \sigma^2$$

$$\left(1-\alpha^4 - \frac{4\alpha^2(1-\alpha^2)}{1+\alpha^2}\right)\gamma_0 = \sigma^2$$

$$(1+\alpha^2-\alpha^4-\alpha^6-4\alpha^2+4\alpha^4)\gamma_0 = (1+\alpha^2)\sigma^2$$

So $\gamma_0 = \dfrac{(1+\alpha^2)}{(1-3\alpha^2+3\alpha^4-\alpha^6)}\sigma^2 = \dfrac{(1+\alpha^2)}{(1-\alpha^2)^3}\sigma^2$

And so $\gamma_1 = \dfrac{2\alpha\gamma_0}{1+\alpha^2} = \left(\dfrac{2\alpha(1+\alpha^2)}{(1-\alpha^2)^3(1+\alpha^2)}\right)\sigma^2 = \dfrac{2\alpha}{(1-\alpha^2)^3}\sigma^2$

And more generally $\gamma_k = 2\alpha\gamma_{k-1} - \alpha^2\gamma_{k-2}$ (D)

(iii) Suppose $\gamma_{k-1} = A\alpha^{k-1} + (k-1)B\alpha^{k-1}$ and $\gamma_{k-2} = A\alpha^{k-2} + (k-2)B\alpha^{k-2}$ and substitute into (D).

$$\gamma_k = 2\alpha A\alpha^{k-1} + 2\alpha(k-1)B\alpha^{k-1} - \alpha^2 A\alpha^{k-2} - (k-2)\alpha^2 B\alpha^{k-2}$$

$$= A(2\alpha^k - \alpha^k) + B(2\alpha^k(k-1) - (k-2)\alpha^k) = A\alpha^k + Bk\alpha^k$$

Which is of the correct form, so the general form of the expression holds.

Setting $k = 0$ we get $\gamma_0 = A$

So $A = \dfrac{(1+\alpha^2)}{(1-\alpha^2)^3}\sigma^2$

Setting $k = 1$ we get $\gamma_1 = (A+B)\alpha$

So $B = \dfrac{\gamma_1}{\alpha} - A = \left(\dfrac{2\alpha}{\alpha(1-\alpha^2)^3}\right)\sigma^2 - \dfrac{(1+\alpha^2)}{(1-\alpha^2)^3}\sigma^2 = \left(\dfrac{1-\alpha^2}{(1-\alpha^2)^3}\right)\sigma^2 = \dfrac{\sigma^2}{(1-\alpha^2)^2}$

[Alternatively, solve using the formula on page 4 of the Tables:

We have $g_k = 2\alpha g_{k-1} + \alpha^2 g_{k-2} = 0$

Using the Tables formula, the roots are $\lambda_1 = \lambda_2 = \alpha$ so we have a solution of the form $g_k = (A + Bk)\lambda^k = (A + Bk)\alpha^k$

Set $k = 0$ and $k = 1$ to get the same equations as before.]

Another good differentiator, with strong candidates scoring well, and weaker candidates struggling with parts (ii) and (iii) in particular.

END OF EXAMINERS' REPORT