INSTITUTE AND FACULTY OF ACTUARIES

EXAMINERS' REPORT

April 2014 examinations

Subject CT6 – Statistical Methods
Core Technical

Introduction

The Examiners' Report is written by the Principal Examiner with the aim of helping candidates, both those who are sitting the examination for the first time and using past papers as a revision aid and also those who have previously failed the subject.

The Examiners are charged by Council with examining the published syllabus. The Examiners have access to the Core Reading, which is designed to interpret the syllabus, and will generally base questions around it but are not required to examine the content of Core Reading specifically or exclusively.

For numerical questions the Examiners' preferred approach to the solution is reproduced in this report; other valid approaches are given appropriate credit. For essay-style questions, particularly the open-ended questions in the later subjects, the report may contain more points than the Examiners will expect from a solution that scores full marks.

The report is written based on the legislative and regulatory context pertaining to the date that the examination was set. Candidates should take into account the possibility that circumstances may have changed if using these reports for revision.

D C Bowie
Chairman of the Board of Examiners

June 2014

General comments on Subject CT6

The examiners for CT6 expect candidates to be familiar with basic statistical concepts from CT3 and so to be comfortable computing probabilities, means, variances etc for the standard statistical distributions. Candidates are also expected to be familiar with Bayes' Theorem, and be able to apply it to given situations. Many of the weaker candidates are not familiar with this material.

The examiners will accept valid approaches that are different from those shown in this report. In general, slightly different numerical answers can be obtained depending on the rounding of intermediate results, and these will still receive full credit. Numerically incorrect answers will usually still score some marks for method providing candidates set their working out clearly.

Comments on the April 2014 paper

The examiners felt that this paper was broadly in line with other recent papers. The quality of solutions was often good, with questions 2 and 8 providing the greatest challenge to most students.

1 (i)

- policyholder has an interest in the risk
- risk is of a financial nature and reasonably qualifiable
- independence of risks
- probability of event is relatively small
- pool large numbers of potentially similar risks
- ultimate limit on liability of insurer
- moral hazards eliminated as far as possible
- claim amount must bear some relationship to financial loss
- sufficient data to reasonably estimate extent of risk / likelihood of occurence

(ii)

- policy lasts for a fixed term
- policy lasts for a relatively short period of time
- policyholder pays a premium
- insurer pays claims that arise during the policy term
- option (but no obligation) to renew policy
- claim does not bring policy to an end

Other sensible points received full credit. This question was generally well answered.

2 A is better than B since Ruth has a capital buffer at the start of the week which can offset later journeys, whereas under B a high fare on Monday causes Ruth to run out of funds.

B and C are the same – the net funds available under C are always exactly ½ of those available under B.

So overall A gives the lowest probability of running out of cash.

Many candidates did not attempt this question which required a qualitative analysis of the situation set out. Those candidates who had a good understanding of the basic principles underlying the material on ruin theory were able to score well.

3 (i) D_1 dominates D_4 since D_1 gives a higher outcome for every state of nature.

D_1 has the best result under θ_1 and so is not dominated.

D_2 has the best result under θ_2 and so is not dominated.

Similarly D_3 has the best result under θ_3 and so is not dominated.

So only D_4 is dominated (by D_1).

(ii) The maximum losses are

D_1	–8
D_2	–3
D_3	–7
D_4	–10

The highest of these is –3 under D_2. Therefore D_2 is the optimal strategy under the minimax criterion

(iii) $E(D_1) = 10 \times 0.5 - 5 \times 0.3 - 8 \times 0.2 = 1.9$
$E(D_2) = 3 \times 0.5 + 12 \times 0.3 - 3 \times 0.2 = 4.5$
$E(D_3) = -7 \times 0.5 + 6 \times 0.3 + 13 \times 0.2 = 0.9$

So D_2 is the optimal decision under the Bayes criterion.

This question was well answered.

4 $E(X_i) = E(E(X_i|\alpha))$

$$= E\left(\frac{\alpha}{\lambda}\right) = \frac{1}{3}E(\alpha)$$

$$= \frac{1}{3}(0.7 \times 300 + 0.3 \times 600) = \frac{1}{3} \times 390$$

$$= 130$$

$\text{Var}(X_i) = \text{Var}(E(X_i|\alpha)) + E(\text{Var}(X_i|\alpha))$

$$= \text{Var}\left(\frac{\alpha}{\lambda}\right) + E\left(\frac{\alpha}{\lambda^2}\right)$$

$$= \frac{1}{\lambda^2}\text{Var}(\alpha) + \frac{1}{\lambda^2}E(\alpha)$$

$$= \frac{1}{9}(0.7 \times 300^2 + 0.3 \times 600^2 - 390^2) + \frac{1}{9} \times 390$$

$$= 2100 + \frac{390}{9} = 2143.33$$

so overall

$$E(S) = \lambda E(X) = 500 \times 130 = 65{,}000$$
$$\text{Var}(S) = \lambda E(X^2) = 500 \times (2143.33 + 130^2)$$

$$= 9{,}521{,}665$$

There are other approaches which can be taken to calculating the variance, all of which were given full credit. Whilst most candidates were able to calculate the mean only the better candidates were able to accurately calculate the variance.

5 Mean claim is $50 \times 0.3 + 100 \times 0.5 + 200 \times 0.2$

$$= 15 + 50 + 40$$

$$= 105$$

Also $E(X^2) = 50^2 \times 0.3 + 100^2 \times 0.5 + 200^2 \times 0.2$

$$= 13{,}750$$

so over 1 year the mean aggregate claim amount is

$$25 \times 105 = 2625$$

and the variance of aggregate claims is

$$25 \times 13{,}750 = 586.30^2$$

Using a Normal approximation we need to find θ such that

$$P(N(2625, 586.3^2) > 240 + 25 \times 105 \times (1 + \theta)) = 0.1$$

i.e. $P(N(2625, 586.3^2) > 240 + 2625(1 + \theta)) = 0.1$

i.e. $$P\left(N(0,1) > \frac{240 + 2625\theta}{586.3}\right) = 0.1$$

so $$\frac{240 + 2625\theta}{586.3} = 1.2816$$

i.e. $$\theta = \frac{1.2816 \times 586.3 - 240}{2625}$$

$$= 0.1948$$

This question was well answered with many candidates scoring well.

6 (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 $\quad e^{\sigma^2} = 1 + \dfrac{110^2}{230^2} = 1.22873$

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

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.

Most candidates were able to apply the method of moments in part (i) but many struggled to apply the method of percentiles. In particular, it was clear that many candidates could not relate the lognormal distribution back to the underlying normal distribution (this was also a common issue in Q8). Alternative comments on the data were given credit in part (ii).

7 (i) $P(\mu > 180) = P(N(187, 10^2) > 180)$

$$= P\left(N(0,1) > \frac{180-187}{10}\right)$$

$$= P(N(0,1) > -0.7)$$

$$= 0.75804$$

(ii) We know that $\mu|x \sim N(\mu_*, \sigma_*^2)$

$$\text{Where } \mu_* = \left(\frac{80\times 182}{15^2} + \frac{187}{10^2}\right)\bigg/\left(\frac{80}{15^2} + \frac{1}{10^2}\right) = 182.14$$

$$\text{And } \sigma_*^2 = \frac{1}{\frac{80}{15^2} + \frac{1}{10^2}} = 2.73556 = 1.6540^2$$

so $P(\mu > 180) = P(N(182.14, 1.654^2) > 180)$

$$= P\left(N(0,1) > \frac{180-182.14}{1.654}\right)$$

$$= P(N(0,1) > -1.29192)$$

$$= 0.38 \times 0.9032 + 0.62 \times 0.90147$$

$$= 0.90180$$

(iii) The probability has risen, reflecting our much greater certainty over the value of μ as a result of taking a large sample.

This is despite the fact that our mean belief about μ has fallen, which a priori might make a lower value of μ more likely.

The posterior distribution has thinner tails / lower volatility, since we have increased credibility around the mean

This question was mostly well answered. A small number of candidates were not aware that the formulae for the Normal / Normal model are given in the tables, and therefore struggled with the algebra required to derive the posterior distribution.

8 (i) (a) Let u_1 and u_2 be independent samples from a $U(0,1)$ distribution.

Then
$$Z_1 = \sqrt{-2\log u_1}\cos(2\pi u_2)$$
$$Z_2 = \sqrt{-2\log u_1}\sin(2\pi u_2)$$

are independent standard normal variables.

(b) Advantage – generates a sample of every pair of u_1 and u_2 – no possibility of rejection.

Disadvantage – requires calculation of sin and cos functions which is more computationally intensive.

(ii) Generate Z as in (i). Then

$$Y = \exp(\mu + \sigma Z)$$

is a sample from the required Lognormal distribution.

(iii) Set $X = 0$, $k = 0$

Step 1 generate a sample u from $U(0,1)$, set $k = k + 1$

Step 2 If $u \le p$ then go to step 3 else go to step 4

Step 3 Generate a sample Y from the Lognormal distribution in (ii) and set $X = X + Y$

Step 4 If $k = n$ finish else go to step 1

X represents aggregate claims on the portfolio.

(iv) The standard error will be approximately $\sqrt{\dfrac{\hat{\theta}(1-\hat{\theta})}{n}} = \sqrt{\dfrac{0.09}{n}}$.

We want $\sqrt{\dfrac{0.09}{n}} \times 1.96 < 0.01$

i.e. $\sqrt{n} > \dfrac{\sqrt{0.09} \times 1.96}{0.01} = 58.8$

i.e. $n > 3457.44$

so 3,458 simulations are needed.

(v) The insurer should use the same pseudo-random numbers so that any variation in simulation results is due to the impact of the reinsurance and not just due to random variation in the simulation process.

Parts (i) and (v) were well answered. The remaining parts were found by many candidates to be the hardest questions on the paper. In part (ii) many candidates could not relate the Lognormal distribution to the Normal distribution from which samples had been generated in (i). Only the best candidates attempted parts (iii) and (iv).

9 Incremental claims in mid 2013 prices are given by:

Accident year	*Development year* 0	1	2
2011	1543.3	550	142
2012	1766.17	811	
2013	1912		

Cumulative claims in mid 2013 prices:

Accident year	*Development year* 0	1	2
2011	1543.3	2093.3	2235.3
2012	1766.17	2577.17	
2013	1912		

$DF_{0,1}$ = (2093.3 + 2577.17) / (1543.3 + 1766.17) = 1.4112441

$DF_{1,2}$ = 2235.3 / 2093.3 = 1.067835

Completed cumulative claims

Accident year	*Development year* 0	1	2
2011			
2012			2751.99
2013		2698.30	2881.34

Incremental claims (mid 2013 prices)

Accident year	*Development year* 0	1	2
2011			
2012			174.82
2013		786.30	183.04

And projecting for inflation, outstanding claims = $(174.82 + 786.30) \times \frac{113}{110}$

$$+ 183.04 \times \frac{117}{110} = 1182.02$$

This question was well answered, with many candidates scoring full marks.

10 (i) The likelihood is given by

$$L = \theta_{ij}(1-\theta_{ij})^{y_{ij}}$$

Taking logs gives

$$l = \log L = \log\theta_{ij} + y_{ij}\log(1-\theta_{ij})$$

Differentiating with respect to θ_{ij} gives

$$\frac{\partial l}{\partial \theta_{ij}} = \frac{1}{\theta_{ij}} - \frac{y_{ij}}{(1-\theta_{ij})}$$

and setting $\frac{\partial l}{\partial \theta_{ij}} = 0$ we have

$$\frac{1}{\hat{\theta}_{ij}} = \frac{y_{ij}}{1-\hat{\theta}_{ij}}$$

so $\quad 1-\hat{\theta}_{ij} = y_{ij}\hat{\theta}_{ij}$

so $\quad 1 = (1+y_{ij})\hat{\theta}_{ij}$

i.e. $\quad \hat{\theta}_{ij} = \dfrac{1}{1+y_{ij}}$

and since $\dfrac{\partial^2 l}{\partial\theta_{ij}^2} = -\dfrac{1}{\theta_{ij}^2} - \dfrac{y_{ij}}{(1-\theta_{ij})^2} < 0$ (since $y_{ij} > 0$)

we do have a maximum.

(ii) $\quad P(Y_{ij} = y) = \theta_{ij}(1-\theta_{ij})^y$

$$= \exp[\log\theta_{ij} + y\log(1-\theta_{ij})]$$

$$= \exp[y\log(1-\theta_{ij})] + \log\theta_{ij}]$$

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

where $\theta = \log(1-\theta_{ij})$ is the natural parameter

$b(\theta) = -\log\theta_{ij} = -\log[1-e^{\theta}]$

$\varphi = 1 \qquad a(\varphi) = 1$

$c(y, \varphi) = 0$

(iii) The Pearson residuals are often skewed for non normal data which makes the interpretation of residual plots difficult.

Deviance residuals are usually more likely to be symmetrically distributed and are preferred for actuarial applications.

This question was, for the most part, answered well. A common mistake in part (i) was to try to sum across either years or policies when the question specifically referred to a single data point.

11 (i) We have $\mu = \dfrac{\alpha}{\alpha+\beta}$ so $\beta = \dfrac{\alpha(1-\mu)}{\mu}$

and $\sigma^2 = \dfrac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} = \mu \times \dfrac{\beta}{(\alpha+\beta)(\alpha+\beta+1)} = \dfrac{\mu(1-\mu)}{\alpha+\beta+1}$

substituting for $\alpha + \beta = \frac{\alpha}{\mu}$ gives

$$\sigma^2 = \frac{\mu(1-\mu)}{\frac{\alpha}{\mu}+1}$$

so $\sigma^2 = \dfrac{\mu^2(1-\mu)}{\alpha+\mu}$

so $\alpha = \dfrac{\mu^2(1-\mu)-\mu\sigma^2}{\sigma^2}$

and $\beta = \dfrac{(\mu^2(1-\mu)-\mu\sigma^2)(1-\mu)}{\mu\sigma^2} = \dfrac{(\mu(1-\mu)-\sigma^2)(1-\mu)}{\sigma^2}$

(ii) (a) $f(\theta|\underline{x}) \;\alpha\; f(\underline{x}|\theta)\, f(\theta)$

$$\alpha\; \theta^{d}(1-\theta)^{n-d}\, \theta^{\alpha-1}(1-\theta)^{\beta-1}$$

$$\alpha\; \theta^{\alpha+d-1}(1-\theta)^{n-d+\beta-1}$$

which is the pdf of a Beta distribution with parameters $\alpha + d$ and $\beta + n - d$.

(b) The posterior mean is given by

$$\frac{\alpha+d}{\alpha+d+\beta+n-d} = \frac{\alpha+d}{\alpha+\beta+n}$$

$$= \frac{\alpha}{\alpha+\beta}\times\frac{\alpha+\beta}{\alpha+\beta+n}+\frac{d}{n}\times\frac{n}{\alpha+\beta+n}$$

$$= \frac{\alpha}{\alpha+\beta}\times(1-Z)+\frac{d}{n}Z$$

where $Z = \dfrac{n}{\alpha+\beta+n}$.

This is in the form of a credibility estimate since $\frac{\alpha}{\alpha+\beta}$ is the mean of the prior distribution and $\frac{d}{n}$ is the MLE.

(iii) $Z = \frac{n}{\alpha+\beta+n}$ decreases as $\alpha + \beta$ increases and increases as $\alpha + \beta$ decreases.

From (i) $\alpha + \beta = \frac{\mu^2(1-\mu)-\mu\sigma^2+\mu(1-\mu)^2-(1-\mu)\sigma^2}{\sigma^2}$

$$= \frac{\mu(1-\mu)(\mu+1-\mu)-\mu\sigma^2-\sigma^2+\mu\sigma^2}{\sigma^2}$$

$$= \frac{\mu(1-\mu)}{\sigma^2}-1$$

so increasing σ^2 decreases $\alpha + \beta$ and increases Z.

(iv) Higher σ^2 implies less certainty in the prior estimate / prior is less reliable and so should lead to more weight on the observed data – which it does via a higher Z.

Only the best candidates were able to complete the algebra in part (i). Many candidates nevertheless scored well on parts (ii) and (iv).

12 (i) $\hat{\phi}_1 = \hat{\rho}_1 = \frac{0.68}{0.9} = 0.755556$

$$\hat{\phi}_2 = \frac{\hat{\rho}_2-\hat{\rho}_1^2}{1-\hat{\rho}_1^2} = \frac{\frac{0.55}{0.9}-0.755556^2}{1-0.755556^2} = 0.093786$$

(ii) (a) The Yule-Walker equations for this model give

$$\gamma_0 = a_1\gamma_1 + \sigma^2$$
$$\gamma_1 = a_1\gamma_0$$

so we have $\hat{a}_1 = \frac{\hat{\gamma}_1}{\hat{\gamma}_0} = \hat{\rho}_1 = 0.755556$

and $\hat{\sigma}^2 = \hat{\gamma}_0 - \hat{a}_1\hat{\gamma}_1 = \hat{\gamma}_0(1-\hat{a}_1\hat{\rho}_1)$

$= 0.9 - 0.755556 \times 0.68 = 0.38622$

Finally we let $\mu = E(Y_t)$ and observe that

$\mu = a_0 + a_1\mu$

so $\mu = \dfrac{a_0}{1-a_1}$

so $\hat{a}_0 = \hat{\mu}(1-\hat{a}_1) = 1.35(1-0.755556) = 0.33000$

(b) For this model the Yule-Walker equations are

$\gamma_0 = a_1\gamma_1 + a_2\gamma_2 + \sigma^2$ (1)
$\gamma_1 = a_1\gamma_0 + a_2\gamma_1$ (2)
$\gamma_2 = a_1\gamma_1 + a_2\gamma_0$ (3)

substituting the observed values in (2) and (3) gives

$0.68 = 0.9\hat{a}_1 + 0.68\hat{a}_2$ (4)
$0.55 = 0.68\hat{a}_1 + 0.9\hat{a}_2$ (5)

(4) $\times 0.68$ – (5) $\times 0.9 \Rightarrow 0.68^2 - 0.55\times 0.9 = \hat{a}_2\left(0.68^2 - 0.9^2\right)$

So $\hat{a}_2 = \dfrac{-0.0326}{-0.3476} = 0.09379$

And $\hat{a}_1 = \dfrac{0.68(1-0.09378596)}{0.9} = 0.68470$

substituting into (1)

$\hat{\sigma}^2 = 0.9 - 0.68470 \times 0.68 - 0.09379 \times 0.55 = 0.38283$

and finally setting $\mu = E(Y_t)$ we have $\mu = a_0 + a_1\mu + a_2\mu$

so $\mu = \dfrac{a_0}{1-a_1-a_2}$

so $\hat{a}_0 = \hat{\mu}\times(1-\hat{a}_1-\hat{a}_2) = 1.35(1-0.68470-0.09379) = 0.29905$

(iii) Stationarity is necessary for both models since the Yule-Walker equations do not hold without the existence of the auto-covariance function.

(iv) Model (a) does satisfy the Markov property since the current value depends only on the previous value.

This does not hold for Model (b).

Most candidates were able to derive the Yule-Walker equations and therefore scored marks on this question. Only the best candidates were able to use these equations to derive numerical values of the parameters. Part (iv) was generally well answered.

Although the question stated that the given values were for the auto-covariance function, many candidates calculated as if the given values came from the auto-correlation function. The Examiners noted that the core reading does use the abbreviation ACF for the auto-correlation function, and therefore gave full credit to candidates who interpreted the question in this way. The numerical values of the estimated parameters taking this approach are as follows:

(i) $\hat{\phi}_1 = 0.68$

$\hat{\phi}_2 = 0.1629$

(ii) (a) $\hat{a}_0 = 0.432$

$\hat{a}_1 = 0.68$

$\hat{\sigma}^2 = 0.48384$

(b) $\hat{a}_0 = 0.3617$

$\hat{a}_1 = 0.5692$

$\hat{a}_2 = 0.1629$

$\hat{\sigma}^2 = 0.4710$

END OF EXAMINERS' REPORT