INSTITUTE AND FACULTY OF ACTUARIES

EXAMINERS' REPORT

September 2015

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.

F Layton
Chairman of the Board of Examiners
December 2015

© Institute and Faculty of Actuaries

A. General comments on the *aims of this subject and how it is marked*

1. The aim of the Statistical Methods subject is to provide a further grounding in mathematical and statistical techniques of particular relevance to financial work.

2. Errors carried over are only penalised once.

3. Generally arithmetic errors are not treated as harshly as method errors.

4. Markers exercise judgement when answers are partly correct and can award partial marks if appropriate. In particular, where a candidate has not used the method in the marking schedule, but has shown some understanding by their working, some credit is given.

5. Errors just due to rounding are not penalised unless the rounding is excessive (e.g. rounding an interim step to just 2 sig fig, say) and significantly compromises accuracy.

B. General comments on *student performance in this diet of the examination*

Stronger candidates with a good all-round knowledge of the subject and an ability and willingness to apply theoretical concepts to practical applications scored well in this diet. Candidates who did not score well typically had gaps in their knowledge of the bookwork for the earlier questions and had difficulty in approaching and answering the later questions, which required a thorough understanding of the relevant topics.

C. Comparative pass rates for the past 3 years for this diet of examination

Year	*%*
September 2015	53
April 2015	46
September 2014	67
April 2014	59
September 2013	59
April 2013	64

Reasons for any significant change in pass rates in current diet to those in the past:

The pass rate for this session is very much in line with the recent average pass rate.

Solutions

Q1 $n > \dfrac{z_{\alpha/2}^2 \, \hat{\tau}^2}{\epsilon^2}$

$$n > 2.5758^2 * \frac{0.12}{0.01^2} = 7{,}961.87$$

so $n = 7{,}962$

> Many candidates scored well on this straightforward question, although a disappointing number of candidates were unfamiliar with the theory and so scored poorly.

Q2 (i) The model assumes that the mean and standard deviation of the claim amounts are known with certainty.

Model assumes that claims are settled as soon as the incident occurs, with no delays.

No allowance for expenses is made.

No allowance for interest.

(ii) Car insurance, contents insurance (or other similar examples)

> Part (i) was typically poorly answered, as the majority of candidates gave the characteristics of insurable risks in general, rather than focusing on short term contracts.
> Part (ii) was generally well answered.

Q3 (i) Because total profits are fixed, whatever one company makes the other can be thought of as having lost, and vice-versa.

(ii) A randomised strategy is where the player randomly chooses between different strategies, rather than adopting a fixed approach.

(iii) Raspberry Inc. will randomly choose the cautious approach with probability p, and the aggressive approach with probability $(1 - p)$.

Robots/Raspberry	*Cautious*	*Aggressive*
Cautious	700	400
Aggressive	200	700

In order to determine the optimal strategy we need to equate the payoffs:

$700p + 400(1-p) = 200p + 700(1-p)$
$800p = 300$
so $p = 3/8$

So Raspberry Inc. should adopt the Cautious approach 3/8 of the time.

Candidates familiar with zero-sum two person games were able to score very well on this question. Weaker candidates were unfamiliar with randomised strategies.

Q4 (i) A suitable distribution is $U(0,1)$ as theta must lie between 0 & 1.
$f(\theta) = 1$ for $0 \le \theta \le 1$

Binomial distribution so likelihood function
$L(\theta) = (30\ C\ 16)\ \theta^{16}\ (1-\theta)^{14}$

Bayes theorem: PDF (posterior) = PDF (prior) * likelihood

PDF (Posterior) proportional to $\theta^{16}\ (1-\theta)^{14}$

So distribution of theta|sample is Beta (17,15)

Under quadratic loss estimate of theta is mean so
$17/(17 + 15) = 0.53125$

(ii) prior PDF proportional to $\theta^{\alpha-1} * (1-\theta)^{\alpha-1}$

so using Bayes again posterior proportional to $\theta^{15+\alpha} * (1-\theta)^{13+\alpha}$

Under all-or-nothing we need the mode of the posterior

Take logs $(15 + \alpha) \log\theta + (13 + \alpha) \log (1 - \theta)$

differentiate $(15 + \alpha) / \theta - (13 + \alpha) / (1 - \theta) = 0$

so $(15 + \alpha) (1 - \theta) = (13 + \alpha) \theta$ so $\theta (2\alpha + 28) = 15 + \alpha$

so $\theta = (15 + \alpha) / (28 + 2\alpha)$

since $f(0) = f(1) = 0$ this must be a maximum

> Well prepared candidates had little difficulty with this straightforward question on Bayes' theory. Most candidates were able to pick up at least a few marks by showing knowledge of the basic theory.

Q5 (i) Given

$$f(\theta) \propto \exp - \frac{(\theta - \mu)^2}{2\sigma^2} \propto \exp - \frac{1}{2\sigma^2} (\theta^2 - 2\mu\theta)$$

and

$$p(\underline{x} \mid \theta) \propto \prod_{i=1}^{n} \exp - \frac{(x_i - \theta)^2}{2\tau^2} \propto \exp - \frac{1}{2\tau^2} \sum_{i=1}^{n} (\theta^2 - 2x_i\theta)$$

$$\propto \exp - \frac{1}{2\tau^2} (n\theta^2 - 2n\bar{x}\theta) \qquad \left(\sum_{i=1}^{n} x_i = n\bar{x} \right),$$

we want

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

$$\propto \exp\left[-\left\{ \left(\frac{1}{2\sigma^2} + \frac{n}{2\tau^2} \right) \theta^2 - \left(\frac{\mu}{\sigma^2} + \frac{n\bar{x}}{\tau^2} \right) \theta \right\} \right]$$

$$\propto \exp\left[-\frac{\tau^2 + n\sigma^2}{2\sigma^2\tau^2} \left\{ \theta^2 - 2\left(\frac{\mu\tau^2 + n\bar{x}\sigma^2}{\tau^2 + n\sigma^2} \right) \theta \right\} \right]$$

$$\propto \exp\left[-\frac{\tau^2 + n\sigma^2}{2\sigma^2\tau^2} \left(\theta - \frac{\mu\tau^2 + n\bar{x}\sigma^2}{\tau^2 + n\sigma^2} \right)^2 \right]$$

$$\Rightarrow \theta \mid \underline{x} \sim N\left(\frac{\tau^2}{\tau^2 + n\sigma^2}\mu + \frac{n\sigma^2}{\tau^2 + n\sigma^2}\bar{x}, \frac{\sigma^2\tau^2}{\tau^2 + n\sigma^2} \right)$$

(ii) The posterior mean (the point estimator under quadratic loss) is

$$E(\theta|\underline{x}) = \frac{\tau^2}{\tau^2 + n\sigma^2}\mu + \frac{n\sigma^2}{\tau^2 + n\sigma^2}\overline{x}$$

(iii) $E(\theta|\underline{x}) = (1 - Z)\mu + Z\overline{x}$

where

$$Z = \frac{n\sigma^2}{\tau^2 + n\sigma^2} = \frac{n}{n + \dfrac{\tau^2}{\sigma^2}}$$

is the credibility factor. Hence $E(\theta|\underline{x})$ can be expressed in the form of a credibility estimate.

> Again well prepared candidates who had learnt the relevant bookwork scored very well on this question. Some candidates attempted to "fudge" the result or only quoted the result, making no attempt to derive it, and hence scored poorly.

Q6 (i) The saturated model is one where the number of parameters is the same as the data points,
i.e. the fitted values are the same as the fitted data.

(ii) The scaled deviance is twice the difference between the log likelihood values between the model in consideration and the saturated model.

(iii) (a) Pearson residuals are $\dfrac{y-\hat{\mu}}{\sqrt{\text{var}(\hat{\mu})}}$ where $\hat{\mu}$ is the fitted response estimator.

The deviance residuals are $\text{sign}(y-\hat{\mu})d_i$ where d_i is the contribution of the i-th to the total deviances,
i.e. Σd_i^2 is the scaled deviance.

(b) The Pearson residuals tend to be skewed in non normal data
while the deviance residuals tend to be symmetric and hence the normal assumption is more appropriate.
For that reason the latter is preferred in actuarial applications.

(c) In the normal data, normal residuals these are identical.

Most candidates were able to score at least some of the marks here, but only the stronger candidates had sufficient recall and understanding of the full detail of the bookwork in order to score very well.

Q7 (i) $$Z_3 = \frac{\sum_{j=1}^{6} P_{3,j}}{\sum_{j=1}^{6} P_{3,j} + E\left(s^2(\theta)\right) / \mathrm{Var}\left(m(\theta)\right)}$$

$$\bar{P} = \sum_i \bar{P}_i = 2265$$

$n = 6$, $N = 3$

$$P^* = \frac{1}{(Nn-1)} \sum_{i=1}^{N} \bar{P}_i \left(1 - \bar{P}_i / \bar{P}\right)$$

= 1/17 * (648 * (1 – 648/2265) + 981 * (1 – 981/2265) + 636 * (1 – 636/2265))

= 86.831 944

$$E\left(s^2(\theta)\right) = \frac{1}{N} \sum_{i=1}^{N} \left\{ \frac{1}{n-1} \sum_{j=1}^{n} P_{i,j} \left(X_{i,j} - \bar{X}_i\right)^2 \right\}$$

= 1/15{30.966 692 + 4.689 264 + 62.449 512) = 6.540 364 5

$$\mathrm{Var}\left(m(\theta)\right) = \frac{1}{P^*} \left\{ \frac{1}{Nn-1} \sum_{i=1}^{N} \sum_{j=1}^{n} P_{i,j} \left(X_{i,j} - \bar{X}\right)^2 - E\left(s^2(\theta)\right) \right\}$$

= 1/86.831 944 * (1/17{64.392683 + 42.240804+66.467182} – 6.540 365)

= 0.041 943

so Z_3 = 636/(636 + 6.5403645/0.04194341) = 0.803 098

$X_{3,7}$ = 0.803 098 * 0.370 370 + (1 – 0.803098) * 0.297 572 = 0.356 036

So $Y_{3,7} = 100 * X_{3,7}$ = 35.60

(ii) Disadvantages of Model 1
Does not make use of the risk volumes
Requires more assumptions about the data

Advantages of Model 1
Requires less information (does not take account of risk volumes)
EBCT Model 1 is likely to be computationally more straightforward

> This was one of the best answered questions on the paper, with the majority of candidates able to score most or all of the marks in part (i). Only the stronger candidates were able to pick up both marks in part (ii), although again most candidates picked up some marks here.

Q8 Assume claims fully run off by the end of development year 3.
Each year develops in the same way
The weighted average past inflation is repeated
The loss ratio is appropriate

DF 2–3 = 2,310/2,212 = 1.044 304
DF 1–2 = (2,951 + 2,212) / (2,251 + 2,034) = 1.204 901
DF 0–1 = (2,034 + 2,251 + 1,851) / (1,528 + 1,812 + 1,693) = 1.219 154

Ultimate loss for 2014 = 0.91 * 4,023 = 3,660.93

Emerging liability = 3660.93 * (1 – 1/(1.044304 * 1.204901 * 1.219 154))
= 1,274.466

So total amount for policies written in 2014 is
2,125 + 1,274.466 = 3,399.466

so remaining

3,399.466 – 572 = 2,827.5

> This straightforward chain ladder question was the best answered question on the paper, although a disappointing number of candidates slipped up towards the end. Weaker candidates also failed to state the assumptions required.

Q9 (i) $M(t) = E[e^{tx}] = \int_0^\infty e^{tx} * \frac{1}{\Gamma(\alpha)} \lambda^\alpha x^{\alpha-1} e^{-\lambda x} dx$

$$= \int_0^\infty \frac{1}{\Gamma(\alpha)} \lambda^\alpha x^{\alpha-1} e^{-(\lambda-t)x} dx$$

$$= \frac{\lambda^\alpha}{(\lambda-t)^\alpha} \int_0^\infty \frac{1}{\Gamma(\alpha)} (\lambda-t)^\alpha x^{\alpha-1} e^{-(\lambda-t)x} dx$$

The integral is PDF of Γ (α, λ - t) $\Rightarrow$ = 1

$$\Rightarrow M(t) = \left(\frac{\lambda}{\lambda-t}\right)^\alpha = \left(1-\frac{t}{\lambda}\right)^{-\alpha}, t < \lambda.$$

(ii) Coefficient $= \frac{E[(x-\mu)^3]}{\left(\sqrt{\text{Var}(x)}\right)^3}$, given $E[x] - \mu$

$E[(x-\mu)^3] = E[x^3 - 3x^2\mu + 3x\mu^2 - \mu^3]$

$= E[x^3] - 3\mu\, E[x^2] + 3\mu^2\, E[x] - \mu^3$

$= E[x^3] - 3\mu\, E[x^2] + 2\mu^3$

$= E[x^3] - 3 * \frac{\alpha}{\lambda} * E[x^2] + 2(E[x])^3$

$M(t) = \left(1-\frac{t}{\lambda}\right)^{-\alpha}$ from (i)

$$M'(t) = -\alpha * \left(-\frac{1}{\lambda}\right)^1 * \left(1-\frac{t}{\lambda}\right)^{-\alpha-1} = \frac{\alpha}{\lambda} * \left(1-\frac{t}{\lambda}\right)^{-\alpha-1}$$

$$M''(t) = \frac{\alpha}{\lambda} * [-(\alpha+1)] * \left(-\frac{1}{\lambda}\right)^1 * \left(1-\frac{t}{\lambda}\right)^{-\alpha-2}$$

$$= \frac{\alpha * (\alpha+1)}{\lambda^2} * \left(1-\frac{t}{\lambda}\right)^{-\alpha-2}$$

$$M'''(t) \;=\; \frac{\alpha*(\alpha+1)}{\lambda^2} * [-(\alpha+2)] * \left(-\frac{1}{\lambda}\right)^1 * \left(1-\frac{t}{\lambda}\right)^{-\alpha-3}$$

$$= \frac{\alpha*(\alpha+1)*(\alpha+2)}{\lambda^3} * \left(1-\frac{t}{\lambda}\right)^{-\alpha-3}$$

$$E[x^3] \;=\; M'''(0) = \frac{\alpha*(\alpha+1)*(\alpha+2)}{\lambda^3}$$

$$E[x^2] \;=\; M''(0) \;=\; \frac{\alpha(\alpha+1)}{\lambda^2}$$

So

$$E[(x-\mu)^3] \;=\; \frac{\alpha*(\alpha+1)*(\alpha+2)}{\lambda^3} - \frac{3\alpha}{\lambda} * \frac{\alpha(\alpha+1)}{\lambda^2} + 2*\left(\frac{\alpha}{\lambda}\right)^3$$

$$= \frac{1}{\lambda^3}\left(\alpha^3 + 3\alpha^2 + 2\alpha - 3\alpha^3 - 3\alpha^2 + 2\alpha^3\right) = \frac{2\alpha}{\lambda^3}$$

$$\text{Coefficient} = \frac{\frac{2\alpha}{\lambda^3}}{\left(\sqrt{\frac{\alpha}{\lambda^2}}\right)^3} \;=\; \frac{2\alpha}{\alpha^{3/2}} \;=\; \frac{2}{\sqrt{\alpha}}$$

Most candidates were able to score well on part (i), but many candidates struggled with part (ii), particularly those who had forgotten or were unfamiliar with CT3 concepts. Many stronger candidates made use of the cumulant generating function in part (ii), which simplified the algebra considerably, and were awarded full credit.

Q10 (i) $\lambda + cR = \lambda M_X(R)$

$$c = (1+\vartheta)\lambda E(X) = 1.12\lambda * 25 = 28\,\lambda$$

$$M_X(R) = \int_0^{50} e^{Rx}\, 0.02\, dx = \frac{\left(e^{50R}-1\right)}{50R}$$

So (dividing by lambda)

$$1+28R-\frac{\left(e^{50R}-1\right)}{50R}=0$$

At R = 0.00665, fn = -0.000 115
At R = 0.00655 fn = 0.000 209

(ii) (a) $$E(Z)=\int_M^{50}(x-M)0.02dx=0.02\frac{x^2}{2}-Mx_M^{50}=25-M+0.01M^2$$

(b) $$c_{\text{net}}=(1+\vartheta)\lambda E(X)-1.15\lambda E(Z)$$
$$=28\lambda-1.15\lambda\left(25-M+0.01M^2\right)$$
$$=\lambda\left(-0.75+1.15M-0.0115M^2\right)$$

net claims = $$\lambda E(X)-\lambda E(Z)=25\lambda-\lambda\left(25-M+0.01M^2\right)$$
$$=\lambda\left(M-0.01M^2\right)$$

Need income > claims so

$-0.75+1.15\text{M}-0.0115M^2>M-0.01M^2$

$-0.0015M^2+0.15M-0.75>0$

$M>5.279$

(iii) It decreases M_{min} since reinsurance is less of a drag.

Again most candidates were able to score well on part (i), but only stronger candidates were able to apply reinsurance theory to score well on parts (ii) and (iii).

Q11 (i) (a) It follows that

$$\begin{pmatrix} 1 & -\beta \\ -\beta & 1 \end{pmatrix}\begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix}\begin{pmatrix} X_{t-1} \\ Y_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_t^1 \\ \varepsilon_t^2 \end{pmatrix}$$

(b) Multiplying both sides by

$$\begin{pmatrix} 1 & -\beta \\ -\beta & 1 \end{pmatrix}^{-1} = \frac{1}{(1-\beta^2)}\begin{pmatrix} 1 & \beta \\ \beta & 1 \end{pmatrix}$$

we then have

$$\begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \frac{1}{2\left(1-\beta^2\right)}\begin{pmatrix} 1 & \beta \\ \beta & 1 \end{pmatrix}\begin{pmatrix} X_{t-1} \\ Y_{t-1} \end{pmatrix} + \frac{1}{\left(1-\beta^2\right)}\begin{pmatrix} 1 & \beta \\ \beta & 1 \end{pmatrix}\begin{pmatrix} \epsilon_t^1 \\ \epsilon_t^2 \end{pmatrix}.$$

Which is a stationary VAR(1) model if the eigenvalues of

$$\mathbf{A_1} = \frac{1}{2(1-\beta^2)}\begin{pmatrix} 1 & \beta \\ \beta & 1 \end{pmatrix}$$

are those λ such that

$$\det\begin{pmatrix} 1-\lambda & \beta \\ \beta & 1-\lambda \end{pmatrix} = 0 \text{ or } \lambda_{1,2} = 1 \pm \beta$$

then the eigenvalues of A_1 are less than one in absolute value if

$$\left|\frac{1 \pm \beta}{2(1-\beta^2)}\right| < 1 \text{ i.e.}$$

$$\left|\frac{1}{2(1-\beta)}\right| < 1$$

and

$$\left|\frac{1}{2(1+\beta)}\right| < 1$$

which implies that $|\beta| < \frac{1}{2}$ or $|\beta| > \frac{3}{2}$

(ii) Here we have a VAR(2) where

$$\boldsymbol{A_1} = \begin{pmatrix} \alpha & \alpha \\ \beta & 0 \end{pmatrix} \qquad \boldsymbol{A_2} = \begin{pmatrix} 0 & 0 \\ -\beta & 0 \end{pmatrix}$$

since

$$\begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{pmatrix} \alpha & \alpha \\ \beta & 0 \end{pmatrix}\begin{pmatrix} X_{t-1} \\ Y_{t-1} \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ -\beta & 0 \end{pmatrix}\begin{pmatrix} X_{t-2} \\ Y_{t-2} \end{pmatrix} + \begin{pmatrix} \epsilon_t^1 \\ \epsilon_t^2 \end{pmatrix}.$$

This was by far the poorest answered question on the paper, with many candidates scoring few marks. Only the strongest candidates were able to derive the required eigenvalues and score well.

Q12 (i) Here $k^{-1} = \int_0^2 e^{-\frac{x}{2}} dx + \int_2^\infty e^{-x} dx = -2\ e^{-(\frac{x}{2})}\Big]_2^\infty = 2(1 - e^{-1}) + e^{-2} - 0$

$= 1.399576$

So $k = 1.399576^{-1} = 0.714502$

(ii) The distribution function here is

$$= \begin{cases} k\int_0^x e^{-\frac{u}{2}} du = 2k\left(1 - e^{-\frac{x}{2}}\right), & 0 < x < 2 \\ k\left(2\left(1-e^{-1}\right) + \int_2^x e^{-u} du\right) = k\left(2\left(1-e^{-1}\right) + e^{-2} - e^{-x}\right), & x \geq 2 \end{cases}$$

At $x = 2$, $(2) = 2 * k * (1 - \exp(-1)) = 0.9033029$, therefore the inversion function is

$$X = F^{-1}(U) = \begin{cases} -2\ln\left(1 - \dfrac{U}{2k}\right), & 0 < U < 0.9033029 \\ -\ln\left(2(1-e^{-1}) + e^{-2} - \dfrac{U}{k}\right), & U \geq 0.9033029 \end{cases}$$

The inversion algorithm is then:

Sample U from $U(0,1)$

Take $X = F^{-1}(U)$ as above.

(iii) $$M = \max_{x>0} \frac{f(x)}{e^{-x}} = \max_{x>0} \frac{f(x)}{e^{-x}} \begin{cases} ke^{+\frac{x}{2}}, & 0 < x < 2 = ke^1 \\ k, & x \geq 2 \end{cases}$$

The rejection function is then

$$h(x) = \frac{f(x)}{Me^{-x}} = \begin{cases} e^{-1}e^{\frac{x}{2}} & ,0 < x < 2 \\ e^{-1}, & x \geq 2 \end{cases}$$

The algorithm is then:

1 – Simulate U_1 from $U(0,1)$, so that $Y = -\log U_1$ is Exp(1)

2 – Simulate U_2 from $U(0,1)$

If $U_2 < h(Y)$ take $X = Y$ otherwise start again.

Most candidates were able to score well in part (i), and also at least partially in part (iii). Only the better prepared candidates were able to apply the inversion theory to part (ii), and only the strongest candidates correctly included the term $2(1 - e^{-1})$ for the second half of the distribution.

END OF EXAMINERS' REPORT