INSTITUTE AND FACULTY OF ACTUARIES

EXAMINERS' REPORT

September 2019 Examinations

Subject CS1 – Actuarial Statistics Core Principles (Part A)

Introduction

The Examiners' Report is written by the Chief 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.

Mike Hammer
Chair of the Board of Examiners
September 2019

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

1. The aim of the Actuarial Statistics subject is to provide a grounding in mathematical and statistical techniques that are of particular relevance to actuarial work.
2. Some of the questions in the examination paper admit alternative solutions from these presented in this report, or different ways in which the provided answer can be determined. All mathematically correct and valid alternative solutions or answers received credit as appropriate.
3. Rounding errors were not penalised, but candidates lost marks where excessive rounding led to significantly different answers.
4. In cases where the same error was carried forward to later parts of the answer, candidates were given full credit for the later parts.
5. In questions where comments were required, valid comments that were different from those provided in the solutions also received full credit where appropriate.

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

1. Performance was satisfactory in general, but varied considerably among candidates. Well prepared candidates were able to score highly.
2. This is a relatively new subject under the recently introduced curriculum, and combines a number of topics from previous CT subjects (CT3 and CT6). A number of candidates appeared to be inadequately prepared, in terms of not having covered sufficiently the entire breadth of the subject.

C. Pass Mark

The combined pass mark for CS1 in this exam diet was 55.

Solutions Subject CS1 – A

Q1

If X is the number of people who have at least two investments, X follows a binomial (300, 0.4) distribution and:

$$E[X] = 300 \times 0.4 = 120 \text{ and } V[X] = 72.$$

[1]

Then, using continuity correction, [½]

$$P(X > 100) = P(X \geq 100.5) = 1 - \Phi\left(\frac{100.5 - 120}{\sqrt{72}}\right) = 1 - \Phi(-2.298) = \Phi(2.298)$$
$$= 0.989$$

[1.5]

[Total 3]

The question was answered well by most candidates. Attention should be given to applying the continuity correction properly.

Q2

(i) $E(\bar{X}) = E\left(\frac{\sum_{i=1}^{n} X_i}{n}\right) = \frac{\sum_{i=1}^{n} E(X_i)}{n} = \frac{n\mu}{n} = \mu$ [1]

(ii) $V(\bar{X}) = V\left(\frac{\sum_{i=1}^{n} X_i}{n}\right) = \frac{\sum_{i=1}^{n} V(X_i)}{n^2}$ because of independence [1]

$= \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$ [1]

(iii) The variance of the sample mean is smaller compared to the variance of individual variables. [1]

(iv) Individual values are less precise than the average of a sample. [1]

Larger sample leads to smaller variance. [1]

[Total 6]

Parts (i)-(iii) were answered very well. In part (ii), independence must be mentioned for a fully justified derivation. Part (iv) was not well answered, with many answers being vague.

Q3

(i) $E[s^2(\theta)]$ is estimated by the average of the sample variances, therefore:

$$\frac{3{,}959{,}980+7{,}543{,}626+3{,}151{,}286}{3}=4{,}884{,}964$$

[½]

The sample mean of the $\bar{X}_i$'s is:

$$\bar{X}=\frac{2{,}109+6{,}152+3{,}016}{3}=3{,}759$$

[½]

And the sample variance of the $\bar{X}_i$'s is:

$$\frac{1}{3-1}\sum_{i=1}^{4}(\bar{X}_i-\bar{X})=0.5\ \times((2{,}109-3{,}759)^2+(6{,}152-3{,}759)^2+(3{,}01-3{,}759)^2)$$

$= 4{,}500{,}499$ [1]

So $Var[m(\theta)]$ is estimated by:

$$\frac{1}{2}\sum(\bar{X}_i-\bar{X})^2-\frac{1}{4}E[s^2(\theta)]\ =4{,}500{,}499-\frac{1}{4}\text{ x }4{,}884{,}964=3{,}279{,}258$$

[1]

The credibility factor,

$$Z=\frac{n}{n+\frac{E[s^2(\theta)]}{Var[m(\theta)]}}$$

is then estimated by:

$$Z=\frac{4}{4+\frac{4{,}884{,}964}{3{,}279{,}258}}=0.72864$$

[1]

(ii) Z is an increasing function of *n*, the number of years of past data. If we have more than 4 years of past data, the credibility factor will increase. [1]

Z is a decreasing function of $[s^2(\theta)]$. If $E[s^2(\theta)]$ increases, e.g. if the variance of the claim amounts from one or more of the risks were to increase, then the value of the credibility factor will fall. [1]

[Total 6]

Answers in part (i) were satisfactory, with a small number of calculation or arithmetic errors. A common error in part (ii) was trying to justify that credibility increases as variance increases.

Q4

(i) **(a)** E(Y | X = 1)

$= \sum_y y\, P(Y = y \mid X = 1)$ [½]

$= \sum_y y \frac{P(Y = y,\ X = 1)}{P(X = 1)}$ [½]

$= (-1 \times \frac{0.03}{0.46}) + (3 \times \frac{0.11}{0.46}) + (4.5 \times \frac{0.2}{0.46})$

= 2.6087 [1]

(b) $\text{Var}(X \mid Y = 3) = E(X^2 \mid Y = 3) - (E(X \mid Y = 3))^2$ [1]

$= (1 \times \frac{0.11}{0.28}) + (9 \times \frac{0.06}{0.28}) - ((1 \times \frac{0.11}{0.28}) + (3 \times \frac{0.06}{0.28}))^2$ [1]

$= 2.3214 - (1.0357)^2$
$= 1.2487$ [1]

(ii) Summing columns gives:

$P(X = 0) = 0.26,\ P(X = 1) = 0.46,\ P(X = 3) = 0.28$ [1]

Summing rows gives:

$P(Y = -1) = 0.11,\ P(Y = 0) = 0.35,\ P(Y = 3) = 0.28,$
$P(Y = 4.5) = 0.26$ [1]

(iii) Show that this result $P(X = x, Y = y) = P(X = x)\, P(Y = y)$ does not hold for one pair, for example: [1]

$P(X = 0, Y = -1) = 0.08 \neq \text{P}(X = 0) \times P(Y = -1)$

Correct conclusion that X and Y are NOT independent. [1]

[Total 9]

The question was reasonably well attempted. A common error in part (i) was not applying the expectation correctly for a conditional probability, e.g. by missing the division element.

Q5

(i) $L(\mu_{ij};\ y_{ij}) = \prod_{i=1}^{n}\prod_{j=1}^{m}\frac{\mu_{ij}{}^{y_{ij}}e^{-\mu_{ij}}}{y_{ij}!}$

$$l(\mu_{ij};\ y_{ij}) = \log\left(L(\mu_{ij};\ y_{ij})\right) = \sum_{i=1}^{n}\sum_{j=1}^{m}\left(y_{ij}\log(\mu_{ij}) - \mu_{ij} - \log(y_{ij}!)\right)$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{m}\left(y_{ij}\beta_i - e^{\beta_i} - \log(y_{ij}!)\right)$$ [2]

$$\frac{d}{d\beta_i}l(\mu_{ij};\ y_{ij}) = \sum_{j=1}^{m}y_{ij} - me^{\beta_i}$$ [1]

And,

$$\frac{d}{d\beta_i}l(\mu_{ij};\ y_{ij}) = 0 \Rightarrow e^{\hat{\beta}_i} = \sum_{j=1}^{m}\frac{y_{ij}}{m} \Rightarrow \hat{\beta}_i = (\bar{y}_i)$$ [1]

where:

$$\bar{y}_i = \sum_{j=1}^{m}\frac{y_{ij}}{m}$$

(ii) For the deviance we have:

$$l_s = \sum_{i=1}^{n}\sum_{j=1}^{m}\left(y_{ij}\log(y_{ij}) - y_{ij} - \log(y_{ij}!)\right)$$ [1]

$$l_c = \sum_{i=1}^{n}\sum_{j=1}^{m}\left(y_{ij}\log(\bar{y}_i) - \bar{y}_i - \log(y_{ij}!)\right)$$ [1]

$$\begin{aligned} D &= 2(l_s - l_c) \\ &= 2\left\{\sum_{i=1}^{n}\sum_{j=1}^{m}\left(y_{ij}\log(y_{ij}) - y_{ij} - \log(y_{ij}!)\right) - \sum_{i=1}^{n}\sum_{j=1}^{m}\left(y_{ij}\log(\bar{y}_i) - \bar{y}_i - \log(y_{ij}!)\right)\right\} \\ &= 2\sum_{i=1}^{n}\sum_{j=1}^{m}\{y_{ij}\log\frac{y_{ij}}{\bar{y}_i} - (y_{ij} - \bar{y}_i)\} \end{aligned}$$

where:

$$\bar{y}_i = \sum_{j=1}^{m}\frac{y_{ij}}{m}$$

[2]

(iii) In this case we have:

$y_{ij} = 7, \bar{y}_i = 18.95$ [1]

$D_{ij} = 2\left\{7\log\left(\frac{7}{18.95}\right) - (7 - 18.95)\right\} = 9.957$ [1]

[Total 10]

Answers to this question were weak in general. The question concerns the MLE and deviance of a simplified Poisson GLM. Part (iii) requires a calculation by inserting numerical values in a given expression.

Q6

(i) The regression slope suggests a positive relationship between the two variables, while the correlation coefficient shows a strong negative relationship. [2]

(ii) The histogram suggests a non-symmetric distribution for the residuals [1]
Non-symmetric about zero. [1]

(iii) $\hat{\beta} = \frac{S_{xy}}{S_{xx}} = \frac{-331.05}{82.5} = -4.013$ [1]

$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} = -19.124 + 4.013\left(\frac{45}{10}\right) = -1.067$ to 4 s.f. [1]

Line given as: $\hat{y} = -1.066 - 4.013x$ [1]

(iv) Predicted value is: $\hat{y} = -1.066 - 4.013 * 11 = -45.207$ [½]

$$\hat{\sigma}^2 = \frac{\left(S_{yy} - \frac{{S_{xy}}^2}{S_{xx}}\right)}{n-2} = \frac{\left(1329.523 - \frac{(-331.05)^2}{82.5}\right)}{8} = 0.1387045$$

[1]

$V(\hat{y}) = \left(\frac{1}{n} + \frac{(xnew - \bar{x})^2}{S_{xx}}\right) \times \hat{\sigma}^2 = \left(\frac{1}{10} + \frac{(11 - \frac{45}{10})^2}{82.5}\right) \times 0.1387595 = 0.08490399$ [1.5]

And $t_{8,0.025} = 2.306$ [½]

95% CI for mean $\hat{y}$ is given by: $-45.207 \pm 2.306 \times (0.08490399)^{1/2}$

i.e. $(-45.879, -44.535)$. [1.5]

(v) The width of the interval is only affected by $V(\hat{y})$, which depends on the new x value through the term $(x_{new} - \bar{x})^2$. This term will now be smaller as the new $x_{new} = 3.5$ value is closer to $\bar{x}$ than $x = 11$. Therefore the interval will be narrower.

[2]

[Total 14]

The question was reasonably well answered by most candidates. In part (i) many candidates provided a reasonable algebraic argument using known formulae. In part (iv) a common issue was using a normal or chi-squared pivotal quantity. In part (v) most candidates identified correctly the impact on the interval. However, note that an appropriate explanation of why the interval is narrower is required here.

Q7

(i) Based on data from two years (*y*), the likelihood is:

$f(y|\theta) \propto \theta^{39}(1-\theta)^{261}\theta^{60}(1-\theta)^{240+x} = \theta^{99}(1-\theta)^{501+x}$ [2]

Prior for θ is: $f(\theta) \propto \theta^2(1-\theta)^4$ [1]

So the posterior density is given by:

$f(\theta|y) \propto f(y|\theta) \times f(\theta) = \theta^{101}(1-\theta)^{505+x}$ [2]

which is the density of a Beta(102, 506 + *x*) distribution. [2]

(ii) The Bayesian estimate under quadratic loss is the posterior mean, so

$\hat{\theta} = E(\theta|y) = \frac{102}{102+506+x} = \frac{102}{608+x}$ [2]

(iii) The Bayesian estimate under all-or-nothing loss is the posterior mode, so we now need to maximise the posterior density:

$\frac{d}{d\theta} f(\theta|y) = 101\,\theta^{100}(1-\theta)^{505+x} - \theta^{101}(505+x)(1-\theta)^{504+x}$

$= \theta^{100}(1-\theta)^{504+x}[101(1-\theta)-(505+x)\theta]$ [2]

$\frac{d}{d\theta} f(\theta|y) = 0 \Rightarrow 101\left(1-\tilde{\theta}\right) - (505+x)\tilde{\theta} \Rightarrow \tilde{\theta} = \frac{101}{606+x}$ [2]

(iv) In this case, for $\hat{\theta} = \tilde{\theta}$ we need:

$$\frac{102}{608+x} = \frac{101}{606+x} \Rightarrow 102x + 61812 - 101x - 61408 = 0 \Rightarrow x = -404$$

[1]

This means that the number of apartments in year 2 would be 300 – 404 = – 104 which is not possible. [1]

[Total 15]

The question was not well answered overall. Parts (i) and (ii) were reasonably well attempted. In part (iii) some candidates worked with the logarithm of the posterior density, which is a valid alternative way to answer the question. The justification in part (iv) was poor in many cases.

Q8

(i) Each house must have the same probability of being burgled. [1]

Whether a house is burgled or not is independent of other houses being burgled. [1]

(ii) $L(p) = [P(x=0)]^{39}[P(x=1)]^{36}[P(x=2)]^{19}[P(x=3)]^{4}[P(x=4)][P(x=5)]$ [1]

Using a *Bin*(6 , p) distribution to calculate the probabilities:

$$L(p) = c\ \times [(1-p)^6]^{39}[p(1-p)^5]^{36}[p^2(1-p)^4]^{19}[p^3(1-p)^3]^4[p^4(1-p)^2]p^5(1-p)$$

$$= c\ \times p^{95}(1-p)^{505}$$ [1]

$$\log L(p) = \log c + 95\ \log p + 505 \log(1-p)$$

$$\frac{\partial}{\partial p}\log L(p) = \frac{95}{p} - \frac{505}{1-p}$$ [1]

Setting the differential equal to zero to obtain the maximum:

$$\frac{95}{\hat{p}} - \frac{505}{1-\hat{p}} = 0 \ \rightarrow\ 95(1-\hat{p}) - 505\hat{p} = 0$$

$$\hat{p} = \frac{95}{95+505} = 0.158333$$ [1]

Checking it's a maximum:

$$\frac{\partial^2}{\partial p^2}\log L(p) = -\frac{95}{p^2} - \frac{505}{(1-p)^2} < 0\ \rightarrow\ Max$$ [1]

(iii) Using the estimate of $\hat{p}$ we get the frequencies of 35.55 , 40.13 , 18.87 , 4.73 , 0.67 , 0.05 , 0.0 , using $P(X=x) = \binom{6}{x}\hat{p}^x(1-\hat{p})^{6-x}$. [2]

(iv) These are fairly similar to the observed frequencies – implying that it is a good fit. [1]

(v) Using $p = 0.13$ and $P(X = x) = \binom{6}{x} 0.13^x (0.87)^{6-x}$ we get

	0	1	2	3	4	5	6
Observed	39	36	19	4	1	1	0
Expected	43.36	38.88	14.52	2.89	0.32	0.02	0.00

[2]

Since the expected frequencies are less than 5 for 3, 4, 5 and 6 houses burgled, we need to combine these columns:

	0	1	2	3+
Observed	39	36	19	6
Expected	43.36	38.88	14.52	3.23

[1]

Calculating the statistic:

$\chi^2 = \frac{(39-43.36)^2}{43.36} + \cdots + \frac{(6-3.23)^2}{3.23} = 4.409516$ [1]

There are now 4 groups so the number of degrees of freedom is 4 – 1 = 3. No further reduction is made for p, as this was given rather than estimated. [1]

Carry out a one-sided test. The observed value of the test statistic is less than the 5% critical value of 7.815. [1]

So there is insufficient evidence to reject H_0 at the 5% level. Therefore it is reasonable to conclude that the model is a good fit. [1]

[Total 17]

The question was generally not well answered. In part (i) many candidates failed to give the standard assumptions required for using a binomial distribution. Answers to part (ii) were generally satisfactory. A number of candidates did not attempt parts (iii) and (iv), while many candidates that attempted them failed to derive the frequencies correctly. Part (v) concerns a standard goodness of fit chi-squared test, which many candidates failed to apply correctly. Note that in part (v), an alternative valid answer can be provided by combining the 2 and 3+ groups. This gives a different value for the statistic, but the same conclusion.

Q9

(i) $M_X(t) = (1 - \theta t)^{-\alpha}$

(a) $M_X'(t) = \alpha\theta(1 - \theta t)^{-\alpha-1}$, hence, $E[X] = M_X'(0) = \alpha\theta$ [1]

(b) $M_X''(t) = \alpha(\alpha+1)\theta^2(1-\theta t)^{-\alpha-2}$, therefore, $E[X^2] = M_X''(0) = \alpha(\alpha+1)\theta^2$ [1]

(c) $M_X'''(t) = \alpha(\alpha+1)(\alpha+2)\theta^3(1-\theta t)^{-\alpha-3}$
implies $E[X^3] = M_X'''(0) = \alpha(\alpha+1)(\alpha+2)\theta^3$. [1]

(ii) (a) With $\alpha = 4$, we have:

$$E[X] = 4\theta, \qquad E[X^2] = 20\theta^2, \qquad E[X^3] = 120\theta^3$$

Hence:

$$\sigma^2 = E[X^2] - (E[X])^2 = 20\theta^2 - (4\theta)^2 = 4\theta^2$$

[2]

(b)

$$\mu_3 = E[X^3] - 3\mu E[X^2] + 2\mu^3 = 120\theta^3 - 3(4\theta)(20\theta^2) + 2(4\theta)^3 = 8\theta^3$$

Coefficient of skewness $= \frac{\mu_3}{\sigma^3} = \frac{8\theta^3}{(2\theta)^3} = 1$ [2]

(iii)

$$L(\theta) = \prod_{i=1}^{n} \frac{x_i^3}{6\theta^4} e^{-\frac{x_i}{\theta}} = \frac{\prod_{i=1}^{n} x_i^3}{6^n\theta^{4n}} e^{-\frac{\sum_{i=1}^{n} x_i}{\theta}}$$

$\log L(\theta) = \log(\prod_{i=1}^{n} x_i^3) - n\log(6) - 4n\log(\theta) - \frac{\sum_{i=1}^{n} x_i}{\theta}$ [1]

$\frac{\partial}{\partial\theta}\log L(\theta) = \frac{-4n}{\theta} + \frac{\sum_{i=1}^{n} x_i}{\theta^2} = 0$ [1]

Solving this equation leads to:

$$\hat{\theta} = \frac{\sum_{i=1}^{n} X_i}{4n} = \frac{\bar{X}}{4}$$ [1]

and this is maximum since:

$$\left.\frac{\partial^2}{\partial\theta^2}\log L(\theta)\right|_{\hat{\theta}} < 0$$

(iv) $E[\hat{\theta}] = \frac{1}{4}E[\bar{X}] = \frac{1}{4}E[X] = \frac{1}{4}4\theta = \theta$, hence, $\hat{\theta}$ is unbiased. [1]

(v) $\bar{X} = \frac{796.2}{100} = 7.962$, implies $\hat{\theta} = \frac{7.962}{4} = 1.9905$ [1]

(vi) (a) $s^2 = \frac{1}{99}\left(8189.4 - \frac{796.2^2}{100}\right) = 18.69$ [1]

(b) $\sigma^2 = 4\theta^2$ and $4\hat{\theta}^2 = 15.848$, s^2 is a bit larger than the variance at $\hat{\theta}$. [1]

(vii) Sample coefficient 1.12 is close to the distribution value 1. [1]

(viii) Approximate 95% CI for μ is $\bar{x} \pm 1.96\sqrt{\frac{s^2}{n}}$ [1]

Since $\mu = 4\theta$, we obtain an approximate 95% CI for θ:

$$\frac{1}{4}\left(\bar{x} \pm 1.96\sqrt{\frac{s^2}{n}}\right) \qquad [1]$$

We obtain:

$$\frac{1}{4}\left(7.962 \pm 1.96\sqrt{\frac{18.69}{100}}\right) \text{ i.e. } (1.779, 2.202) \qquad [1]$$

(ix) (a)
The lower limit of the variance is $4 \times 1.779^2 = 12.66$ and the upper limit is $4 \times 2.202^2 = 19.40$. [1]

(b)
The value s^2 falls within these values, confirming that s^2 is close to $4\hat{\theta}^2$. [1]

[Total 20]

Performance on this question was mixed. Part (i) was generally well answered – some candidates attempted to derive the MGF which is not required here. In part (ii) there were several algebraic errors. Parts (iii) and (iv) were well attempted, while answers in parts (v)-(vii) were generally weak for those candidates that attempted them. There were some reasonable attempts in part (viii). Many candidates failed to scale the CI down by a quarter, while another common error was not using the sample standard deviation. Note that a valid alternative answer can be given in part (viii) with the use of asymptotic normality and the Cramér-Rao lower bound for the variance. Part (ix) was poorly answered.

END OF EXAMINERS' REPORT