



Institute
and Faculty
of Actuaries

EXAMINERS' REPORT

CS1A – Actuarial Statistics

Core Principles

Paper A

April 2023

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.

For some candidates, this may be their first attempt at answering an examination using open books and online. The Examiners expect all candidates to have a good level of knowledge and understanding of the topics and therefore candidates should not be overly dependent on open book materials. In our experience, candidates that spend too long researching answers in their materials will not be successful either because of time management issues or because they do not properly answer the questions.

Many candidates rely on past exam papers and examiner reports. Great caution must be exercised in doing so because each exam question is unique. As with all professional examinations, it is insufficient to repeat points of principle, formula or other text book works. The examinations are designed to test "higher order" thinking including candidates' ability to apply their knowledge to the facts presented in detail, synthesise and analyse their findings, and present conclusions or advice. Successful candidates concentrate on answering the questions asked rather than repeating their knowledge without application.

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.

Sarah Hutchinson
Chair of the Board of Examiners
July 2023

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

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.

Some of the questions in the examination paper accept alternative solutions from those 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.

Rounding errors were not penalised. However, candidates may have lost marks where excessive rounding led to significantly different answers.

In cases where the same error was carried forward to later parts of the answer, candidates were given appropriate credit for the later parts.

In questions where comments were required, valid comments that were different from those provided in the solutions also received full credit where appropriate.

The paper included a number of multiple choice questions, where showing working was not required as part of the answer. In all multiple choice questions, the details provided in the answers in this report (e.g. calculations) are for information.

In all numerical questions that were not multiple-choice, full credit was given for correct answers that also included appropriate workings.

Standard keyboard typing was accepted for mathematical notation.

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

Well prepared candidates were able to score highly, however, a number of candidates appeared to be inadequately prepared, in terms of not having covered sufficiently the entire breadth of the subject.

Candidates are encouraged to practise more on the fundamentals of mathematical calculus and probability. For example, mixed answers in Question 7 suggest that a number of candidates would benefit from additional work on the derivation of the likelihood function, as well as standard calculus (e.g. differentiation).

Candidates are reminded that in questions requiring numerical answers, sufficient details must be shown in the calculations for achieving full marks (e.g. Question 6(ii), Question 10(vi)).

C. Pass Mark

The Pass Mark for this exam was 62.
1460 presented themselves and 746 passed.

Solutions for Subject CS1A – April 2023**Q1**

(i)

Y follows an exponential distribution with rate λ [1]

(ii)

The Correct answer is B. [3]

The two random variables X and Y are independent, therefore the density of their sum can be calculated by convolution

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(z-y)f_Y(y) dy = \int_0^z \lambda^2(z-y)e^{-\lambda(z-y)}\lambda e^{-\lambda y} dy,$$

since $z-y \geq 0$ and $y \geq 0$.

$$f_Z(z) = \lambda^3 e^{-\lambda z} \int_0^z (z-y) dy = \frac{\lambda^3}{2} z^2 e^{-\lambda z}.$$

[Total 4]

Part (i) was very well answered. A number of candidates failed to state the parameter of the distribution.

Part (ii) was generally well answered.

Q2

Two messages are sent in the first minute from the options:

both from the first friend or

both from the second friend or

one from each friend.

[1]

The probability is therefore:

$$P = P(N^A = 2, N^B = 0) + P(N^A = 0, N^B = 2) + P(N^A = 1, N^B = 1). \quad [1]$$

Hence

$$P = e^{-5} \left(\frac{3^2 2^0}{2! 0!} + \frac{3^0 2^2}{0! 2!} + \frac{3^1 2^1}{1! 1!} \right) \quad [1]$$

$$= e^{-5} \frac{25}{2} = 0.084 \quad [1]$$

Alternative answer:

$N^A \sim \text{Poi}(3)$, $N^B \sim \text{Poi}(2)$ independently

$N = N^A + N^B \sim \text{Poi}(5)$,

by independence

$$P(N = 2) = \exp(-5) * 5^2 / 2!$$

$$= 0.084$$

[Total 4]

Very well answered by the majority of candidates. A common error in the alternative answer was not mentioning independence.

Q3

$$M'_X(t) = 4(2500)(1 - 2500t)^{-5} \quad [1/2]$$

$$M''_X(t) = 20(2500)^2(1 - 2500t)^{-6} \quad [1]$$

$$E(X) = M'_X(0) = 10,000 \quad [1/2]$$

$$E(X^2) = M''_X(0) = 125,000,000 \quad [1]$$

$$\text{Var}(X) = 125,000,000 - 10,000^2 = 25,000,000$$

$$\text{SD}(X) = \sqrt{\text{Var}(X)} = 5,000 \quad [1]$$

Alternative answer:

MGFs are unique –

this is a gamma(4, 1/2500).

$$\text{Variance}(X) = 4/(1/2500)^2 = 25,000,000$$

$$\text{SD}(X) = 5,000$$

[Total 4]

Generally well answered. There were errors with the derivation of the derivatives. For the alternative, only a small number of candidates mentioned the uniqueness property.

Q4

(i)

The correct answer is A. [3]

We know that $\frac{9S_X^2}{\sigma^2} \sim \chi_9^2$

Hence

$$P(S_X^2 > 1.5\sigma^2) = P\left(\frac{9S_X^2}{\sigma^2} > 9 \times 1.5\right) = 1 - P(\chi_9^2 \leq 13.5) = 0.14$$

(ii)

We have $\frac{16S_Y^2}{\sigma^2} \sim \chi_{16}^2$. [1]

$$P(S_Y^2 > 1.5\sigma^2) = P\left(\frac{16S_Y^2}{\sigma^2} > 16 \times 1.5\right) \quad [1]$$

$$= 1 - P(\chi_{16}^2 \leq 24) = 0.09 \quad [1]$$

[Total 6]

Both parts were well answered, with no particular issues.

Q5

The random variable h has a discrete prior distribution. Therefore, the posterior distribution of h must also have a discrete distribution.

Let X denote the number of heads obtained when the coin is tossed 15 times.

Then $X \sim \text{Bin}(15, h)$ and: [1]

$$P(h = 0.35 | X = 9) = \frac{P(h=0.35 \text{ and } X=9)}{P(X=9)}$$

$$= \frac{P(X=9 | h=0.35) P(h=0.35)}{P(X=9 | h=0.35) P(h=0.35) + P(X=9 | h=0.85) P(h=0.85)}$$
 [2]

Given $X \sim \text{Bin}(15, h)$, the required probabilities are:

$$P(X = 9 | h = 0.35) = \binom{15}{9} 0.35^9 \cdot 0.65^6 = 0.02975$$
 [1]

$$P(X = 9 | h = 0.85) = \binom{15}{9} 0.85^9 \cdot 0.15^6 = 0.01320$$
 [1]

Therefore, the posterior probabilities for h are:

$$P(h = 0.35 | X = 9) = \frac{0.02975 \times 0.7}{0.02975 \times 0.7 + 0.0132 \times 0.3} = 0.84023$$
 [1]

$$P(h = 0.85 | X = 9) = \frac{0.0132 \times 0.3}{0.02975 \times 0.7 + 0.0132 \times 0.3} = 0.15977$$
 [1]

Alternative answer for $P(h = 0.85 | X = 9)$ is:

$$P(h = 0.85 | X = 9) = 1 - P(h = 0.35 | X = 9) = 1 - 0.84023 = 0.15977$$

[Total 7]

Many candidates answered this question well. However, a number of candidates did not attempt it. There were some calculation errors and some candidates did not present enough details in the various steps of the answer.

Q6

(i)

Yes, the analyst is right. [1]

The distribution of X for $X > 3$ is not known, so the exact expectation of X cannot be determined. [1]

The expectation will be smallest if we assume that $X = 4$ whenever $X > 3$, that is, there are no families with more than 4 children. In that case we can calculate the expectation. [1]

(ii)

$$E[X] = 1 \times 0.42 + 2 \times 0.4 + 3 \times 0.1 + 4 \times p_4 + 5 \times p_5 + \dots$$
 [1]

$$\geq 1 \times 0.42 + 2 \times 0.4 + 3 \times 0.1 + 4 \times (p_4 + p_5 + \dots)$$
 [1]

$$= 0.42 + 0.8 + 0.3 + 4 \times 0.03 = 1.64$$
 [1]

(iii)

The correct answer is C

[3]

$$E[X|X > 3] = 4P[X = 4|X > 3] + 5P[X = 5|X > 3] + \dots$$

$$= 4 \frac{P[X=4]}{P[X>3]} + 5 \frac{P[X=5]}{P[X>3]} + \dots = \frac{1}{P[X>3]} (4p_4 + 5p_5 + \dots)$$

Therefore, $(4p_4 + 5p_5 + \dots) = 4.5 * 0.03 = 0.135$.

So, the expectation of X is: $0.42 + 0.8 + 0.3 + 0.135 = 1.655$.

[Total 9]

In part (i), the justification of the existence of a lower limit was frequently missed.

Part (ii): in many cases candidates did not show sufficient workings to justify their numerical answer. A common error was ignoring the part for $X > 3$ in the calculation.

Part (iii): There were mixed answers in this part.

Q7

(i)

The correct answer is B.

[3]

Likelihood:

$$L(\theta) = \prod_{i=1}^n f(Y_i, X_i; \theta) \quad (\text{jointly independent})$$

$$= \prod_{i=1}^n f(Y_i|X_i = x_i; \theta) \times f(x_i),$$

$$= \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{(y_i - x_i\theta)^2}{2}\right] \times \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{x_i^2}{2}\right]$$

Therefore:

$$L(\theta) = \prod_{i=1}^n \frac{1}{2\pi} \exp\left[-\frac{(y_i - x_i\theta)^2 + x_i^2}{2}\right]$$

(ii)

Log-likelihood:

$$l(\theta) = \log(L(\theta))$$

$$= -n \log(2\pi) - \sum \frac{(y_i - x_i\theta)^2}{2} - \sum \frac{x_i^2}{2} \quad [1]$$

Taking the first derivative gives

$$l'(\theta) = \sum x_i (y_i - x_i\theta) \quad [1\frac{1}{2}]$$

$$\text{Equating to 0 gives } \hat{\theta} = \frac{\sum x_i y_i}{\sum x_i^2} \quad [1\frac{1}{2}]$$

(iii)

Second derivative

$$l''(\theta) = -\sum_i x_i^2 \quad [1]$$

$$E(l''(\theta)) = -\sum_i E(X_i^2) = -n, \quad [1\frac{1}{2}]$$

X_i identically distributed $N(0,1)$. [1/2]

Therefore $CRLB = 1/n$. [1]

(iv)

The asymptotic distribution is $\hat{\theta} \sim N(\theta, CRLB)$ therefore $\hat{\theta} \sim N\left(\theta, \frac{1}{n}\right)$ [2]

[Total 13]

Parts (i)-(iii): there were mixed answers here, with many candidates failing to identify the correct form of the likelihood function. There were also frequent errors with the mathematics parts, e.g. taking the logarithm of the function and calculating the first and second derivatives.

Part (iv) was well answered.

Q8

(i)

Test $H_0: \beta_1 = 0$ v. $H_1: \beta_1 \neq 0$. [1]

Test statistic = $-0.2162 / 0.0532 = -4.06$. [1]

p-value = $2 * P(Z < -4.06) = 2 * 0.00002 = 0.00004$ [1]

We have very strong evidence against H_0 [1]
and conclude that air temperature significantly affects the number of damaged components. [1]

(ii)

Using the parameter estimates, the linear predictor is:
 $11.6630 - 0.2162 * 31 = 4.9608$. [1/2]

Using the natural link for the binomial GLM we have:

$p.\hat{a}t = \exp(4.9608) / \{1 + \exp(4.9608)\}$ [1]

i.e. $p.\hat{a}t = 0.99$ [1/2]

(iii)

The expected value of the number of components that will be damaged is
 $6 * 0.99 = 5.94$. [1]

(iv)(a)

$P(\text{at least 5 not damaged}) = P(\text{at most 1 damaged})$ [1/2]

with number damaged components following Binomial(6, 0.99). [1/2]

So, $P(\text{at least 5 not damaged}) = P(\text{at most 1 damaged}) =$

$P(0 \text{ damaged}) + P(1 \text{ damaged}) =$

$(1 - 0.99)^6 + 6 * 0.99 * (1 - 0.99)^5 = 5.95e-10$ [2]

(b)
Launch is not safe when the air temperature is as cold as 31 degrees Fahrenheit [1]

(v)
This approach is not suitable [1]
as the analysis may give probability estimates that are greater than 1 [1]

[Total 14]

Part (i): Mixed answers, with common errors including incorrect p-values (often involving a t distribution).

Part (ii): Most candidates calculated the value of the linear predictor. However, many answers did not use the logistic link function for the required probability.

Part (iii): Generally well answered.

Part (iv): Mixed answers, with many comments being inconsistent with earlier answers.

Part (v): A number of candidates failed to give the correct justification for the approach not being suitable.

Q9

(i)
Correct answer is D [3]

Likelihood function: $L(p) \propto \left(\frac{1}{2}p\right)^{n_0} p^{n_1} \left(\frac{1}{4}p\right)^{n_2} \left(\frac{1}{4}p\right)^{n_3} (1 - 2p)^{n_4}$

Log-likelihood:

$$\begin{aligned} l(p) &= \text{Constant} + n_0 \log\left(\frac{1}{2}p\right) + n_1 \log p + n_2 \log\left(\frac{1}{4}p\right) + n_3 \log\left(\frac{1}{4}p\right) \\ &\quad + n_4 \log(1 - 2p) \\ &= \text{Constant} + n_0 \log\left(\frac{1}{2}\right) + n_2 \log\left(\frac{1}{4}\right) + n_3 \log\left(\frac{1}{4}\right) + (n_0 + n_1 + n_2 + n_3) \log p \\ &\quad + n_4 \log(1 - 2p) \\ &= \text{Constant} + (n_0 + n_1 + n_2 + n_3) \log p + n_4 \log(1 - 2p) \end{aligned}$$

(ii)

First derivative:

$$l'(p) = (n_0 + n_1 + n_2 + n_3) \frac{1}{p} + n_4 \frac{-2}{1-2p} = 0 \quad [2\frac{1}{2}]$$

$$\hat{p} = \frac{n_0 + n_1 + n_2 + n_3}{2(n_0 + n_1 + n_2 + n_3 + n_4)} = \frac{n - n_4}{2n} \quad [1\frac{1}{2}]$$

where $n = n_0 + n_1 + n_2 + n_3 + n_4$ is the total number of policyholder in the sample

(iii)

The probability for having at least one car is:

$$P[X \geq 1] = p + \frac{1}{4}p + \frac{1}{4}p + (1 - 2p) = 1 - \frac{1}{2}p$$

$$\text{Or } P[X \geq 1] = 1 - P[X = 0] = 1 - \frac{1}{2}p \quad [1\frac{1}{2}]$$

$$\text{Likelihood function: } L(p) \propto \left(\frac{1}{2}p\right)^{m_0} \left(1 - \frac{1}{2}p\right)^{m_1} \quad [1\frac{1}{2}]$$

Log-likelihood:

$$l(p) = m_0 \log\left(\frac{1}{2}p\right) + m_1 \log\left(\frac{2-p}{2}\right) + \text{Constant} \quad [1]$$

(iv)

Correct answer is C [3]

First derivative:

$$l'(p) = \frac{1}{2}m_0 \frac{2}{p} - \frac{1}{2}m_1 \frac{2}{2-p} = m_0 \frac{1}{p} - m_1 \frac{1}{2-p} = 0$$

$$(2 - p)m_0 - pm_1 = 0$$

$$2m_0 = p(m_0 + m_1)$$

$$\hat{p} = \frac{2m_0}{m_0 + m_1}$$

(v)

$$\hat{p} = \frac{n - n_4}{2n} = \frac{50 + 37 + 17 + 16}{2(50 + 37 + 17 + 16 + 10)} = \frac{6}{13} = 0.4615 \quad [1]$$

(vi)

$$\hat{p} = \frac{2m_0}{m_0 + m_1} = \frac{100}{50 + 37 + 17 + 16 + 10} = \frac{10}{13} = 0.769 \quad [1]$$

[Total 16]

This question was answered well generally.

In part (ii), there were various derivation errors for $p_{\hat{}}$.

Q10

(i)

$$S_{xx} = \sum x_i^2 - (\sum x_i)^2/n = 80.309 \quad [1\frac{1}{2}]$$

$$S_{xy} = \sum x_i y_i - (\sum x_i)(\sum y_i)/n = 36.729 \quad [1\frac{1}{2}]$$

$$\hat{\beta} = \frac{S_{xy}}{S_{xx}} = 0.457, \hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} = 37.067 \quad [2]$$

The fitted regression line is

$$\hat{y} = 37.067 + 0.457x \quad [1]$$

(ii)

$$S_{yy} = \sum y_i^2 - \left(\sum y_i\right)^2 / n = 24.409$$

Pearson's correlation coefficient is:

$$r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} = 0.830 \quad [1]$$

(iii)

$$H_0: \rho = 0.8 \text{ vs } H_1: \rho \neq 0.8 \quad [1]$$

Under H_0 , $W = \frac{1}{2} \log \left(\frac{1+r}{1-r} \right)$ has (approximately) a normal distribution with mean $\frac{1}{2} \log \left(\frac{1+\rho}{1-\rho} \right)$ and standard deviation $\frac{1}{\sqrt{n-3}}$. [1]

With $r = 0.83$, we obtain $W = 1.188$ and under H_0 W has a normal distribution with mean 1.099 and standard deviation $1/3$ [1]

Hence the p-value (since this is a two-sided test) is $2P(W > 1.188) = 2P(Z > (1.188 - 1.099)/(1/3)) = 2P(Z > 0.267) = 0.790$ [1]

We have no evidence even at 1% against the null hypothesis that $\rho = 0.8$ [1]

(iv)

$$\hat{\sigma}^2 = \frac{1}{n-2} (S_{yy} - S_{xy}^2/S_{xx}) = \frac{1}{10} (24.409 - 36.729^2/80.309) = 0.761 \quad [1]$$

$$\text{Hence, s.e.}(\hat{\beta}) = (\hat{\sigma}^2/S_{xx})^{0.5} = 0.097 \quad [1/2]$$

The 99% confidence interval for β is

$$\begin{aligned} \hat{\beta} \pm \{t_{0.005,10} \times \text{s.e.}(\hat{\beta})\} &= 0.457 \pm 3.169 \times 0.097 & [1] \\ &= (0.150, 0.764) & [1/2] \end{aligned}$$

(v)

In part (iv), the confidence interval does not contain 0 and we can conclude there is a linear relationship between the model and experts scores [1]

The test in part (iii) concluded no evidence against 0.8 correlation between the model's scores and experts' scores [1/2]

The two results are consistent [1/2]

(vi)

Since all model scores are unique and separately all experts scores are also unique, the Spearman's rank correlation coefficient can be calculated as:

$$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2-1)}, \text{ where } d_i = \text{rank}(x_i) - \text{rank}(y_i) \quad [1]$$

rank(x_i)	4	2	7	3	9	1	11	5	8	6	10	12
rank(y_i)	8	4	7	2	10	3	9	1	11	5	6	12
d_i	-4	-2	0	1	-1	-2	2	4	-3	1	4	0
d_i^2	16	4	0	1	1	4	4	16	9	1	16	0

[4]

$$r_s = 0.748 \quad [1]$$

(vii)

The Spearman correlation seems slightly lower than the Pearson's correlation [1]
but the model's alignment with the experts' opinion is good [1]

Alternative comment:

Both measures are high and positive, suggesting strong alignment

[Total 23]

The question was answered well overall.

In part (iii), a common error was calculating the p-value incorrectly, often using a one-sided test.

The comments in parts (v) and (vii) were unclear in many cases.

In part (vi) many candidates did not show sufficient details in their calculations.

[Paper Total 100]

END OF EXAMINERS' REPORT



Institute
and Faculty
of Actuaries

Beijing

14F China World Office 1 · 1 Jianwai Avenue · Beijing · China 100004
Tel: +86 (10) 6535 0248

Edinburgh

Level 2 · Exchange Crescent · 7 Conference Square · Edinburgh · EH3 8RA
Tel: +44 (0) 131 240 1300

Hong Kong

1803 Tower One · Lippo Centre · 89 Queensway · Hong Kong
Tel: +852 2147 9418

London (registered office)

7th Floor · Holborn Gate · 326-330 High Holborn · London · WC1V 7PP
Tel: +44 (0) 20 7632 2100

Oxford

1st Floor · Belsyre Court · 57 Woodstock Road · Oxford · OX2 6HJ
Tel: +44 (0) 1865 268 200

Singapore

5 Shenton Way · UIC Building · #10-01 · Singapore 068808
Tel: +65 8778 1784

www.actuaries.org.uk

© 2021 Institute and Faculty of Actuaries