Teacher Support Materials 2009

Maths GCE

Paper Reference MS04

Question 1

1 A teacher believes that the Verbal Reasoning Quotient (VRQ) of first-born children in a family is higher than that of subsequent children in the family. The teacher randomly selects ten families, which each have two children, and then records the VRQ for each child. The results are shown in the table.

Family	A	B	C	D	E	F	G	H	I	J
First-born VRQ	110	108	99	107	121	128	107	110	123	106
Second-born VRQ	110	108	100	101	123	127	103	106	122	103

Assuming that differences in the VRQ are normally distributed, investigate the teacher's belief at the 5% level of significance. *(10 marks)*

Student Response

1 d = Second-born − first-born

d: 0 0 1 −6 2 −1 −4 −4 −1 −3 M 1

$\bar{d} = -1.6$ ✓ A 1

$s_d = 2.5473$. ✓ A 1

H_0: $\bar{d}$ is normally distributed. X B 0

H_1: $\bar{d}$ is not normally distributed. X B 0

$T = \dfrac{\bar{d} - 0}{s_d/\sqrt{n}} = \dfrac{-1.6 - 0}{\frac{2.5473}{\sqrt{10}}} = -1.99$ ✓ M 1 A 1

$v = n - 1 = 10 - 1 = 9$ ✓ B 1

5% ; 2 tails $\Rightarrow t = 2.262$. X

B 0

6/

Commentary

The candidate correctly finds the differences between the first-born and second-born VRQ's. The mean and standard deviation of these differences is also calculated correctly. The hypotheses, however, are stated incorrectly. The test statistic was correctly calculated and the number of degrees of freedom was stated correctly. The critical value quoted is for a two-tail test and in this case a one-tail test is required. Having no correct hypotheses the candidate was unable to state any conclusion. There were many completely correct responses to this question, although this is a relatively weak answer.

Mark scheme

1	Differences are: 0, 0, –1, 6, –2, 1, 4, 4, 1, 3	M1		
	H₀: $\mu_d = 0$	B1		$\overline{d}$ for μ_d and other poor notation B1B0
	H_1: $\mu_d > 0$	B1		
	$\overline{d} = 1.6$	A1		
	$s = 2.547$	A1		
	$t_{\text{calc}} = \dfrac{1.6-0}{\left(\dfrac{2.547}{\sqrt{10}}\right)} = 1.986$	M1 A1F		
	$v = 9$	B1		
	$t_{\text{crit}} = 1.833$	B1		
	Reject H_0. Evidence at 5% level to suggest 1st born has higher VR	A1F	10	
	Total		**10**	

Question 2

2 The random variable X denotes the number of trials necessary in order to obtain the first success.

(a) State **three** conditions which must apply in order that X may be modelled by a geometric distribution. *(3 marks)*

(b) The discrete random variable Y is such that $Y \sim \text{Geo}(p)$ and $P(Y = 1 \text{ or } 2) = 0.2775$.

(i) Determine the value of p. *(4 marks)*

(ii) Hence find the values of $E(Y)$ and $\text{Var}(Y)$. *(2 marks)*

Student response

2. a) • there are only 2 possible out comes / success or fail. ✓ (9)

• Probablity of success is constant ✓

• Trials are independent ✓ E 3

b) i) $P(Y = 1 \text{ or } 2) = p + p(1-p) = -p^2 + 2p = 0.2775$

$\Rightarrow p^2 - 2p + 0.2775 = 0$

$p = 0.15$ or ~~...~~ 1.85 (not valid) ✓

$\Rightarrow p = 0.15$ 4

ii) $E(Y) = \frac{1}{p} = \frac{1}{0.15} = \frac{20}{3}$ ✓

$Var(Y) = \frac{1-p}{p^2} = \frac{1-0.15}{0.15^2} = \frac{340}{9}$ ✓ 2

Commentary

This question was the least well done question on the paper and this candidate has produced a complete solution. In part (a) there are three correct conditions for the geometric distribution. In part (b) (i) the direct approach is adopted and both probabilities are expressed in terms of *p*, the probability of success in a single trial. Thus a quadratic equation is formed and the value between 0 and 1 is the correct value for *p*. Notice that the candidate clearly rejects the invalid solution. In part (b) (ii) the value of *p* is inserted in the formulae for the mean and variance of the geometric distribution.

Mark Scheme

2(a)	Independent trials Two outcomes OE Constant probability of success Unlimited number of trials	E1 × 3	3	Any three
(b)(i)	$p + p(1-p) = 0.2775$ $p^2 - 2p + 0.2775 = 0$ $p = 0.15 \quad (0 < p < 1)$	M1 m1 m1 A1	4	$1-(1-p)^2 = 0.2775$ $(1-p)^2 = 0.2775$ $(1-p) = 0.85$ $p = 0.15$
(ii)	$\mathrm{E}(Y) = \dfrac{1}{0.15} = 6.67$ $\mathrm{Var}(Y) = \dfrac{0.85}{0.15^2} = 37.8$	B1F B1	2	ft on $0 < p < 1$
	Total		**9**	

Question 3

3 Fourteen randomly selected physics students are asked to determine independently the density of copper by a particular method. Their results, in $\mathrm{kg\,m^{-3}}$, are shown below.

8924	8929	8931	8926	8928	8925	8929
8921	8925	8927	8930	8920	8923	8922

(a) Construct a 98% confidence interval for the standard deviation of the density of copper as determined by a physics student using this particular method. *(6 marks)*

(b) State **one** assumption that you have made in constructing this confidence interval. *(1 mark)*

Student Response

Leave blank

3. a) c.i. $\nu = 4.107, 27.688$ (χ^2 dist) B 1

$\nu = 13$

$\frac{3.4513^2}{4.107}$ ✗, $\frac{3.4513^2}{27.688}$ ✗

$s = 3.4513$ B 1

$(0.359, 2.42)$ B 1

~~$\frac{3.4513}{\sqrt{3}} = 3.304$~~ M 0

~~$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$~~

b) densities are normally distributed. E 1

E(X_4)

4

Commentary

This candidate has picked up marks for knowing that a chi-squared distribution with 13 degrees of freedom is required here. Furthermore, the correct chi-squared values for a 98% confidence interval have been found from tables, or calculator. The final mark is also gained by stating the assumption that has been made. Crucially the candidate is unable to recall the formula for the confidence limits, viz. $\sqrt{\frac{(n-1)s^2}{\chi^2}}$. This is a fine example of good examination technique, where the candidate has maximised the mark, despite not fully knowing what to do.

Mark Scheme

3(a)	$s = 3.451$	B1		$s^2 = 11.9123 \quad \sum(x-\bar{x})^2 = 154.86$
	$\nu = 13$	B1		
	$\chi^2_{13}(0.01) = 4.107$ $\chi^2_{13}(0.99) = 27.688$	B1		
	98% CL for σ are $\sqrt{\frac{13\times 3.451^2}{27.688}}$ and $\sqrt{\frac{13\times 3.451^2}{4.107}}$	M1 A1√		ft on χ^2 values
	98% CI is (2.36, 6.14)	A1	6	AWFW (2.36, 2.37) and (6.135, 6.145)
(b)	Sample is from a normal distribution	E1	1	
	Total		7	

Question 4

4 The rateable values of businesses in a certain town are distributed with mean μ and variance σ^2. Two trainee valuers are asked to estimate μ.

Trainee valuer A intends to take a random sample of 15 businesses and to calculate the mean rateable value, $\overline{X}_A$, of the businesses in this sample.

(a) Write down expressions for the mean and the variance of $\overline{X}_A$. *(2 marks)*

(b) Trainee valuer B intends to take a random sample of 10 businesses and to calculate the mean rateable value, $\overline{X}_B$, of the businesses in this sample.

The principal valuer suggests that they could obtain a better estimate of μ by combining their individual estimators.

The trainee valuers consider two possible combinations, $\overline{X}_L$ and $\overline{X}_M$, of their estimators, where

$$\overline{X}_L = \tfrac{1}{2}\overline{X}_A + \tfrac{1}{2}\overline{X}_B \quad \text{and} \quad \overline{X}_M = \tfrac{3}{5}\overline{X}_A + \tfrac{2}{5}\overline{X}_B$$

(i) Show that the estimator $\overline{X}_M$ has mean μ and variance $\dfrac{\sigma^2}{25}$. *(3 marks)*

(ii) Show that $\overline{X}_L$ is an unbiased estimator of μ. *(1 mark)*

(iii) Calculate the relative efficiency of $\overline{X}_L$ with respect to $\overline{X}_M$, and give a reason why $\overline{X}_M$ should be preferred to $\overline{X}_L$ as an estimator of μ. *(5 marks)*

Student Response

4 a) $E(\bar{X}_A) = \mu$ ✓

$Var(\bar{X}_A) = \dfrac{\sigma^2}{15}$ ✓ B1

B1

b) i) $E(\bar{X}_M) = E\left(\frac{3}{5}\bar{X}_A + \frac{2}{5}\bar{X}_B\right)$

$E(\bar{X}_B) = \mu \Rightarrow E(\bar{X}_M) = \frac{3}{5}E(\bar{X}_A) + \frac{2}{5}E(\bar{X}_B)$

$= \frac{3}{5}\mu + \frac{2}{5}\mu$

$= \mu$ ✓ B1

$Var(\bar{X}_B) = \dfrac{\sigma^2}{10}$

$Var(\bar{X}_M) = Var\left(\frac{3}{5}\bar{X}_A + \frac{2}{5}\bar{X}_B\right)$

$= \frac{9}{25}Var(\bar{X}_A) + \frac{4}{25}Var(\bar{X}_B)$

$$= \frac{9}{25} \times \frac{\sigma^2}{15} + \frac{4}{25} \times \frac{\sigma^2}{10}$$ M1

$$= \frac{\sigma^2}{25}$$ A1

b(ii) $E(\bar{X}_L) = E\left(\frac{1}{2}\bar{X}_A + \frac{1}{2}\bar{X}_B\right)$

$$= \frac{1}{2}E(\bar{X}_A) + \frac{1}{2}E(\bar{X}_B)$$

$$= \frac{1}{2}\mu + \frac{1}{2}\mu = \mu$$ B1

$\Rightarrow \bar{X}_L$ is an unbiased estimator of μ.

iii) The relative efficiency of $\bar{X}_L$ with respect to $\bar{X}_M$:

~~RE~~

$$RE = \frac{\frac{1}{Var(\bar{X}_L)}}{\frac{1}{Var(\bar{X}_M)}} = \frac{Var(\bar{X}_M)}{Var(\bar{X}_L)}$$

$$Var(\bar{X}_L) = Var\left(\frac{1}{2}\bar{X}_A + \frac{1}{2}\bar{X}_B\right)$$

$$= \frac{1}{4}Var(\bar{X}_A) + \frac{1}{4}Var(\bar{X}_B)$$

$$= \frac{1}{4} \times \frac{\sigma^2}{15} + \frac{1}{4} \times \frac{\sigma^2}{10}$$ M1

$$= \cancel{\frac{\sigma^2}{24}}\ \frac{\sigma^2}{24}$$ A1

Hence: $RE = \dfrac{Var(\bar{X}_M)}{Var(\bar{X}_L)} = \dfrac{\frac{\sigma^2}{25}}{\frac{\sigma^2}{24}} = \dfrac{24}{25}$ M1 A1

$RE = \frac{24}{25} < 1 \Rightarrow \bar{X}_M$ should be ~~prefea~~ preferred to $\bar{X}_L$ as an estimator of μ E1

(7)

Commentary

The responses to this question frequently suffered from poor notation and lack of explanation as to what candidates were doing. In this case the candidate exhibits exemplary explanation and correct notation. In part (a) the required results are clearly stated. In part (b) (i), since the answers are given on the question paper, sufficient working must be shown and this is the case here. Part (b) (ii) requires a proof that $E(\bar{X}_L) = \mu$, which the candidate has done in the same fashion as performed in part (a). The final part requires $Var(\bar{X}_L)$ to be calculated in the same fashion as in part (a) and hence obtain the relative efficiency of $\bar{X}_L$ with respect to $\bar{X}_M$. For the comment, the candidate has learned that if the relative efficiency is <1 then $\bar{X}_M$ is preferred. Other candidates referred to the sizes of the variances, which was equally acceptable.

Mark Scheme

4(a)	$\mathrm{E}(\bar{X}_A) = \mu$	B1		
	$\mathrm{Var}(\bar{X}_A) = \dfrac{\sigma^2}{15}$	B1	2	
(b)(i)	$\mathrm{E}(\bar{X}_M) = \dfrac{3}{5}\mu + \dfrac{2}{5}\mu = \mu$	B1		AG
	$\mathrm{Var}(\bar{X}_M) = \dfrac{9}{25} \times \dfrac{\sigma^2}{15} + \dfrac{4}{25} \times \dfrac{\sigma^2}{10}$	M1		
	$= \dfrac{\sigma^2}{25}$	A1	3	AG
(ii)	$\mathrm{E}(\bar{X}_L) = \dfrac{1}{2}\mu + \dfrac{1}{2}\mu = \mu$	B1	1	
(iii)	$\mathrm{Var}(\bar{X}_L) = \dfrac{1}{4} \times \dfrac{\sigma^2}{15} + \dfrac{1}{4} \times \dfrac{\sigma^2}{10}$	M1		
	$= \dfrac{\sigma^2}{24}$	A1		
	Rel. Eff. $= \dfrac{24}{\sigma^2} \div \dfrac{25}{\sigma^2} = \dfrac{24}{25}$	M1 A1F		ft on $\mathrm{Var}(\bar{X}_L)$
	$< 1 \Rightarrow$ prefer $\bar{X}_M$	E1F	5	OE eg $\mathrm{Var}(\bar{X}_M) < \mathrm{Var}(\bar{X}_L)$
	Total		**11**	

Question 5

5 Henrietta, a statistician, buys boxes of 6 eggs from a local farm shop. She believes that the number of light brown eggs in each box follows a binomial distribution. In order to test this belief, she selects a random sample of 100 boxes of eggs from the farm shop and records the number of light brown eggs in each box, with the following results.

Number of light brown eggs per box	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**Total**
Frequency	23	32	23	17	4	1	0	100

(a) (i) Use these data to estimate the mean number of light brown eggs in a box of 6 eggs. *(1 mark)*

(ii) Hence show that a suitable estimate of p, the probability that a randomly chosen egg is light brown, is 0.25 . *(1 mark)*

(b) By calculating an appropriate χ^2-statistic, test, at the 5% level of significance, whether a binomial distribution is an appropriate model for the number of light brown eggs in a box of 6 eggs. *(10 marks)*

Student Response

5. a) (i) $\bar{X} = 1.5$ ✓ B 1

(ii) B 0

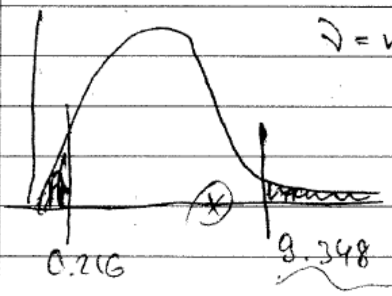
b) H_0: Binomial ✓ $Bin(\cancel{100}\ 6, 0.25)$ B 1

H_1: Not binomial

O	E	$\frac{(O-E)^2}{E}$			
23	${}^6C_0\, 0.25^0 (0.75)^6 \times 100 = 18$	1.389		M	1
32	${}^6C_1\, 0.25^1 (0.75)^5 \times 100 = 36$ ✓	0.444		M	1
23	${}^6C_2\, 0.25^2 (0.75)^4 \times 100 = 30$ ✓	1.633			
17	${}^6C_3\, 0.25^3 (0.75)^3 \times 100 = 13$ ✓	1.231	Not 1dp.	A	0
4	${}^6C_4\, 0.25^4 (0.75)^2 \times 100 = 3$ ✓	0.333			
1	0	0	No classes combined	M	0
0	0	0		M	1

5.031 ✗ A 0

$\nu = n - 3 = 6 - 3 = 3$ ✗ B 0



$2\frac{1}{2}\%$

Accept H_0. Evidence that Binomial. ✓ B 0

A 1

6

Commentary

This type of question is frequently done very well and high marks are obtained on it. This candidate has made numerous small errors, which have resulted in the loss of a good number of marks. Although the mean has been correctly stated in part (a) (i) the candidate was unable to use *np* to calculate the probability of success for the binomial distribution concerned (S1 specification). The value of the test statistic is incorrect because the candidate has only calculated expected frequencies to the nearest integer (1 dp was required), nor have classes been combined when expected frequencies fell below 5. The result of these errors was the loss of 3 marks. Two further marks were lost since the degrees of freedom were incorrect and although a follow through mark might have been available, it was lost since the critical value quoted is for a 2.5% level rather than a 5% level. With more thorough learning, 5 or 6 marks (nearly a grade) could easily have been gained.

Mark Scheme

5(a)(i)	$\bar{x} = 1.5$	B1	1	
(ii)	$6p = 1.5 \Rightarrow p = 0.25$	B1	1	AG
(b)	H_0: distribution is binomial	B1		
	O_i / E_i: 23 / 17.80; 32 / 35.60; 23 / 29.66; 17 / 13.18; 4 / 3.30; 1 / 0.44; 0 / 0.02	M1 M1 A1		Attempt at probabilities Probabilities × 100 ≥ 4 correct (1dp)
	Combine classes	M1		
	$\chi^2_{\text{calc}} = \frac{5.20^2}{17.80} + \frac{3.60^2}{35.60} + \frac{6.66^2}{29.66} + \frac{5.06^2}{16.94}$	M1		Use of formula
	$= 4.89$	A1		AWFW (4.85, 4.95)
	$v = 4 - 2 = 2$	B1		
	$\chi^2_{\text{crit}} = 5.991$	B1F		ft on v ($2 \le v \le 6$) (not 5%) ($v = 3 \Rightarrow 7.815$)
	Accept H_0. Evidence to suggest binomial distribution is an appropriate model	A1√	10	
	Total		**12**	

Question 6

6 The heights in centimetres, X, of men from police forces in England may be assumed to be normally distributed with variance $\sigma_X{}^2$. The heights in centimetres, Y, of men from police forces in Scotland may also be assumed to be normally distributed but with variance $\sigma_Y{}^2$.

The heights, measured to the nearest centimetre, x, of a random sample of 11 men from police forces in England were

179 189 191 179 188 178 189 181 174 186 187

The heights, measured to the nearest centimetre, y, of a random sample of 9 men from police forces in Scotland were

182 180 178 175 177 180 182 181 181

(a) Calculate unbiased estimates of $\sigma_X{}^2$ and $\sigma_Y{}^2$. *(2 marks)*

(b) (i) Hence determine a 95% confidence interval for the variance ratio $\frac{\sigma_X{}^2}{\sigma_Y{}^2}$. *(7 marks)*

(ii) Comment on the suggestion that the heights of men from police forces in England are more variable than those of men from police forces in Scotland. *(2 marks)*

Student Response

6a) $\sigma x^2 = 32.21818182$ ✓ $\sigma y^2 = 5.\dot{7}$ ✓ 2

$n = 11$ $n = 9$

6b)i) $\frac{\sigma x^2}{\sigma y^2} \times \frac{1}{F_1}, \frac{\sigma x^2}{\sigma y^2} \times F_2$

$F_1 = {}^{v_1}_{v_2} = F_1{}^{10}_{8} = 4.295$

$F_2 = {}^{v_2}_{v_1} = F_2{}^{8}_{10} = 3.855$

$\frac{\sigma x^2}{\sigma y^2} \times \frac{1}{F_1} = \frac{32.2181\ldots}{5.7} \times \frac{1}{4.295}$

$= 1.298305885$

$= \cancel{1.298}\ 1.30$

$\frac{\sigma x^2}{\sigma y^2} \times F_2 = \frac{32.2181}{5.7} \times 3.855$

$= 21.49634266$

$= 21.5$

$= (1.30, 21.5)$ cm ✓ 7

6b)ii) It could be viable, as the variance is larger for police ~~pol~~ forces in England, which widens the confidence interval.

Leave blank

E0

e0

Commentary

The candidate has correctly calculated the variances for both groups of policemen in part (a). (Errors here, such as confusion of variance and standard deviation or incorrect data entry on the calculator, can cost marks later.) In part (b) (i) the candidate has remembered the correct method for obtaining confidence limits and has avoided the variety of problems that can occur, including degrees of freedom, which is v_1 and which is v_2, whether we require a fraction or its reciprocal and confusion between F_{upper} and F_{lower}. The marks for the final part of the question have been lost, however. With many of these comments it is best to refer to what has just been calculated. The key is to state that the confidence interval does not contain 1 (not 0 as some candidates thought). It is then possible to assert that either the suggestion made in the question is correct, or that the heights of policemen from England are more variable than those of policemen in Scotland.

Mark Scheme

6(a)	$S_X^2 = 32.218$	M1		Either
	$S_Y^2 = 5.778$	A1	2	Both correct; condone 2 sf SC: B1 for ≥ 1 sd
(b)(i)	$\nu_1 = 10$, $\nu_2 = 8$	B1		
	$F_{10,8} = 4.295$, $F_{8,10} = 3.855$	B1,B1		
	$F_{\text{calc}} = \dfrac{32.218}{5.778} = 5.576$	M1		
	$\dfrac{1}{4.295} \le \dfrac{VR}{5.576} \le 3.855$	m1 A1√		ft on v_1 and v_2
	$\Rightarrow 1.30 \le VR \le 21.5$	A1	7	Accept 1.3
(ii)	$1 \notin$ CI $\Rightarrow$ Variability greater among men from police forces in England	E1 E1	2	Dependent
	Total		**11**	

Question 7

7 The continuous random variable X is modelled by an exponential distribution with probability density function

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geqslant 0 \\ 0 & x < 0 \end{cases}$$

(a) Write down the cumulative distribution function, $F(x)$. *(2 marks)*

(b) Show that the **exact** value of the interquartile range of X is given by $\frac{1}{\lambda}\ln 3$. *(5 marks)*

(c) (i) Use integration to prove that $E(X^2) = \frac{2}{\lambda^2}$. *(4 marks)*

(ii) Hence, given that $E(X) = \frac{1}{\lambda}$, show that $Var(X) = \frac{1}{\lambda^2}$. *(1 mark)*

(d) (i) Find the **exact** value of λ for which the value of the interquartile range of X is four times the value of the variance of X. *(2 marks)*

(ii) Describe what happens to the interquartile range of X as λ increases without bound. *(1 mark)*

MS04

Student Response

Leave blank

7a) $F(x) = 1 - e^{-\lambda x}$ ✓

B 1

b) $\cancel{\int_0^u \lambda e^{-\lambda x}\,dx = 0.5}$

$\cancel{\left[-e^{-\lambda x}\right]_l^u = 0.5}$

$\cancel{\left(-e^{-\lambda u}\right) - \left(-e^{-\lambda l}\right) = 0.5}$

$\cancel{e^{-\lambda l} - e^{-\lambda u} = 0.5}$

$\cancel{-\lambda l - \lambda u = \ln 0.8}$

$\cancel{l + u =}$

$\int_0^u \lambda e^{-\lambda x}\,dx = 0.75$

$\left[-e^{-\lambda x}\right]_0^u = 0.75$

$\left(-e^{-\lambda u}\right) - \left(-e^{0}\right) = 0.75$

$1 - e^{-\lambda u} = 0.75$

$-e^{-\lambda u} = -0.25$

$e^{-\lambda u} = 0.25$

$-\lambda u = \ln 0.25$

$u = -\frac{1}{\lambda}\ln 0.25$ ✓

$\int_0^l \lambda e^{-\lambda x}\,dx = 0.25$

$\left(-e^{-\lambda l}\right) + 1 = 0.25$

$e^{-\lambda l} = 0.75$ ✓

$-\lambda l = \ln 0.75$

$l = -\frac{1}{\lambda}\ln 0.75$ ✓

IQ range $= u - l$

$= -\frac{1}{\lambda}\left(\ln 0.25 - \ln 0.75\right)$

$= -\frac{1}{\lambda}\left(-\ln 3\right)$

$= \frac{1}{\lambda}\ln 3$ ✓

5

Leave blank

c)(i) $\int x^2 \cdot \lambda e^{-\lambda x}\,dx$

$u = x^2 \quad \frac{du}{dx} = 2x$

$\frac{dv}{dx} = \lambda e^{-\lambda x} \quad v = -e^{-\lambda x}$

allow limits appearing later M 1

$= \left[-x^2 e^{-\lambda x}\right]_0^\infty - \int -2x e^{-\lambda x}\,dx$ ✓ A 1

$= \left[-x^2 e^{-\lambda x}\right]_0^\infty + \left[(2x)\left(\frac{e^{-\lambda x}}{-\lambda}\right)\right]_0^\infty + \int 2\,\frac{e^{-\lambda x}}{-\lambda}\,dx$ ✓ A 1

$= [0] + [0] + \left[\frac{-2e^{-\lambda x}}{\lambda^2}\right]_0^\infty = \frac{2}{\lambda^2}$ ✗ A 0

AG

c)(ii) $Var(X) = E(X^2) - (E(X))^2$

$= 2/\lambda^2 - (1/\lambda)^2$

$= 2/\lambda^2 - 1/\lambda^2 = 1/\lambda^2$ ✓ 1

d)(i)

$1/\lambda \ln 3 = 4\left(1/\lambda^2\right)$

$\frac{1}{\lambda}\ln 3 = \frac{4}{\lambda^2}$

$\ln 3 = 4/\lambda$

$\lambda = \frac{4}{\ln 3}$ ✓ 2

(ii) It gets smaller. and tends to what? E 0

12

Commentary

Part (a) of this student's response lacked sufficient detail. It was necessary to state that $\mathrm{F}(x) = 1 - \mathrm{e}^{-\lambda x}$ when *x* is ≥ 0, and is 0 for other values of *x*. In part (b) the candidate successfully uses the cumulative distribution function to obtain quartiles and hence the interquartile range. In part (c) (i), to earn any marks, limits must be used for the integration. Although they do not initially appear, the candidate subsequently realises that they are needed. The final mark in this section, however, is lost, since full working is not shown and the answer is given on the paper. In part (c) (ii) $\mathrm{E}(X^2)$ is correctly used to find $\mathrm{Var}(X)$. The interquartile range and $\mathrm{Var}(X)$ are correctly used in part (d) (i), but in part (d) (ii) the comment again lacks detail. It is necessary to state that the interquartile range tends to zero as $\lambda \to \infty$. So here we have an essentially correct solution, which could have gained 3 more marks with attention to finer details.

Mark Scheme

7(a)	$\mathrm{F}(x) = 1 - \mathrm{e}^{-\lambda x}\ ,\ x \ge 0$	B1		$\mathrm{F}(x) = 1 - \mathrm{e}^{-\lambda x}$ B1B0
	$\mathrm{F}(x) = 0\ ,\ x < 0$	B1	2	Dependent
(b)	$1 - \mathrm{e}^{-\lambda x} = \frac{3}{4}$	M1		For either Q_1 or Q_3
	$Q_3 = \frac{1}{\lambda}\ln 4$	m1A1		m1 for attempting to solve for either Q_1 or Q_3
	$1 - \mathrm{e}^{-\lambda x} = \frac{1}{4}$			
	$Q_1 = \frac{1}{\lambda}\ln\frac{4}{3}$	A1		
	$\mathrm{IQR} = \frac{1}{\lambda}\ln 3$	A1	5	AG
(c)(i)	$\mathrm{E}(X^2) = \int_0^\infty \lambda x^2 \mathrm{e}^{-\lambda x}\,\mathrm{d}x$	M1		Limits required
	$= \left[-x^2\mathrm{e}^{-\lambda x}\right]_0^\infty + \int_0^\infty 2x\mathrm{e}^{-\lambda x}\,\mathrm{d}x$	A1		
	$= \left[-\frac{2x}{\lambda}\mathrm{e}^{-\lambda x}\right]_0^\infty + \int_0^\infty \frac{2}{\lambda}\mathrm{e}^{-\lambda x}\,\mathrm{d}x$	A1		
	$= \left[-\frac{2}{\lambda^2}\mathrm{e}^{-\lambda x}\right]_0^\infty$	A1	4	
	$= \frac{2}{\lambda^2}$			AG
(ii)	$\mathrm{Var}(X) = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$	B1	1	AG
(d)(i)	$\frac{1}{\lambda}\ln 3 = \frac{4}{\lambda^2}$	M1		
	$\lambda = \frac{4}{\ln 3}$	A1	2	
(ii)	$\mathrm{IQR} \to 0$ as $\lambda \to \infty$	E1	1	
	Total		**15**	