Teacher Support Materials 2008

Maths GCE

Paper Reference MS04

Question 1

1 The volume of fuel consumed by an aircraft making an east–west transatlantic flight was recorded on 10 occasions with the following results, correct to the nearest litre.

68 860	71 266	69 476	68 973	69 318
70 467	71 231	68 977	70 956	69 465

These volumes of fuel may be assumed to be a random sample from a normal distribution with standard deviation σ.

(a) Construct a 99% confidence interval for σ. *(6 marks)*

(b) State one factor that may cause the volume of fuel consumed to vary. *(1 mark)*

Student Response

Leave blank

1.

$$\frac{(n-1)s^2}{\chi_u^2} < \sigma^2 < \frac{(n-1)s^2}{\chi_L^2}$$

~~$s^2 = (n-1)$~~ ~~$s = 976.1$~~

~~$s^2 = 952745.9$~~

~~$\frac{(10-1)\,952745.9}{}$~~ ~~$< \sigma^2 <$~~

$s = 976.1$ ✓ B 1

$s^2 = 952{,}745.9$ ✓

degrees of freedom = 8 ✗ B 0

$$\frac{(10-1)\,952745.9}{21.955} < \sigma^2 < \frac{(10-1)\,952745.9}{1.344}$$ B 0

M 1

$390558.6 < \sigma^2 < 6379994.9$ A 1

$624.9 < \sigma < 2525.9$

$624 < \sigma < 2530$ to 3sf A 0

B. Wind speed. ✓ E 1

4

Commentary

The candidate loses marks here by having 8 degrees of freedom, instead of 9. This means he also gets the wrong values for chi-squared and loses the final accuracy mark. He earns the final mark for a correct comment. A number of candidates lost this mark when they said that more fuel was used when the plane travels further, which is not addressing the concept of variability, which is being tested in this question.

Mark scheme

Q	Solution	Marks	Total	Comments
1(a)	$s = 976.09$	B1		
	$\nu = 9$	B1		
	$\chi_9^2\left(0.005\right) = 1.735$			
	$\chi_9^2\left(0.995\right) = 23.589$	B1		
	99% CL for σ are:			Or σ^2
	$\sqrt{\dfrac{9\times 976.09^2}{23.589}}$ and $\sqrt{\dfrac{9\times 976.09^2}{1.735}}$	M1 A1√		√ on s^2 and χ^2 (with or without $\sqrt{\ }$)
	99% CI is (603, 2220)	A1	6	AWRT
(b)	eg: Weather conditions Load Pilot	E1	1	Any sensible alternative
	Total		7	

Question 2

2 (a) The discrete random variable X follows a geometric distribution with parameter p.

Prove that $\mathrm{E}(X) = \dfrac{1}{p}$. *(3 marks)*

(b) A fair six-sided die is thrown repeatedly until a six occurs.

(i) State the expected number of throws required to obtain a six. *(1 mark)*

(ii) Calculate the probability that the number of throws required to obtain a six is greater than the expected value. *(3 marks)*

(iii) Find the least value of r such that, when the die is thrown repeatedly, there is more than a 90% chance of obtaining a six on or before the rth throw. *(4 marks)*

Student response

2. (a) $X \sim \text{Geo}(p)$

$P(X=x) = p(1-p)^{x-1}$ Let $q = 1-p$

$= pq^{x-1}$

X	1	2	3	4	5	...
$P(X=x)$	p	pq	pq^2	pq^3	pq^4	...

$E(X) = p + 2pq + 3pq^2 + 4pq^3 + 5pq^4 + \cdots$

$= p(1 + 2q + 3q^2 + 4q^3 + 5q^4 + \cdots)$

$= p(1-q)^{-2}$

$= \dfrac{p}{(1-q)^2} = \cancel{=}$

$\because q = 1-p \Rightarrow p = 1-q$

$\therefore \Rightarrow \dfrac{p}{(1-q)^2} = \dfrac{p}{p^2} = \dfrac{1}{p}$

3

Leave blank

(b) (i) 6 — 1

(ii) $\cancel{P(X > 6) = \frac{1}{6}(1-p)^{x-1}}$

$\cancel{= (1-}$

$X \sim Geo(1/6)$

$P(X > 6) = (1-p)^{x-1}$

$= (1 - 1/6)^{5}$

$= 0.402$ — B C

(iii) $P(X \le r) = 1 - (1 - 1/6)^{r-1} > 0.9$

$-(1 - 1/6)^{r-1} > 0.9 - 1$

$(1 - 1/6)^{r-1} < 0.1$ — M 1

$(r-1)\ln\frac{5}{6} < \ln 0.1$ — ~~A~~ 0

$\cancel{r <}$ $r - 1 > \frac{\ln 0.1}{\ln 5/6}$ — M 1

$r > 13.6$ — ~~A~~ 0

$\therefore \; r = 14$

6

Commentary

The candidate correctly answers parts (a) and (b) (i), but trys to go for the quick solution in (b) (ii).The index is wrong, however, (5 instead of 6). This loses all the marks. If the method whereby a geometric progression was summed, 1 or 2 marks could have beeen earned, if a mistake had been made.

Mark Scheme

2(a)	$\mathrm{E}(X) = p + 2pq + 3pq^2 + \ldots$	M1		
	$= p\left(1 + 2q + 3q^2 + \ldots\right)$	A1		
	$= \dfrac{p}{(1-q)^2}$			
	$= \dfrac{p}{p^2}$			
	$= \dfrac{1}{p}$	A1	3	AG (working required)
(b)(i)	6	B1	1	
(ii)	$\mathrm{P}(X \leqslant 6) = \dfrac{\left(\frac{1}{6}\right)\left(1 - \left(\frac{5}{6}\right)^6\right)}{\left(1 - \frac{5}{6}\right)}$	M1		
	$= 0.665$	A1		
	$\mathrm{P}\left(X > 6\right) = 0.335$	A1	3	$\left(\frac{5}{6}\right)^6 = 0.335$ B3 AWRT
(iii)	$1 - \left(\frac{5}{6}\right)^r > 0.9$	M1		
	$\Rightarrow \left(\frac{5}{6}\right)^r < 0.1$	A1		
	$\Rightarrow r > \dfrac{\log 0.1}{\log\left(\frac{5}{6}\right)}$	M1		
	$= 12.6$ $\therefore\ r = 13$	A1	4	CAO
	Total		**11**	

Question 3

3 A geologist is studying the effect of exposure to weather on the radioactivity of granite. He collects, at random, 9 samples of freshly exposed granite and 8 samples of weathered granite. For each sample, he measures the radioactivity, in counts per minute. The results are shown in the table.

	Counts per minute								
Freshly exposed granite	226	189	166	212	179	172	200	203	181
Weathered granite	178	171	141	133	169	173	171	160	

(a) Assuming that these measurements come from two independent normal distributions with a common variance, construct a 95% confidence interval for the difference between the mean radioactivity of freshly exposed granite and that of weathered granite. *(9 marks)*

(b) Comment on a claim that the difference between the mean radioactivity of freshly exposed granite and that of weathered granite is 10 counts per minute. *(2 marks)*

Student Response

3a) $\bar{x} = 192$, $\bar{y} = 162$, $n = 9, 8$, $s_x = 19.7357$, $s_y = 16.349$, $s_x^2 = 389.5$, $s_y^2 = 267.714$ B 1

B 1

$\bar{x} - \bar{y} = 30$

$s^2 = \dfrac{(n-1)s_x^2 + (n-1)s_y^2}{n+n-2}$ M 1

$= \dfrac{(8 \times 389.5) + (7 \times 267.714)}{15}$

$= \dfrac{3164 + 1874}{15}$

$= 335.867$ → 18.3 A 0

$v = 9+8-2 = 15$ 95% = 0.975 = 1.753 × B 1

B 0

$\bar{x}\bar{y} \pm t_{crit} \sqrt{\dfrac{s^2}{n+n}}$

$30 \pm 1.753 \times \sqrt{\dfrac{335.867}{9+8}}$ × M 0

$30 \pm 1.753 \times \sqrt{79.3}$

30 ± 15.61

$(14.39, 45.61)$ ×

3b) 10 is not included in the interval so reject it. 2

Commentary

This question was generally done very well, but this candidate makes a slip with the value for the pooled variance. An incorrect *t*-value is looked up in the tables. Finally the standard error of the differences is incorrect. In all 5 marks are lost.

Mark Scheme

Q	Solution	Marks	Total	Comments
3(a)	$\overline{x_1} = 192 \quad s_1 = 19.736$	B1		Both CAO
	$\overline{x_2} = 162 \quad s_2 = 16.362$	B1		Both (AWRT 19.7, 16.4)
	$s^2 = \dfrac{8 \times 19.736^2 + 7 \times 16.362^2}{8+7}$	M1		
	$= 332.\dot{6} \ \left(= 18.239^2\right)$	A1		AWRT 18.2
	$v = 15$	B1		
	$t = 2.131$	B1		AWRT 2.13
	$192 - 162 = 30$ $\therefore$ 95% CL are:			
	$30 \pm \left(2.13 \times 18.239 \sqrt{\dfrac{1}{9} + \dfrac{1}{8}}\right)$	M1 A1√		√ on t and s^2
	95% CI is (11.1, 48.9)	A1	9	AWRT
(b)	$10 \notin$ CI	E1√		
	$\therefore$ reject claim	e1√	2	√ on (a)
	Total		**11**	

Question 4

4 The lifetimes of electrical components follow an exponential distribution with mean 200 hours.

(a) Calculate the probability that the lifetime of a randomly selected component is:

(i) less than 120 hours; *(2 marks)*

(ii) more than 160 hours; *(2 marks)*

(iii) less than 160 hours, given that it has lasted more than 120 hours. *(3 marks)*

(b) Determine the median lifetime of these electrical components. *(3 marks)*

Student Response

4 (a) $E(x) = 200 \Rightarrow \lambda = \frac{1}{200}$

(i) $P(x < 120) = P(x \leq 119) = 1 - e^{-\frac{1}{200}(119)}$ M 1

A 0

$= 0{,}4484$ X

(ii) $P(x > 160) = 1 - P(x \leq 160) = 1 - \left(1 - e^{-\frac{160}{200}}\right) =$

$= 1 - 1 + e^{-\frac{160}{200}} = 0{,}4493$ 2

Leave blank

(iii) $P(\cancel{\leq}x < 160) = P(x \leq 159) = 1 - e^{-\frac{159}{200}} =$

$= 0.5484$

M 0

(b) $\cancel{\lambda e^{-\lambda x} = 0.5}$

$\cancel{\frac{1}{200} e^{-\frac{1}{200}x} = 0.5}$

$\cancel{e^{-\frac{1}{200}x} = 100}$

$\cancel{-\frac{1}{200}x = \ln 100}$

$\cancel{x = -200 \ln 100}$

$\cancel{x =}$

$P(x \leq n) = 0.5$

$1 - e^{-\frac{n}{200}} = 0.5$

$e^{-\frac{n}{200}} = 0.5$

$-\frac{n}{200} = \ln 0.5$

$n = -200 \ln 0.5$

$n = 138.629$

median lifetime is 138 hours

3

2

Commentary

In part (a) (i) the candidate writes 119 instead of 120., losing a mark. Part (a) (ii) is correct, but like many other students, the candidate cannot calculate a conditional probability in part (a) (iii). What is required is P(120<X<160)/P(X>120), making use of the answers from the previous two parts (see mark scheme below). The candidate makes a good recovery in part (b), however, handling the negative signs and logarithms perfectly.

4(a)(i)	$\mathrm{F}(x)=1-\mathrm{e}^{-\frac{x}{200}}$			May be quoted
	$\mathrm{P}(X<120)=1-\mathrm{e}^{-0.6}$	M1		
	$=0.451$	A1	2	AWRT
(ii)	$\mathrm{P}(X>160)=\mathrm{e}^{-0.8}$	M1		
	$=0.449$	A1	2	AWRT
(iii)	$\mathrm{P}(X<160 \mid X>120)$			
	$=\dfrac{1-[0.4512+0.4493]}{1-0.4512}$	M1 A1		or $=\mathrm{P}(X<40)$ $=1-\mathrm{e}^{-0.2}$
	$=0.181$	A1	3	AWRT
(b)	$1-\mathrm{e}^{-\frac{m}{200}}=0.5$			
	$\Rightarrow \mathrm{e}^{-\frac{m}{200}}=0.5$	M1		
	$\Rightarrow m=\ln 0.5 \times(-200)$	M1		
	$= 139$ hours	A1	3	AWRT
	Total		**10**	

Question 5

5 It is thought that the marks in an examination may be modelled by a triangular distribution with probability density function

$$f(x) = \begin{cases} \frac{1}{1875}x & 0 \leqslant x < 50 \\ \frac{6}{75} - \frac{2}{1875}x & 50 \leqslant x \leqslant 75 \\ 0 & \text{otherwise} \end{cases}$$

(a) Sketch the graph of f. *(2 marks)*

(b) A school enters 60 candidates for the examination. The results are summarised in the table.

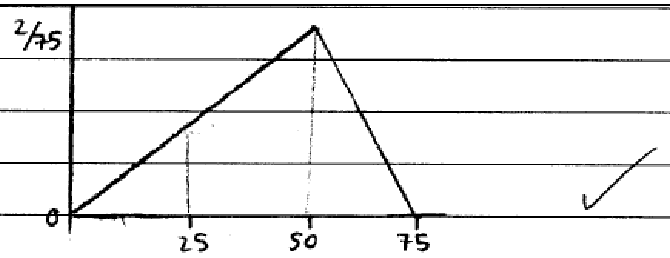
Marks	0–	25–	50–75
Number of candidates	7	28	25

(i) Investigate, at the 5% level of significance, whether the triangular distribution in part (a) is an appropriate model for these data. *(9 marks)*

(ii) Describe, with a reason, how the test procedure in part (b)(i) would differ for a school entering 15 candidates, assuming that its results are summarised using the same mark ranges as in the table above. *(2 marks)*

Student Response

5a.



2

bi. H_0: $X \sim$ Triangular distribution. B 1

H_1: X is not a triangular distribution.

Expected 0-25

$\frac{1}{75} \times 25 = \frac{25}{75} \div 2 = \frac{1}{6} \times 60 = \underline{\underline{10}}$

~~Expected 25-50~~

$\cancel{\frac{2}{75} \times 25 = \frac{50}{75} \div 2 = \frac{1}{3} \times 60 = 20}$

~~Expected 25-50~~

$\cancel{10 + \frac{2}{75} \times 25 = 10 + \left(\frac{50}{75} \times 60\right) = 10 + 40 = 50}$

Expected 25-50

$10 + \left(\frac{1}{75} \times 25 \times 60\right) = 10 + 20 = \underline{\underline{30}}$ ✓ M 1

Expected 50-75

~~[illegible]~~ $30 + 10 = 40$ A 1

$60 - 40 = \underline{\underline{20}}$ ✓ A 1

	0 - 25	25 - 50	50 - 75
O	7	28	25
E	10	30	20
$\frac{(O-E)^2}{E}$	0.9	0.1333	1.25

m 1

$\sum \frac{(O-E)^2}{E} = 2.2833$ ✓ A 1

2 degrees of freedom ✓ B 1

$p = 0.95$

$\chi^2 = 5.991$ ✓ B 1

$5.991 > 2.2833$ ✓

Therefore accept H_0, marks can be modelled by a triangular distribution. ✓ A 1

Expected ~~values~~ value would be less than 5 between 0-25 ~~and~~ ✓ So they would need to be grouped meaning less degrees of freedom and loss accuracy. ✓ 2

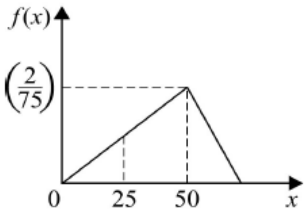
would have to be 1 as only 2 previously.

13

Commentary

The candidate has a clear and accurate diagram. (Many cadidates thought that 50 was halfway between 0 and 75, or worse.) This was marked leniently as only a sketch was required. The candidate uses areas of triangles and a trapezium to find the expected frequencies in part (b) (i), which is easier and quicker than integration, which was often employed. The calculation of chi-squared is accurate, as are the degrees of freedom,the critical value and the conclusion of the test, in context. The comment in part (b) (ii) is correct and gains both marks, some candidates dropped the second mark by not referring to the new number of degrees of freedom.

Mark Scheme

Q	Solution	Marks	Total	Comments
5(a)		B1		Shape
		B1	2	x–scale
(b)(i)	H_0 : triangular distribution fits	B1		
	Areas $\frac{1}{6}, \frac{1}{2}, \frac{1}{3}$	M1 A1		
	O_i 7 28 25 E_i 10 30 20	A1√		
	$\chi^2_{\text{calc}} = \frac{9}{10} + \frac{4}{30} + \frac{25}{20}$	M1		
	$= 2.28$	A1		Accept 2.25 to 2.30
	$v = 3 - 1 = 2$	B1		
	$\chi^2_{\text{crit}} = 5.991$	B1		AWRT 5.99
	$2.28 < 5.991 \Rightarrow$ Accept H_0 Triangular distribution fits data at 5% level of significance	A1√	9	
(ii)	$E_1 < 5 \Rightarrow$ combine classes $\left(\frac{1}{6} \times 15 = 2.5\right)$	M1		
	$v = 2 - 1 = 1$	A1	2	Or gives new χ^2_{calc}
	Total		**13**	

Question 6

6 (a) The IQs of a random sample of 15 students have a standard deviation of 9.1 .

Test, at the 5% level of significance, whether this sample may be regarded as coming from a population with a variance of 225 . Assume that the population is normally distributed. *(6 marks)*

(b) The weights, in kilograms, of 6 boys and 4 girls were found to be as follows.

Boys	53	37	41	50	57	57
Girls	40	46	37	40		

Assume that these data are independent random samples from normal populations.

Show that, at the 5% level of significance, the hypothesis that the population variances are equal is accepted. *(7 marks)*

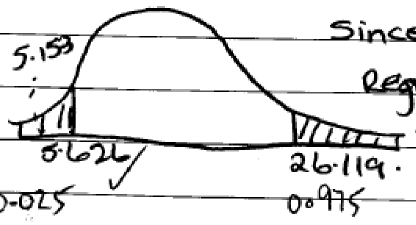
Student Response

6a) $n = 15$ $\sigma = 9.1$

H_0: pop $\sigma^2 = 225$ ✓ 2 tailed 5% B 1

H_1: pop $\sigma^2 \neq 225$ ✓ $\chi^2_{0.975,14} = 5.629$, 26.119.

$\dfrac{14 \times 9.1^2}{225} = \cancel{1.509}\ 5.153$ ✓ M 1 A 1



$\Rightarrow \nu = 14$

Since 5.153 is in the critical region Reject Ho. Significant evidence at the 5% level that the sample B 1 may not be regarded as coming from a population with variance 225. ✓ B 1

A 1

blank

b) H_0: pop $\sigma^2_B = \sigma^2_G$ ✓ 2 tailed 5%

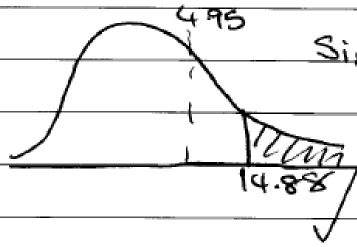
H_1: pop $\sigma^2_B \neq \sigma^2_G$ ✓ B 1

$S_B =$ ~~[illegible]~~ 8.40 n = 6 B 1

$S_G = 3.775$ n = 4

$\frac{8.40^2}{3.775^2} = 4.95$ ✓ M 1 A 1

$F^5_3 = 14.88$



Since 4.95 < 14.88 Accept H_0. B 1

No significant evidence at the 5% level that the population B 1

variances of boys and girls weights are not equal. ✓ A 1

13

Commentary

.In part (a) , testing a population variance, the hypotheses are correctly stated, the critical value and test statistic are clearly displayed and the correct conclusion, in context, is given. Likewise in part (b), testing whether population variances are equal, using an F-test, the same procedure is meticulously carried out and full marks are scored for this question, which is most impressive.

Mark Scheme

Q	Solution	Marks	Total	Comments
6(a)	$H_0 : \sigma^2 = 225 \quad H_1 : \sigma^2 \neq 225$	B1		Both
	$v = 15 - 1 = 14$	B1		
	$\left.\begin{array}{l}\chi^2_{14}(0.025) = 5.629 \\ \chi^2_{14}(0.975) = 26.119\end{array}\right\}$	B1		Both; or $F(\infty, 14) = 2.487$
	$\chi^2 = \dfrac{(n-1)s^2}{\sigma^2} = \dfrac{14 \times 9.1^2}{225} = 5.15$	M1 A1		$F_{calc} = \dfrac{225}{9.1^2} = 2.72$
	$5.15 < 5.629 \Rightarrow$ Reject H_0			$2.72 > 2.487 \Rightarrow$ Reject H_0
	Evidence to suggest that variance is not 225	A1$\surd$	6	
(b)	$H_0 : \sigma_B^2 = \sigma_G^2 \quad H_1 : \sigma_B^2 \neq \sigma_G^2$	B1		Both
	$\left.\begin{array}{l}s_B^2 = 70.567 \\ s_G^2 = 14.25\end{array}\right\}$	B1		Both; or $s_B = 8.400 \quad s_G = 3.7749$
	$F_{calc} = \dfrac{70.567}{14.25} = 4.95$	M1 A1$\surd$		AWRT; $\surd$ on variances
	$v_1 = 5 \quad v_2 = 3$	B1		
	$F_{5,3} = 14.88$	B1		
	$4.952 < 14.88$			
	$\Rightarrow$ Accept H_0			
	Variances are equal	A1$\surd$	7	
	Total		**13**	

Question 7

7 (a) The random variable X has a distribution with unknown mean μ and unknown variance σ^2.

A random sample of size n, denoted by $X_1, X_2, X_3, ..., X_n$, has mean $\overline{X}$ and variance V, where

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{and} \quad V = \left(\frac{1}{n}\sum_{i=1}^{n} X_i^2\right) - \overline{X}^2$$

(i) Show that

$$\mathrm{E}(X_i^2) = \sigma^2 + \mu^2 \quad \text{and} \quad \mathrm{E}(\overline{X}^2) = \frac{\sigma^2}{n} + \mu^2$$ *(3 marks)*

(ii) Hence show that $\dfrac{nV}{n-1}$ is an unbiased estimator for σ^2. *(3 marks)*

(b) A random sample of size 2, denoted by X_1 and X_2, is taken from the distribution in part (a).

Show that $\frac{1}{2}(X_1 - X_2)^2$ is an unbiased estimator for σ^2. *(4 marks)*

Student Response

blank

7. (a) (i) $V = E(X^2) - [E(X)]^2$

$\sigma^2 = E(X^2) - \mu^2$

$E(X^2) = \sigma^2 + \mu^2$ M 1

$V_{(\bar{X})} = E(\bar{X}^2) - [E(\bar{X})]^2$ M 1

$\frac{\sigma^2}{n} = E(\bar{X}^2) - \mu^2$ M 1

$E(\bar{X}^2) = \frac{\sigma^2}{n} + \mu^2$

~~(ii) $E\left(\frac{nV}{n-1}\right)$~~

~~$= \frac{n}{n-1} E(V)$~~

~~$=$~~

~~(ii) $E\left(\frac{nV}{n-1}\right)$~~

~~$= \frac{1}{n-1} E(nV)$~~

~~$= \frac{n}{n-1} E(V)$~~

~~$= \frac{n}{n-1}\left(\frac{1}{n}(\sigma^2+\mu^2) - \left(\frac{\sigma^2}{n}+\mu^2\right)\right)$~~

~~$= \frac{n}{n-1}\left(\frac{\sigma^2}{n} + \frac{\mu^2}{n} - \frac{\sigma^2}{n} - \mu^2\right)$~~

~~$= \frac{n}{n-1}\left(\frac{\mu^2 - n\mu^2}{n}\right)$~~

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

~~$= \mu^2$~~

~~(ii) $E\left(\frac{nV}{n-1}\right)$~~

(ii) $E\left(\frac{nV}{n-1}\right)$

$= \frac{n}{n-1}\left(\frac{1}{n}\cdot n(\sigma^2+\mu^2) - \left(\frac{\sigma^2}{n}+\mu^2\right)\right)$

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

$= \frac{n}{n-1}\left(\frac{n\sigma^2 - \sigma^2}{n}\right)$

$= \frac{\sigma^2(n-1)}{n-1}$ 3

$= \sigma^2$

Leave blank

(b) $E\left(\frac{1}{2}(X_1 - X_2)^2\right)$

$= \frac{1}{2} E(X_1^2 - 2X_2 + X_2^2)$

$= \frac{1}{2}[E(X_1^2) - 2E(X_2) + E(X_2^2)]$

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

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

$= \sigma^2 + \mu^2 - \mu$

$E\left(\frac{1}{2}(X_1 - X_2)^2\right)$

$= \frac{1}{2} E(X_1 - X_2)^2$

$= \frac{1}{2} E(X_1^2 - 2X_1X_2 + X_2^2)$

$= \frac{1}{2}[E(X_1)^2 - 2E(X_1)(X_2) + E(X_2^2)]$

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

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

$= \sigma^2$ ✓

4

10

Commentary

This proved to be the most difficult question on the paper. This candidate is able to rearrange the formulae for both of the preliminary results, which are then needed in part (a) (ii). A number of candidates were able to do part (a) (ii), without first doing part (a) (i). Pleasingly, this candidate did part (a) (ii) after a false start,proving that persistence pays dividends. Likewise a false start was made in part (b), but the question was completely and correctly answered, using the alternative method in the mark scheme. The alternative method proved to be the more common approach, by those who were able to complete this question.

Mark Scheme

Q	Solution	Marks	Total	Comments
7(a)(i)	$\sigma^2 = \mathrm{E}(X_i^{\,2}) - \mu^2$			
	$\Rightarrow \mathrm{E}(X_i^{\,2}) = \sigma^2 + \mu^2$	M1		
	$\mathrm{Var}(\overline{X}) = \mathrm{E}(\overline{X}^2) - \mu^2 = \dfrac{\sigma^2}{n}$	M1M1		
	$\Rightarrow \mathrm{E}(\overline{X}^2) = \dfrac{\sigma^2}{n} + \mu^2$		3	AG
(ii)	$nV = \sum_1^n X_i^{\,2} - n\overline{X}^2$			
	$\Rightarrow \mathrm{E}(nV) = \mathrm{E}\left\{\sum_1^n X_i^{\,2}\right\} - \mathrm{E}(n\overline{X}^2)$	M1		
	$= n(\sigma^2 + \mu^2) - (\sigma^2 + n\mu^2)$	M1		
	$= (n-1)\sigma^2$			
	$\Rightarrow \mathrm{E}\left(\dfrac{nV}{n-1}\right) = \sigma^2$	A1	3	
(b)	$\mathrm{E}(X) = \frac{1}{2}(X_1 + X_2)$			
	$V = \frac{1}{2}(X_1^2 + X_2^2) - \frac{1}{4}(X_1 + X_2)^2$	M1		or $\mathrm{E}\left[\frac{1}{2}(X_1 - X_2)^2\right]$
	$= \frac{1}{4}(X_1^2 - 2X_1X_2 + X_2^2)$			$= \frac{1}{2}\mathrm{E}(X_1^{\,2}) - \mathrm{E}(X_1X_2) + \frac{1}{2}\mathrm{E}(X_2^{\,2})$
	$= \frac{1}{4}(X_1 - X_2)^2$	A1		$= \mathrm{E}(X_1^{\,2}) - \{\mathrm{E}(X_1)\}^2$
	$\dfrac{nV}{n-1} = \dfrac{2}{1} \times \dfrac{(X_1 - X_2)^2}{4}$	M1		$= \sigma^2 + \mu^2 - \mu^2$
	$= \frac{1}{2}(X_1 - X_2)^2$	A1	4	$= \sigma^2 \Rightarrow$ unbiased
	Total		**10**	
	TOTAL		**75**	