OCR
RECOGNISING ACHIEVEMENT

Friday 24 May 2013 – Morning

AS GCE MATHEMATICS (MEI)

4766/01 Statistics 1

QUESTION PAPER

Candidates answer on the Printed Answer Book.

OCR supplied materials:
- Printed Answer Book 4766/01
- MEI Examination Formulae and Tables (MF2)

Other materials required:
- Scientific or graphical calculator

Duration: 1 hour 30 minutes

4715830613

INSTRUCTIONS TO CANDIDATES

These instructions are the same on the Printed Answer Book and the Question Paper.

- The Question Paper will be found in the centre of the Printed Answer Book.
- Write your name, centre number and candidate number in the spaces provided on the Printed Answer Book. Please write clearly and in capital letters.
- **Write your answer to each question in the space provided in the Printed Answer Book**. Additional paper may be used if necessary but you must clearly show your candidate number, centre number and question number(s).
- Use black ink. HB pencil may be used for graphs and diagrams only.
- Read each question carefully. Make sure you know what you have to do before starting your answer.
- Answer **all** the questions.
- Do **not** write in the bar codes.
- You are permitted to use a scientific or graphical calculator in this paper.
- Final answers should be given to a degree of accuracy appropriate to the context.

INFORMATION FOR CANDIDATES

This information is the same on the Printed Answer Book and the Question Paper.

- The number of marks is given in brackets **[ ]** at the end of each question or part question on the Question Paper.
- You are advised that an answer may receive **no marks** unless you show sufficient detail of the working to indicate that a correct method is being used.
- The total number of marks for this paper is **72**.
- The Printed Answer Book consists of **12** pages. The Question Paper consists of **4** pages. Any blank pages are indicated.

INSTRUCTION TO EXAMS OFFICER/INVIGILATOR

- Do not send this Question Paper for marking; it should be retained in the centre or recycled. Please contact OCR Copyright should you wish to re-use this document.

Section A (36 marks)

1 The weights, x grams, of 100 potatoes are summarised as follows.

$$n = 100 \qquad \Sigma x = 24\,940 \qquad \Sigma x^2 = 6\,240\,780$$

(i) Calculate the mean and standard deviation of x. **[3]**

(ii) The weights, y grams, of the potatoes after they have been peeled are given by the formula $y = 0.9x - 15$. Deduce the mean and standard deviation of the weights of the potatoes after they have been peeled. **[3]**

2 Every evening, 5 men and 5 women are chosen to take part in a phone-in competition. Of these 10 people, exactly 3 will win a prize. These 3 prize-winners are chosen at random.

(i) Find the probability that, on a particular evening, 2 of the prize-winners are women and the other is a man. Give your answer as a fraction in its simplest form. **[4]**

(ii) Four evenings are selected at random. Find the probability that, on at least three of the four evenings, 2 of the prize-winners are women and the other is a man. **[4]**

3 The weights of bags of a particular brand of flour are quoted as 1.5 kg. In fact, on average 10% of bags are underweight.

(i) Find the probability that, in a random sample of 50 bags, there are exactly 5 bags which are underweight. **[3]**

(ii) Bags are randomly chosen and packed into boxes of 20. Find the probability that there is at least one underweight bag in a box. **[2]**

(iii) A crate contains 48 boxes. Find the expected number of boxes in the crate which contain at least one underweight bag. **[2]**

4 Martin has won a competition. For his prize he is given six sealed envelopes, of which he is allowed to open exactly three and keep their contents. Three of the envelopes each contain £5 and the other three each contain £1000. Since the envelopes are identical on the outside, he chooses three of them at random. Let £X be the total amount of money that he receives in prize money.

(i) Show that $\mathrm{P}(X = 15) = 0.05$. **[2]**

The probability distribution of X is given in the table below.

r	15	1010	2005	3000
$\mathrm{P}(X = r)$	0.05	0.45	0.45	0.05

(ii) Find $\mathrm{E}(X)$ and $\mathrm{Var}(X)$. **[5]**

5 A researcher is investigating whether people can identify whether a glass of water they are given is bottled water or tap water. She suspects that people do no better than they would by guessing. Twenty people are selected at random; thirteen make a correct identification. She carries out a hypothesis test.

(i) Explain why the null hypothesis should be $p = 0.5$, where p represents the probability that a randomly selected person makes a correct identification. **[2]**

(ii) Briefly explain why she uses an alternative hypothesis of $p > 0.5$. **[1]**

(iii) Complete the test at the 5% significance level. **[5]**

Section B (36 marks)

6 The birth weights in kilograms of 25 female babies are shown below, in ascending order.

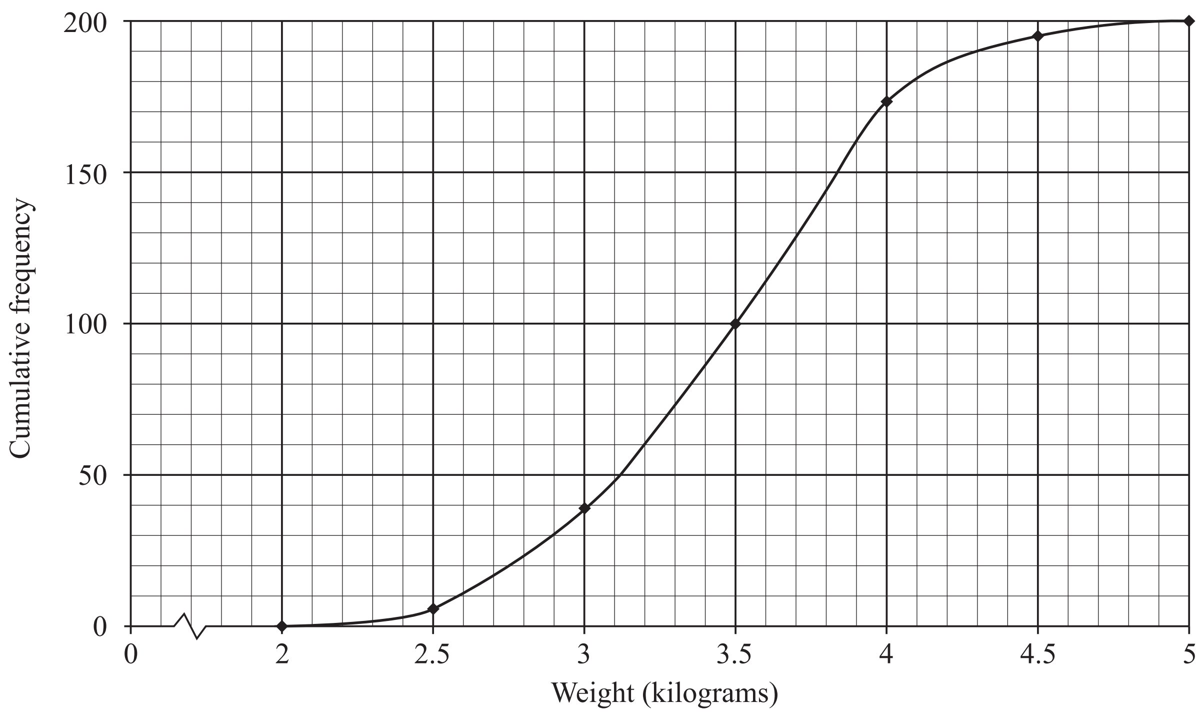
1.39	2.50	2.68	2.76	2.82	2.82	2.84	3.03	3.06	3.16	3.16	3.24	3.32
3.36	3.40	3.54	3.56	3.56	3.70	3.72	3.72	3.84	4.02	4.24	4.34	

(i) Find the median and interquartile range of these data. **[3]**

(ii) Draw a box and whisker plot to illustrate the data. **[3]**

(iii) Show that there is exactly one outlier. Discuss whether this outlier should be removed from the data. **[4]**

The cumulative frequency curve below illustrates the birth weights of 200 male babies.



(iv) Find the median and interquartile range of the birth weights of the male babies. **[3]**

(v) Compare the weights of the female and male babies. **[2]**

(vi) Two of these male babies are chosen at random. Calculate an estimate of the probability that both of these babies weigh more than any of the female babies. **[3]**

7 Jenny has six darts. She throws darts, one at a time, aiming each at the bull's-eye. The probability that she hits the bull's-eye with her first dart is 0.1. For any subsequent throw, the probability of hitting the bull's-eye is 0.2 if the previous dart hit the bull's-eye and 0.05 otherwise.

(i) Illustrate the possible outcomes for her first, second and third darts on a probability tree diagram. **[4]**

(ii) Find the probability that

(A) she hits the bull's-eye with at least one of her first three darts, **[3]**

(B) she hits the bull's-eye with exactly one of her first three darts. **[4]**

(iii) Given that she hits the bull's-eye with at least one of her first three darts, find the probability that she hits the bull's-eye with exactly one of them. **[3]**

Jenny decides that, if she hits the bull's-eye with any of her first three darts, she will stop after throwing three darts. Otherwise she will throw all six darts.

(iv) Find the probability that she hits the bull's-eye three times in total. **[4]**

Question			Answer	Marks	Guidance	
1	(i)		Mean $= \frac{24940}{100} = 249.4$g or 249g	B1	Ignore units	CAO NB 249.40 gets B0 for over-specification
			$Sxx = 6240780 - \frac{24940^2}{100} = 20744$	M1	For Sxx	M1 for 6240780 - 100 × their mean^2 BUT NOTE M0 if their $S_{xx} < 0$
			$s = \sqrt{\frac{20744}{99}} = \sqrt{209.53} = 14.4751 = 14.5$g	A1	CAO ignore units	For s^2 of 210 (or better) allow M1A0 with or without working For RMSD of 14.4 (or better) allow M1A0 provided working seen For RMSD^2 of 207 (or better) allow M1A0 provided working seen Allow 14.48 but NOT 14.47
				[3]		
1	(ii)		New mean = (0.9 × 249.4) – 15 = 209.5g	B1	FT their mean provided answer is positive	If candidate 'starts again' only award marks for CAO Allow 209
			New sd = 0.9 × 14.48 = 13.03g	M1	FT their sd	Or for $0.9^2 \times 14.5^2$
				A1	FT Allow 13.0 to 13.1	Deduct at most 1 mark overall in whole question for over-specification of Mean and 1 mark overall for SD
				[3]		

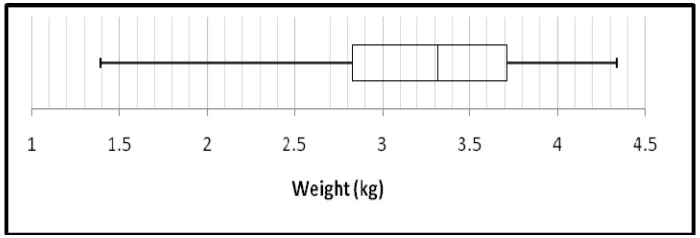
Question			Answer	Marks	Guidance	
2	(i)		$3\times\frac{5}{10}\times\frac{4}{9}\times\frac{5}{8}=\frac{300}{720}=\frac{5}{12}$ = (0.4167)	M1	For 5/10 × 4/9	Correct working but then multiplied or divided by some factor scores M1M1M0A0
				M1	For × 5/8	Zero for binomial Allow M2 for equivalent triple such as $\frac{5}{10}\times\frac{5}{9}\times\frac{4}{8}$
				M1	For 3 × triple product	Or 3 separate equal triplets added
				A1	CAO (Fully simplified)	Answer must be a fraction
				[4]		
			Or $\frac{\binom{5}{2}\times\binom{5}{1}}{\binom{10}{3}}=\frac{10\times5}{120}=\frac{5}{12}$	M1*	For $\binom{5}{2}\times\binom{5}{1}$	Seen
				M1*	For $\binom{10}{3}$	Seen
				M1* dep	For whole fraction	Correct working but then multiplied or divided by some factor scores M1M1M0A0
				A1	CAO (Fully simplified)	
2	(ii)		$4\times\frac{7}{12}\times\left(\frac{5}{12}\right)^3+\left(\frac{5}{12}\right)^4$	M1FT	For first probability	Allow $^4C_3$
				M1FT	For $(5/12)^4$	
				M1FT	For sum of both correct probabilities	Provided sum <1
			$=0.169+0.030=0.199$ Or $=\frac{875}{5184}+\frac{625}{20736}=\frac{1375}{6912}$	A1	CAO Do not allow 0.2, unless fuller answer seen first	Alternative for 1- (P(0)+P(1)+P(2)) allow M1FT for two 'correct' probs, M1 for sum of three 'correct', M1 for 1 – answer, A1 CAO
				[4]		

Question			Answer	Marks	Guidance	
3	(i)		X ~ B(50, 0.1)	M1	For $0.1^5 \times 0.9^{45}$	
			P(5 underweight) = $\binom{50}{5} \times 0.1^5 \times 0.9^{45} = 0.1849$	M1	For $\binom{50}{5} \times p^5 \times q^{45}$	With p **+** q **= 1** Also for $2118760 \times 8.73 \times 10^{-8}$
				A1	CAO	Allow 0.185 or better <u>NB 0.18 gets A0</u>
				[3]		
3	(ii)		X ~ B(20, 0.1) $P(X \geq 1) = 1 - P(X = 0)$ $= 1 - 0.1216 = 0.8784$	M1 A1 **[2]**	For 0.1216 CAO	Allow M1 for 0.9^{20} Allow 0.878 or better See tables at the website <u>http://www.mei.org.uk/files/pdf/formula_book_mf2.pdf</u>
3	(iii)		$E(X) = 48 \times 0.8784 = 42.16$ (= 42.2)	M1		
				A1	FT their probability from part (ii)	If any indication of rounding to 42 or 43 or to another integer on FT allow M1A0 SC1 for 48 × their p giving an integer answer. NB 0.6083 in (ii) leads to 29.20
				[2]		

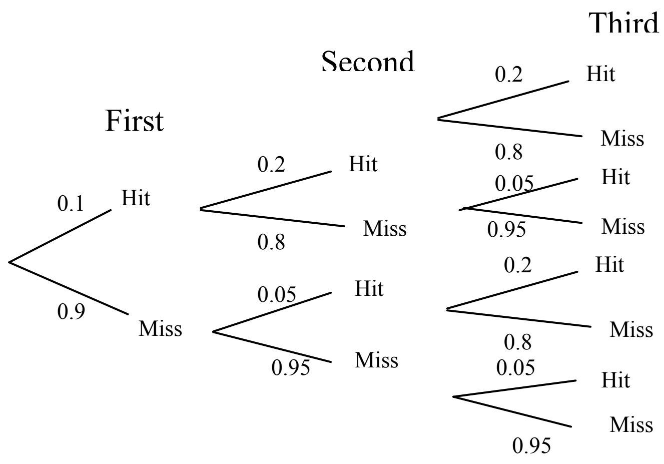
Question			Answer	Marks	Guidance	
4	(i)		$\text{P}(X = 15) = \frac{3}{6}\times\frac{2}{5}\times\frac{1}{4}$	M1	For product of three correct fractions	
			$= = \frac{6}{120} = \frac{1}{20} = 0.05$ **Or** $\frac{1}{{}_6C_3} = \frac{1}{20} = 0.05$ **Or** $\frac{3!\times 3!}{6!} = \frac{1}{20} = 0.05$	A1 **[2]**	NB ANSWER GIVEN NB 1 – (0.45 + 0.45 + 0.05) = 0.05 scores M0A0	Full marks for $3!\times\frac{1}{6}\times\frac{1}{5}\times\frac{1}{4} = \frac{6}{120} = 0.05$ Allow 3×2 in place of 3! SC1 for $6\times\frac{1}{6}\times\frac{1}{5}\times\frac{1}{4} = \frac{6}{120} = 0.05$
4	(ii)		$\text{E}(X) = (15 \times 0.05) + (1010 \times 0.45) + (2005 \times 0.45) + (3000 \times 0.05)$	M1	For Σrp (at least 3 terms correct)	
			= 1507.5 so 1508 (4sf)	A1	CAO	Allow 1507, 1510, 1507.5, 1507.50 or 3015/2
			$\text{E}(X^2) = (15^2 \times 0.05) + (1010^2 \times 0.45) + (2005^2 \times 0.45) + (3000^2 \times 0.05)$ = 2718067.5	M1	For Σr^2p (at least 3 terms correct)	Use of $\text{E}(X-\mu)^2$ gets M1 for attempt at $(x-\mu)^2$ should see $(-1492.5)^2$, $(-497.5)^2$, 497.5^2, 1492.5^2, (if E(X) wrong FT their E(X)) (all 4 correct for M1), then M1 for $\Sigma p(x-\mu)^2$ (at least 3 terms correct with their probabilities) Division by 4 or other spurious value at end gives max M1A1M1M1A0, or M1A0M1M1A0 if E(X) also divided by 4. Unsupported correct answers get 5 marks
			$\text{Var}(X) = 2718067.5 - (1507.5)^2$	M1	dep for – their $\text{E}(X)^2$	
			= 445511.25 so 445500 (4sf)	A1	FT their E(X) provided Var(X) > 0 (and of course $\text{E}(X^2)$ is correct)	Allow 446000
				[5]		

Question			Answer	Marks	Guidance	
5	(i)		Because if people cannot make a correct identification, then the probability that they guess correctly will be 0.5 For 'equally likely to guess right or wrong' or 'two outcomes with equal probability' or '50:50 chance of success' or 'right one in two occasions on average' or 'two (equally likely) outcomes' etc	E1 E1 **[2]**	For idea of a guess or 'chosen at random' For idea of two outcomes	NB The question includes the sentence 'She suspects that people do no better than they would by guessing.', so this on its own does not get the mark for the idea of a guess
5	(ii)		'Because people may do better than they would by guessing' or similar	B1 **[1]**	For idea of selecting correctly /identifying /knowing	No marks if answer implies that it is because there are over half in the sample who make a correct identification
5	(iii)		$P(X \geq 13) = 1 - P(X \leq 12) = 1 - 0.8684 = 0.1316$ **NB PLEASE ANNOTATE THE TOP AND BOTTOM OF THE EXTRA PAGE IF NOT USED**	M1	For notation $P(X \geq 13)$ or $P(X > 12)$ or $1 - P(X \leq 12)$	Notation $P(X = 13)$ scores M0. If they have the correct $P(X \geq 13)$ then give M1 and ignore any further incorrect notation.
				B1*	For 0.1316	Or for 1 – 0.8684 indep of previous mark
			$0.1316 > 0.05$	M1* dep	For comparison with 5%	
			So not significant	A1*		Allow 'accept H_0' or 'reject H_1'
			There is insufficient evidence to suggest that people can make a correct identification.	E1* dep	NB Point probabilities score zero.	Must include 'insufficient evidence' or something similar such as 'to suggest that' ie an element of doubt either in the A or E mark. Must be in context to gain E1 mark. Do not allow 'sufficient evidence to suggest proportion making correct identification is 0.5' or similar

Question			Answer	Marks	Guidance	
			ALTERNATIVE METHOD – follow method above unless some mention of CR seen		Must see some reference to CR to gain any marks	
			Critical region method UPPER TAIL $P(X \geq 14) = 1 - P(X \leq 13) = 1 - 0.9423 = 0.0577 > 5\%$	B1	For either probability	Do not insist on correct notation as candidates have to work out two probabilities for full marks.
			$P(X \geq 15) = 1 - P(X \leq 14) = 1 - 0.9793 = 0.0207 < 5\%$	M1*	For a correct comparison with 5%	Allow comparison in form of statement 'critical region at 5% level is …'
			So critical region is $\{15,16,17,18,19,20\}$	M1* dep	cao dep on the two correct probabilities	No marks if CR not justified Condone $\{15, \ldots 20\}$, $X \geq 15$, oe but not $P(X \geq 15,)$ etc
			13 not in CR so not significant	A1*	Must include '13 not in CR'	Allow 'accept H_0' or 'reject H_1'
			There is insufficient evidence to indicate that people can make a correct identification.	E1* dep on A1	Ignore any work on lower critical region	NB If CR found correctly, **then P(X=13) subsequently found,** but cand says '13 not in CR' then allow up to all five marks. If do not say '13 not in CR' allow no marks
				[5]		

Question			Answer	Marks	Guidance	
6	(i)		Median = 3.32 kg Q1 (= 6.5th value) = 2.83 Q3 (= 19.5th value) = 3.71 Inter-quartile range = 3.71 – 2.83 = 0.88	B1 B1 B1 **[3]**	 For Q1 or Q3 For IQR dep on both quartiles correct	For Q1 allow 2.82 to 2.84 For Q3 allow 3.70 to 3.72 If no quartiles given allow B0B1 for IQR in range 0.86 to 0.90
	(ii)			G1 G1 G1	For reasonably linear scale shown. For boxes in approximately correct positions, with median just to right of centre For whiskers in approximately correct positions in proportion to the box FT their median and quartiles if sensible – guidance above is only for correct values	Dep on attempt at box and whisker plot with at least a box and one whisker. Condone lack of label. Do not award unless RH whisker significantly shorter than LH whisker Allow LH whisker going to 2.5 and outlier marked at 1.39
				[3]		
6	(iii)		Lower limit 2.83 – (1.5 × 0.88) = 1.51 Upper limit 3.71 + (1.5 × 0.88) = 5.03 Exactly one baby weighs less than 1.51 kg and none weigh over 5.03 kg so there is exactly one outlier.	B1 B1 E1*	For 1.51 FT For 5.03 FT Dep on their 1.51 and 5.03	Any use of <u>median</u> ± 1.5 × IQR scores B0 B0 E0 No marks for ± 2 or 3 × IQR In this part FT their values from (i)or (ii) if sensibly obtained but not from location ie 6.5, 19.5 Do not penalise over-specification as not the final answer Do not allow unless FT leads to upper limit above 4.34 and lower limit between 1.39 and 2.50

Question			Answer	Marks	Guidance	
			'Nothing to suggest that this baby is not a genuine data value so she should not be excluded' or 'This baby is premature and therefore should be excluded'.	E1* Dep **[4]**	Any sensible comment in context	For use of mean ± 2sd allow B1 For 3.27 + 2 × 0.62= 4.51 B1 For 3.27 - 2 × 0.62= 2.03 Then E1E1 as per scheme
6	**(iv)**		Median = 3.5 kg Q1 = 50th value = 3.12 Q3 = 150th value = 3.84 Inter-quartile range = 3.84 – 3.12 = 0.72	B1 B1 B1 **[3]**	 For Q1 or Q3 For IQR FT their quartiles	 For Q1 allow 3.11 to 3.13 For Q3 allow 3.83 to 3.85 Dep on both quartiles correct If no quartiles given allow B0B1 for IQR in range 0.70 to 0.74
6	**(v)**		Female babies have lower weight than male babies on the whole Female babies have higher weight variation than male babies	E1 FT E1 FT **[2]**	Allow 'on average' or similar in place of 'on the whole' Allow 'more spread' or similar but not 'higher range' Condone less consistent	Do not allow lower median Do not allow higher IQR, but SC1 for both lower median and higher IQR, making clear which is which
6	**(vi)**		Male babies must weigh more than 4.34 kg			
			Approx 10 male babies weigh more than this.	M1*	For 10 or 9 or 8	Or 200 – 190, 200 –191 or 200 –192
			Probability = $\frac{10}{200}\times\frac{9}{199}=\frac{90}{39800}=\frac{9}{3980}=0.00226$ or $\frac{9}{200}\times\frac{8}{199}=\frac{72}{39800}=0.00181$ or $\frac{8}{200}\times\frac{7}{199}=\frac{56}{39800}=\frac{7}{4975}=0.00141$	M1* dep A1	For first fraction multiplied by any other different fraction (Not a binomial probability) CAO Allow their answer to min of 2 sf	Allow any of these answers For spurious factors, eg 2 × correct answer allow M1M1A0 SC1 for n/200 × (n-1)/199
				[3]		

Question			Answer	Marks	Guidance	
7	(i)			G1 G1 G1 G1 **[4]**	For first set of branches For second set of branches (indep) For third set of branches (indep) For labels	All probabilities correct All probabilities correct All probabilities correct All correct labels for 'Hit' and 'Miss', 'H' and 'M' etc. Condone omission of First, Second, Third. Do not allow misreads here as all FT
7	(ii)	*A*	P(Hits with at least one) = 1 – P(misses with all) = 1 – (0.9 × 0.95 × 0.95) = 1 – 0.81225 = 0.18775	M1* M1* dep A1	For 0.9 × 0.95 × 0.95 For 1 – ans CAO	FT their tree for both M marks, provided three terms 0.188 or better. Condone 0.1877 Allow 751/4000
			ALTERNATIVE METHOD only if there is an attempt to add 7 probabilities At least three correct triple products Attempt to add 7 triple products	M1 M1 A1	CAO	(not necessarily correct triple products)
			FURTHER ALTERNATIVE METHOD 0.1 + 0.9 × 0.05 Above probability + 0.9 × 0.95 × 0.05	M1 M1 A1 **[3]**	CAO	

Question			Answer	Marks	Guidance	
7	(ii)	*B*	P(Hits with exactly one) $= (0.1 \times 0.8 \times 0.95) + (0.9 \times 0.05 \times 0.8) + (0.9 \times 0.95 \times 0.05)$ $= 0.076 + 0.036 + 0.04275 = \frac{19}{250} + \frac{9}{250} + \frac{171}{4000}$ $= \frac{619}{4000} = 0.15475$	M1 M1 M1 A1 **[4]**	For two correct products For all three correct products For sum of all three correct products CAO	FT their tree for all three M marks, provided three terms Allow 0.155 or better
7	(iii)		P(Hits with exactly one given hits with at least one) $= \frac{\text{P}(\text{Hits with exactly one and hits with at least one})}{\text{P}(\text{Hits with at least one})}$ $= \frac{0.15475}{0.18775}$ $= 0.8242$	 M1 M1 A1 **[3]**	 For numerator FT For denominator FT CAO	If answer to (*B*) > than answer to (*A*) then max M1M0A0 Both must be part of a fraction Allow 0.824 or better or 619/751
7	(iv)		P(Hits three times overall) = $(0.1 \times 0.2 \times 0.2) + (0.9 \times 0.95 \times 0.95 \times 0.05 \times 0.2 \times 0.2)$ $= 0.004 + 0.0016245$ $= 0.0056245$	M1 M1 M1* Dep on both prev M1's A1 **[4]**	For $0.1 \times 0.2 \times 0.2$ or 0.004 or 1/250 For $0.9 \times 0.95 \times 0.95 \times 0.05 \times 0.2 \times 0.2$ For sum of both CAO	FT their tree for all three M marks provided three terms in first product and six in second product. Last three probs must be $0.05 \times 0.2 \times 0.2$ unless they extend their tree With no extras Allow 0.00562 or 0.00563 or 0.0056

NOTE RE OVER-SPECIFICATION OF ANSWERS

If answers are grossly over-specified, deduct the final answer mark in every case. Probabilities should also be rounded to a sensible degree of accuracy. In general final non probability answers should not be given to more than 4 significant figures. Allow probabilities given to 5 sig fig.

PLEASE HIGHLIGHT ANY OVER-SPECIFICATION

Please note that there are no G or E marks in scoris, so use B instead

NB PLEASE ANNOTATE EVERY ADDITIONAL ANSWER SHEET EVEN IF FULL MARKS AWARDED OR THE PAGE IS BLANK

Additional notes re Q5 part iii

Comparison with 95% method
If 95% seen anywhere then
M1 for $P(X \leq 12)$
B1 for 0.8684
M1* for comparison with 95% dep on second B1
A1* for not significant oe
E1*

Comparison with 95% CR method
If 95% seen anywhere then
B1 for 0.9423 or 0.9793
M1 for correct comparison with 95%
M1dep for correct CR provided both probs correct
then follow mark scheme for CR method

Smallest critical region method:
Smallest critical region that 13 could fall into is {13, 14, 15, 16, 17, 18, 19, 20} **gets B1 and has size 0.1316 gets B1, This is > 5% gets M1*, A1*, E1* as per scheme**
NB These marks only awarded if 13 used, not other values.

Use of *k* method with no probabilities quoted:
This gets zero marks.

Use of *k* method with one probability quoted:
Mark as per scheme

Line diagram method and Bar chart method
No marks unless correct probabilities shown on diagram, then mark as per scheme..