

**ADVANCED SUBSIDIARY GCE
MATHEMATICS**

Probability & Statistics 1

4732

QUESTION PAPER

Candidates answer on the printed answer book.

OCR supplied materials:

- Printed answer book 4732
- List of Formulae (MF1)

Other materials required:

- Scientific or graphical calculator

**Thursday 26 May 2011
Morning**

Duration: 1 hour 30 minutes

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. 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.
- Give non-exact numerical answers correct to 3 significant figures unless a different degree of accuracy is specified in the question or is clearly appropriate.

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 reminded of the need for clear presentation in your answers.**
- The total number of marks for this paper is **72**.
- The printed answer book consists of **12** pages. The question paper consists of **8** 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 destroyed.

- 1 Five salesmen from a certain firm were selected at random for a survey. For each salesman, the annual income, x thousand pounds, and the distance driven last year, y thousand miles, were recorded. The results were summarised as follows.

$$n = 5 \quad \Sigma x = 251 \quad \Sigma x^2 = 14\,323 \quad \Sigma y = 65 \quad \Sigma y^2 = 855 \quad \Sigma xy = 3247$$

- (i) (a) Show that the product moment correlation coefficient, r , between x and y is -0.122 , correct to 3 significant figures. [3]
- (b) State what this value of r shows about the relationship between annual income and distance driven last year for these five salesmen. [1]
- (c) It was decided to recalculate r with the distances measured in kilometres instead of miles. State what effect, if any, this would have on the value of r . [1]
- (ii) Another salesman from the firm is selected at random. His annual income is known to be £52 000, but the distance that he drove last year is unknown. In order to estimate this distance, a regression line based on the above data is used. Comment on the reliability of such an estimate. [2]
- 2 The orders in which 4 contestants, P , Q , R and S , were placed in two competitions are shown in the table.

Position	1st	2nd	3rd	4th
Competition 1	Q	R	S	P
Competition 2	Q	P	R	S

Calculate Spearman's rank correlation coefficient between these two orders. [5]

- 3 (i) A random variable, X , has the distribution $B(12, 0.85)$. Find
- (a) $P(X > 10)$, [2]
- (b) $P(X = 10)$, [2]
- (c) $\text{Var}(X)$. [2]
- (ii) A random variable, Y , has the distribution $B(2, \frac{1}{4})$. Two independent values of Y are found. Find the probability that the sum of these two values is 1. [4]

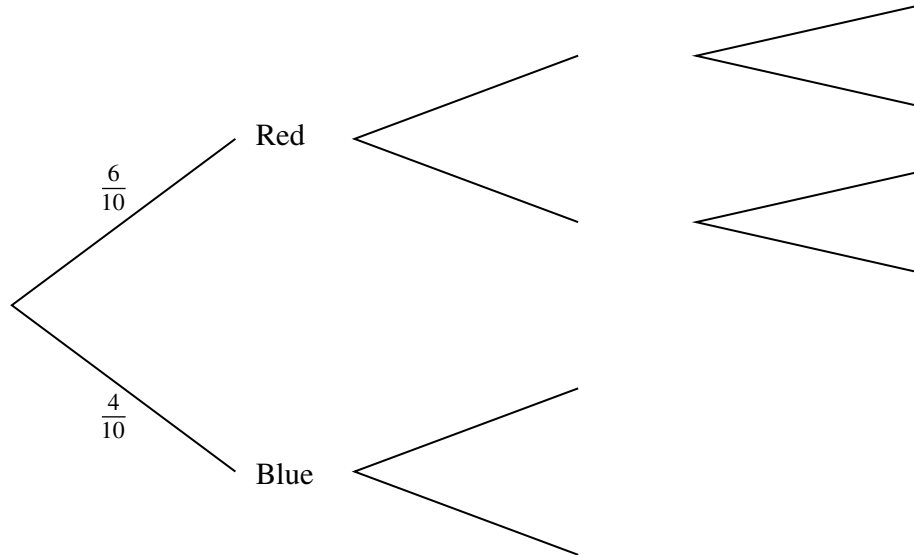
- 4 The table shows information about the time, t minutes correct to the nearest minute, taken by 50 people to complete a race.

Time (minutes)	$t \leq 27$	$28 \leq t \leq 30$	$31 \leq t \leq 35$	$36 \leq t \leq 45$	$46 \leq t \leq 60$	$t \geq 61$
Number of people	0	4	28	14	4	0

- (i) In a histogram illustrating the data, the height of the block for the $31 \leq t \leq 35$ class is 5.6 cm. Find the height of the block for the $28 \leq t \leq 30$ class. (There is no need to draw the histogram.) [3]
- (ii) The data in the table are used to estimate the median time. State, with a reason, whether the estimated median time is more than 33 minutes, less than 33 minutes or equal to 33 minutes. [3]
- (iii) Calculate estimates of the mean and standard deviation of the data. [6]
- (iv) It was found that the winner's time had been incorrectly recorded and that it was actually less than 27 minutes 30 seconds. State whether each of the following will increase, decrease or remain the same:
- (a) the mean, [1]
 - (b) the standard deviation, [1]
 - (c) the median, [1]
 - (d) the interquartile range. [1]

- 5 A bag contains 4 blue discs and 6 red discs. Chloe takes a disc from the bag. If this disc is red, she takes 2 more discs. If not, she takes 1 more disc. Each disc is taken at random and no discs are replaced.

(i) Complete the probability tree diagram in your Answer Book, showing all the probabilities. [2]



The total number of blue discs that Chloe takes is denoted by X .

(ii) Show that $P(X = 1) = \frac{3}{5}$. [2]

The complete probability distribution of X is given below.

x	0	1	2
$P(X = x)$	$\frac{1}{6}$	$\frac{3}{5}$	$\frac{7}{30}$

(iii) Calculate $E(X)$ and $\text{Var}(X)$. [5]

- 6 A group of 7 students sit in random order on a bench.

(i) (a) Find the number of orders in which they can sit. [1]

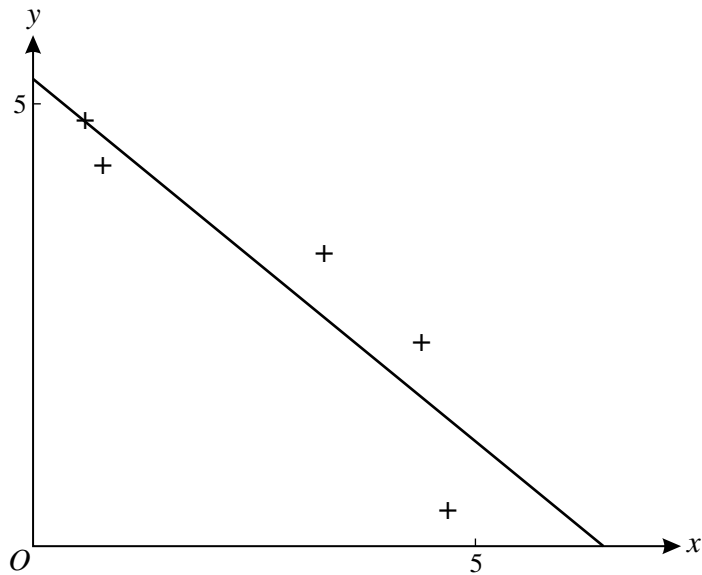
(b) The 7 students include Tom and Jerry. Find the probability that Tom and Jerry sit next to each other. [3]

(ii) The students consist of 3 girls and 4 boys. Find the probability that

(a) no two boys sit next to each other, [2]

(b) all three girls sit next to each other. [3]

- 7 The diagram shows the results of an experiment involving some bivariate data. The least squares regression line of y on x for these results is also shown.



- (i) Given that the least squares regression line of y on x is used for an estimation, state which of x or y is treated as the independent variable. [1]
- (ii) Use the diagram to explain what is meant by 'least squares'. [2]
- (iii) State, with a reason, the value of Spearman's rank correlation coefficient for these data. [2]
- (iv) What can be said about the value of the product moment correlation coefficient for these data? [1]
- 8 Ann, Bill, Chris and Dipak play a game with a fair cubical die. Starting with Ann they take turns, in alphabetical order, to throw the die. This process is repeated as many times as necessary until a player throws a 6. When this happens, the game stops and this player is the winner.

Find the probability that

- (i) Chris wins on his first throw, [1]
- (ii) Dipak wins on his second throw, [3]
- (iii) Ann gets a third throw, [2]
- (iv) Bill throws the die exactly three times. [4]

Note: “(3 sfs)” means “answer which rounds to ... to 3 sfs”. If correct ans seen to ≥ 3 sfs, ISW for later rounding
 Penalise over-rounding only once in paper.

lia	$\frac{3247 - \frac{251 \times 65}{5}}{\sqrt{(14323 - \frac{251^2}{5})(855 - \frac{65^2}{5})}}$ or $\frac{-16}{\sqrt{1722.8 \times 10}}$ $= -0.1219\dots$	M2	M1 for correct subst in any correct S formula M2 for correct subst'n in any correct r formula	or $\frac{-80}{\sqrt{8614 \times 50}}$ Allow -0.1218
b	Poor/no/little/weak/not strong corr'n or rel'nship or link between income & distance oe	A1 3 B1 1	Must see at least 4 sfs or slight neg/weak corr'n (oe) between income & distance In context, ie <u>any</u> comment on income & distance, even if incorrect	eg, Poor neg corr'n, so higher distance, lower income No rel'nship. Low income doesn't cause low distance NOT “Not proportional ...” NOT “negative corr'n ...” No recovery of this mark in (ii)
c	No effect or -0.122 oe	B1 1	eg “Nothing” or “None” oe	Ignore other NOT “Little effect” NOT “Not much effect”
ii	r close to 0, or small, or poor corr'n oe or $r = -0.122$ Unreliable	B1 B1dep 2	or Weak/no corr'n or poor rel'nship oe or No evidence to link sales & distance Condone “innacurate” or “incorrect” or “less reliable” or “not that reliable” “The data is unreliable” Must have correct reason	or because small sample Ignore other Allow: “Unreliable because pts do not fit a st line” “Unreliable because pts are scattered” “Unreliable because not strong neg” “Unreliable because r not close to -1 ” “Unreliable because r smaller than $(-)0.7$ ” NOT “Unreliable because extrapolated”: B0B0 but “Unreliable because extrapolated and poor corr'n”: B1B1
Total		7		

2	<p>Attempt ranks 4 1 2 3 or 1 2 3 4 or 1 2 3 4 oe 2 1 3 4 1 3 4 2 1 4 2 3</p> <p>Σd^2 attempted (or 6) $1 - \frac{6\Sigma d^2}{4(4^2-1)}$ $= \frac{2}{5}$ oe</p>	<p>M1 A1 M1 M1 A1 5</p>	<p>Ignore labels of rows or columns</p> <p>No ranks seen, $d = (0), \pm 1, \pm 1, \pm 2$, or $d^2 = (0), 1, 1, 4$ any order: M1A1</p> <p>NOT $(\Sigma d)^2$</p>	<p>No wking, $\Sigma d^2 = 6$: M1A1M1 No wking, $\Sigma d^2 = \text{eg } 14$: M0A0M0, but can gain 3rd M1</p> <p>No wking, ans $\frac{2}{5}$: Full mks Allow both sets of ranks reversed</p> <p>NB incorrect method: 2 3 4 1 2 1 3 4 OR $d = (0), \pm 2, \pm 1, \pm 3$ any order OR $d^2 = (0), 4, 1, 9$ any order (leading to $\Sigma d^2 = 14$ and $r_s = -\frac{2}{5}$): M0A0M1M1A0</p>
Total		5		
3ia	<p>$(1 - 0.5565)$ or $12 \times 0.85^{11} \times (1 - 0.85) + 0.85^{12}$ $= 0.4435$ or 0.443 or 0.444 (3 sf)</p>	<p>M1 A1 2</p>	<p>or $1 - ((1-0.85)^{12} \dots^{12} C_{10} \times 0.85^{10} (1-0.85)^2)$ ie $1 - (\text{all 11 correct binomial terms})$</p>	<p>or $1 - 0.557$ NB $1 - 0.4435$ (oe): M0A0</p>
b	<p>$0.5565 - 0.2642$ or $^{12}C_{10}(1 - 0.85)^2(0.85)^{10}$ $= 0.2923$ or 0.2924 or 0.292 (3 sf)</p>	<p>M1 A1 2</p>		<p>or $0.557 - 0.264$</p>
c	<p>$12 \times 0.85 \times (1 - 0.85)$ $= 1.53$ oe</p>	<p>M1 A1 2</p>		
ii	<p>$(\frac{3}{4})^2$ AND $\frac{3}{4} \times \frac{1}{4}$ seen (possibly $\times 2$)</p> <p>$(\frac{3}{4})^2 \times 2 \times \frac{3}{4} \times \frac{1}{4}$ oe or $\frac{27}{128}$ or 0.211 $2 \times (\frac{3}{4})^2 \times 2 \times \frac{3}{4} \times \frac{1}{4}$ oe $= \frac{27}{64}$ or 0.422 (3 sfs)</p>	<p>M1 M1 M1 A1 4</p>	<p>eg $(\frac{3}{4})^2 + \frac{3}{4} \times \frac{1}{4}$ or $2 \times (\frac{3}{4})^2 + 2 \times \frac{3}{4} \times \frac{1}{4}$ or $0.5625 + 0.1875$ or $0.5625 + 0.375$</p> <p>or eg 0.5625×0.375</p> <p>Fully correct method</p>	<p>or $\frac{9}{16}$ and $\frac{3}{16}$ or $\frac{9}{16}$ and $\frac{3}{8}$ eg in table or list</p> <p>Allow even if further incorrect wking</p> <p>Ans 0.211: check wking but probably gets M1M1M0A0</p> <p>Use of 0.85 instead of $\frac{1}{4}$: MR max M1M1M1A0</p>
Total		10		

4i	Method is either: Just $4 \div 3$ or $\frac{4}{3}$ or: Use of ratio of correct frequencies AND ratio of widths (correct or 4 and 2)			
4i	$5.6 \times \frac{4}{28} \times \frac{5}{3}$ or $0.8 \times \frac{5}{3}$ or $(5.6 \div \frac{28}{5}) \times \frac{4}{3}$ or $\frac{4}{3}$ or $4 \div 3$ oe $= 1 \frac{1}{3}$ or $\frac{4}{3}$ or 1.33 (3 sf) oe	M2 A1 3	M1 for $5.6 \times \frac{4}{28} \times \frac{4}{2}$ or $0.8 \times \frac{4}{2}$ or $(5.6 \div \frac{28}{4}) \times \frac{4}{2}$ or 0.8×2 oe (= 1.6) No wking, ans 1.3: M2A0 Ans 1.6: Check wking but probably M1M0A0	Correct calc'n using 5.6, 28, 4, 5, 3 oe: M2 Correct calc'n using 5.6, 28, 4, 4, 2 oe: M1 ie fully correct method: M2 or: incorrect class widths, otherwise correct method: M1 $\frac{4}{3}$ correctly obtained (or no wking) then further incorrect: M1M0A0 Use of ratio of widths OR freqs but not both: M0 eg $5.6 \times \frac{4}{28}$ (= 0.8) or $5.6 \times \frac{3}{5}$ (= 3.36): M0 $\frac{4}{2} = 2$: M0M0A0
ii	25 or 26 or 25.5 Med is 21 st (or 22 nd or 21.5 th) in 31-35 class or "25 - 4" Can be implied by calc'n Med > 33 or "more than"	B1 B1 B1 3	or 25 & 26 or med in last ≈ 7 in class or $33 \approx 14^{\text{th}}$ in class or $33 \approx 18^{\text{th}}$ in whole set Can be implied by diagram indep	May be implied, eg by 21 or 22 or 21.5 Calc'ns need not be correct but need to contain relevant figures for gaining B1B1 The " \approx " sign means ± 2 <u>Alternative Method:</u> $33 \approx 18^{\text{th}}$ value B1 More values above 33 than below oe B1 Med > 33 B1 Ignore comment on skew NB Use EITHER the main method OR the <u>Alternative Method</u> (above), not a mixture of the two. Choose the method that gives most marks.

iii	≥ 3 mid-pts attempted $\Sigma fx \div 50$ attempted $(= \frac{1819}{50})$ $= 36.38$ or 36.4 (3 sf) Σfx^2 attempted $(= 68055.5)$ $\sqrt{\frac{68055.5}{50} - (\frac{1819}{50})^2}$ or $\sqrt{1361.11 - 36.38^2}$ $(= \sqrt{37.6056})$ $= 6.13$ (3 sfs) Alt for variance: $\Sigma f(x - \bar{x})^2 (= 1880.28)$ M1 $\sqrt{\frac{1880.28}{50}}$ M1 $= 6.13$ (3 sf) A1	M1 M1 A1 M1 M1 A1 6	seen or implied ≥ 3 terms. or 36 with correct working ≥ 3 terms. completely correct method except midpts & ft their mean, dep not $\sqrt{(\text{neg})}$	Not nec'y correct values (29, 33, 40.5, 53) Allow on boundaries. Not class widths Allow on boundaries. Not class widths (3364, 30492, 22963.5, 11236) Allow class widths for this mark only NB mark is not just for “– mean ² ”, unlike q5(iii) $\Sigma(fx)^2$: M0M0A0 If no wking for Σfx^2 , check using their x and f If no wking or unclear wking: full mks for each correct ans for incorrect ans: $35.8 \leq \mu \leq 36.9$ M0M1A0 $6.0 \leq \text{sd} \leq 6.25$ M1M0A0
iv	(a) Decrease (b) Increase (c) Same (d) Same	B1B1 B1B1 4	Ignore other, eg “slightly” or “probably”	Ignore any comments or reasons, even if incorrect
Total		16		
5	If done with replacement, no marks in any part of this question.			
5i	All correct probs correctly placed, matching labels, if any	B2 2	B1 for 4 correct probs anywhere	Allow B2 with missing labels but only if probs consistently placed, ie R above B throughout
ii	$\frac{4}{10} \times \frac{6}{9} + \frac{6}{10} \times \frac{4}{9} \times \frac{5}{8} + \frac{6}{10} \times \frac{5}{9} \times \frac{4}{8}$ or $\frac{4}{15} + \frac{1}{6} + \frac{1}{6}$ $(= \frac{3}{5} \text{ AG})$	B2 2	B1: two of these products (or their results) added (not multiplied) or $1 - (\frac{6}{10} \times \frac{5}{9} \times \frac{4}{8} + \frac{6}{10} \times \frac{4}{9} \times \frac{3}{8} + \frac{4}{10} \times \frac{3}{9})$ or $1 - (\frac{1}{6} + \frac{1}{10} + \frac{2}{15})$	B1: 1 – two of these products (or results) added (not multiplied) NB incorrect methods can lead to correct ans AG so no wking no mks No ft from tree in (i)

iii	Σxp attempted $= \frac{16}{15}$ oe or 1.07 (3 sfs)	M1 A1	Both non-zero terms $\div 3$ etc or $\frac{1}{\Sigma xp}$: M0	
	$\Sigma x^2 p$ attempted (= $\frac{23}{15}$ or 1.53) $- \frac{(\frac{16}{15})^2}{15}$	M1 M1	Both non-zero terms $\div 3$ etc: or $\frac{1}{\Sigma x^2 p}$: M0 indep but dep +ve result	Not Σxp^2 NB easier to gain than equiv mark in qu 4(iii)
	$= \frac{89}{225}$ oe or 0.395 or 0.396 (3 sfs)	A1 5	Ans 0.388: check wking but probably comes from $\mu = 1.07$; premature rounding: M1M1A0	not 0.395, but check for dot over 5 for recurring
	Alt for Var(X): $\Sigma(x-\bar{x})^2 p$	M2	$\frac{1}{6} \times \frac{16^2}{15} + \frac{3}{5} \times \frac{1^2}{15} + \frac{7}{30} \times \frac{14^2}{15}$ all correct M2, 2 terms correct M1	
Total		9		
6ia	5040	B1 1		
b	$6!$ or $5! \times 6$ or 720 $\div 7!$ or \div "5040" or 1440 or $(5! \text{ or } 6!) \times 2$ $= \frac{2}{7}$ oe or 0.286 (3 sf)	M1 M1 A1 3	$\frac{1}{7} \times \frac{1}{6}$ M1* $\times 6$ or $\times 2$ M1 dep* Any $\div 7!$ or "5040" but NOT any $\times 2$	NOT $6!$ in denom eg $\frac{6!}{5040}$ or $\frac{1}{7}$ or 0.143 or $\frac{1}{21}$ (3 sfs): M1M1A0
iiia	$3! \times 4!$ alone or 144 ($\div 7!$ or "5040") $= \frac{1}{35}$ oe or 0.0286 (3sf)	M1 A1 2	$\frac{4}{7} \times \frac{3}{6} \times \frac{3}{5} \times \frac{2}{4} \times \frac{2}{3} \times \frac{1}{2}$ oe or $\frac{1}{7C3 \text{ or } 7C4}$	Not $3! \times 4! \times \dots$ (eg not $3! \times 4! \times 5$) not $\frac{1}{3! \times 4!}$, not $\frac{1}{144}$ NB no mark for $\div 7!$ or "5040" in this part
b	5 seen or $5!$ seen $3! \times 4! \times 5$ or $5! \times 3!$ or 720 or 5×144 ($\div 7!$ or "5040") $= \frac{1}{7}$ oe or 0.143 (3 sf)	M1 M1 A1 3	$5 \times \frac{3}{7} \times \frac{2}{6} \times \frac{1}{5} (\times \frac{4}{4} \times \frac{3}{3} \times \frac{2}{2})$ oe: M2 or $5 \times \frac{1}{7C3 \text{ or } 7C4}$: M2 or $5 \times$ "(iiia)": M2	or GGGBBBB, BGGGBBB, BBGGGBB, BBBGGGB, BBBBGGG NB no mark for $\div 7!$ or "5040" in this part
Total		9		

7i	x	B1 1	Ignore explanations. “Neither” or “Both”: B0	
ii	Diag showing vertical differences only State that sum of squares of these is min oe	B1 B1 2	Allow description instead of diag: “Distances from pts to line // to y-axis” oe dep vert or horiz lines (not both) drawn or described	Allow \geq one line, from a point to the line Must have Min, Squares, Distances & Sum
iii	-1 Ranks opposite or reversed or <u>perfect</u> neg corr'n between <u>ranks</u> oe	B1 B1dep 2	Not approx -1 As x increases, y decreases	Allow eg: -1 because neg corr'n so ranks must be reversed Ignore other NOT neg corr'n or strong neg rel'nship oe NOT comment about “disagreement” or “agreement”
iv	“Negative” or “Not -1”	B1 1	eg “Strong neg” or any negative value > -1 or “Close to -1”	Any implication of Negative, except NOT “Negative gradient” and NOT “-1” given as the value of r
Total		6		
8	Incorrect p (eg “cubical die means 18 sides hence $p = \frac{1}{18}$ ”): can gain all B & M marks.			
8i	$25/216$ oe or 0.116 (3 sfs)	B1 1		
ii	$(5/6)^7 \times 1/6$ alone = 0.0465 (3 sfs) or $\frac{78125}{1679616}$	M2 A1 3	M1 for $(5/6)^8 \times 1/6$ alone	
iii	$(5/6)^8$ oe alone = 0.233 (3 sfs) or $\frac{390625}{1679616}$	M1 A1 2	$1 - P(X \leq 8)$, with exactly 8 correct terms	NOT $1 - (5/6)^8$, NOT $(5/6)^8 \times \dots$
iv	NB If more than 5 products are added (eg $P(1 \leq X \leq 12)$): no marks			
	$(5/6)^9 \times 1/6 + (5/6)^{10} \times 1/6 + (5/6)^{11} \times 1/6 + (5/6)^{12} \times 1/6$ (= 0.0323 + 0.0268 + 0.0224 + 0.0187) = 0.100 (3 sfs)	M3 A1 4	M3 for all correct or M2 for 3 of these added or these 4 plus 1 extra or 0.0817 or 0.0680 or 0.139 or 0.116 or M1 for ≥ 1 of these terms or values seen; ignore incorrect Allow 0.1 with wking	$(5/6)^9 - (5/6)^{13}$ or $1 - (5/6)^{13} - [1 - (5/6)^9]$ M3 or $(5/6)^{8,9 \text{ or } 10} - (5/6)^{12, 13 \text{ or } 14}$ or $1 - (5/6)^{12, 13 \text{ or } 14} - [1 - (5/6)^{8,9 \text{ or } 10}]$ M2 or $\pm[(5/6)^9 - (1 - (5/6)^{13})]$ or $\pm[1 - (5/6)^9 - (5/6)^{13}]$ M1
Total		10		

Total 72 marks