OCR
RECOGNISING ACHIEVEMENT

Thursday 24 May 2012 – Morning

AS GCE MATHEMATICS

4732 Probability and Statistics 1

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

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. HB pencil may be used for graphs and diagrams only.
- Answer **all** the questions.
- Read each question carefully. Make sure you know what you have to do before starting your answer.
- 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**.
- This 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.

1 For each of the last five years the number of tourists, x thousands, visiting Sackton, and the average weekly sales, £ y thousands, in Sackton Stores were noted. The table shows the results.

Year	2007	2008	2009	2010	2011
x	250	270	264	290	292
y	4.2	3.7	3.2	3.5	3.0

(i) Calculate the product moment correlation coefficient r between x and y. **[4]**

(ii) It is required to estimate the average weekly sales at Sackton Stores in a year when the number of tourists is 280 000. Calculate the equation of an appropriate regression line, and use it to find this estimate. **[4]**

(iii) Over a longer period the value of r is −0.8. The mayor says, "This shows that having more tourists causes sales at Sackton Stores to decrease." Give a reason why this statement is not correct. **[1]**

2 The masses, x kg, of 50 bags of flour were measured and the results were summarised as follows.

$$n = 50 \qquad \Sigma(x - 1.5) = 1.4 \qquad \Sigma(x - 1.5)^2 = 0.05$$

Calculate the mean and standard deviation of the masses of these bags of flour. **[6]**

3 The test marks of 14 students are displayed in a stem-and-leaf diagram, as shown below.

0	
1	2 6
2	1 3 5
3	w x 4 8 y z
4	6 7 7

Key: 1 \| 6 means 16 marks

(i) Find the lower quartile. **[1]**

(ii) Given that the median is 32, find the values of w and x. **[2]**

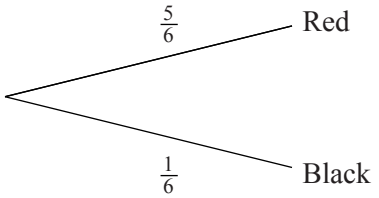
(iii) Find the possible values of the upper quartile. **[2]**

(iv) State one advantage of a stem-and-leaf diagram over a box-and-whisker plot. **[1]**

(v) State one advantage of a box-and-whisker plot over a stem-and-leaf diagram. **[1]**

4 A bag contains 5 red discs and 1 black disc. Tina takes two discs from the bag at random without replacement.

(i) The diagram shows part of a tree diagram to illustrate this situation.



Complete the tree diagram in your Answer Book showing all the probabilities. **[2]**

(ii) Find the probability that exactly one of the two discs is red. **[3]**

All the discs are replaced in the bag. Tony now takes three discs from the bag at random without replacement.

(iii) Given that the first disc Tony takes is red, find the probability that the third disc Tony takes is also red. **[2]**

5 **(i)** Write down the value of Spearman's rank correlation coefficent, r_s, for the following sets of ranks.

(a)

Judge *A* ranks	1	2	3	4
Judge *B* ranks	1	2	3	4

[1]

(b)

Judge *A* ranks	1	2	3	4
Judge *C* ranks	4	3	2	1

[1]

(ii) Calculate the value of r_s for the following ranks.

Judge *A* ranks	1	2	3	4
Judge *D* ranks	2	4	1	3

[3]

(iii) For each of parts **(i)(a)**, **(i)(b)** and **(ii)**, describe in everyday terms the relationship between the two judges' opinions. **[3]**

6 A six-sided die is biased so that the probability of scoring 6 is 0.1 and the probabilities of scoring 1, 2, 3, 4, and 5 are all equal. In a game at a fête, contestants pay £3 to roll this die. If the score is 6 they receive £10 back. If the score is 5 they receive £5 back. Otherwise they receive no money back. Find the organiser's expected profit for 100 rolls of the die. **[5]**

7 **(i)** 5 of the 7 letters A, B, C, D, E, F, G are arranged in a random order in a straight line.

(a) How many different arrangements of 5 letters are possible? **[2]**

(b) How many of these arrangements end with a vowel (A or E)? **[3]**

(ii) A group of 5 people is to be chosen from a list of 7 people.

(a) How many different groups of 5 people can be chosen? **[1]**

(b) The list of 7 people includes Jill and Jo. A group of 5 people is chosen at random from the list. Given that either Jill and Jo are both chosen or neither of them is chosen, find the probability that both of them are chosen. **[3]**

8 **(i)** The random variable X has the distribution B(30, 0.6). Find $P(X \geqslant 16)$. **[2]**

(ii) The random variable Y has the distribution B(4, 0.7).

(a) Find $P(Y = 2)$. **[2]**

(b) Three values of Y are chosen at random. Find the probability that their total is 10. **[6]**

9 **(i)** A clock is designed to chime once each hour, on the hour. The clock has a fault so that each time it is supposed to chime there is a constant probability of $\frac{1}{10}$ that it will not chime. It may be assumed that the clock never stops and that faults occur independently. The clock is started at 5 minutes past midnight on a certain day. Find the probability that the first time it does not chime is

(a) at 0600 on that day, **[3]**

(b) before 0600 on that day. **[3]**

(ii) Another clock is designed to chime twice each hour: on the hour and at 30 minutes past the hour. This clock has a fault so that each time it is supposed to chime there is a constant probability of $\frac{1}{20}$ that it will not chime. It may be assumed that the clock never stops and that faults occur independently. The clock is started at 5 minutes past midnight on a certain day.

(a) Find the probability that the first time it does not chime is at either 0030 or 0130 on that day. **[2]**

(b) Use the formula for the sum to infinity of a geometric progression to find the probability that the first time it does not chime is at 30 minutes past some hour. **[3]**

Question			Answer	Marks	Guidance	
1	(i)		$\Sigma x = 1366$ $\Sigma y = 17.6$ $\Sigma x^2 = 374460$ $\Sigma y^2 = 62.82$ $\Sigma xy = 4784.8$ $S_{xx} = 374460 - \frac{1366^2}{5}$ or 1268.8 $S_{yy} = 62.82 - \frac{17.6^2}{5}$ or 0.868 $S_{xy} = 4784.8 - \frac{1366\times17.6}{5}$ or –23.52 $r = \frac{-23.52}{\sqrt{1268.8\times0.868}}$ or $\frac{-23.52}{33.186...}$ oe = –0.709 (3 sfs)	B1 M1 M1 A1 **[4]**	any three correct; may be implied by 2 *S's* correct sub in any correct *S* formula, ft Σs, $\bar{x}$, $\bar{y}$ corr sub into 3 *S*s and *r*, ft Σs, $\bar{x}$, $\bar{y}$ cao	OR, using $S_{xx} = \Sigma(x - \bar{x})^2$ etc: $\bar{x} = \frac{1366}{5}$ or 273.2, $\bar{y} = \frac{17.6}{5}$ or 3.52, either:B1 $(-23.2)^2 + (-3.2)^2 + (-9.2)^2 + 16.8^2 + 18.8^2$ $0.68^2 + 0.18^2 + (-0.32)^2 + (-0.02)^2 + (-0.52)^2$ $(-23.2)\times0.68 + (-3.2)\times0.18 + (-9.2)\times(-0.32) + 16.8\times(-0.02) + 18.8\times(-0.52)$ If no working seen: –0.71: SC 3; –0.7: SC 1
1	(ii)		$b = \frac{"-23.52"}{"1268.8"}$ or $-\frac{147}{7930}$ or –0.0185 (3 sfs) $y - \frac{"17.6"}{5} =$ "–0.0185"$(x - \frac{"1366"}{5})$ $\Rightarrow y = -0.019x + 8.6$ or better, ie 2 sfs enough ($y = -0.019\times280 + 8.6$ (= 3.39 to 3.41)) Est sales = £3390 to £3410 or 3.39 thousand to 3.41 thousand	M1 M1 A1 A1ft **[4]**	ft their S_{xy} & S_{xx} & Σs from (i) or $a = \frac{"17.6"}{5} -$ "(–0.0185)" $\times \frac{"1366"}{5}$ if *a* incorrect, must see method for M1 cao; must be "$y = \ldots$" coeffs that round to –0.019 & 8.6 to 2 sfs ft their $y\times1000$, dep M1M1, dep sub 280 (not 280000) Allow "k" for thousand No working, ans in range: M1M1A0A1	use of *x* on *y* line: $b' = \frac{"-23.52"}{"0.868"}$ (or –27.1) M0 $x - \frac{"1366"}{5} =$ "–27.1"$(y - \frac{"17.6"}{5})$ or $a' = \frac{"1366"}{5} -$ "(–27.1)" $\times \frac{"17.6"}{5}$) M1 (if a' incorrect, must see method for M1) $x = -27.1y + 369$ cao A1 3277 or 3278 A0
1	(iii)		There may be other factors oe Correlation does not imply causation oe	 B1 **[1]**	or any suggestion of another factor that could be involved, eg Depends on state of the economy oe Must state or clearly imply: EITHER corr'n does not imply causation OR there could be another factor involved Ignore all else	NOT: Tourists & sales not nec'y linked Sales are not entirely dep on tourists Could be a coincidence Might be different other years More tourists wd incr sales –0.8 is not strong corr'n Only shows good neg corr'n Sample is small Could be affected by extremes Neg corr'n not nec'y imply neg relnship

Question			Answer	Marks	Guidance	
2			$\frac{1.4}{50}$ (= 0.028) $1.5 + \frac{1.4}{50}$ = 1.528 or $\frac{191}{125}$ or 1.53 (3 sf) $\frac{0.05}{50} - (\frac{1.4}{50})^2$ or 0.000216 seen √0.000216 = 0.0147 (3 sf)	M1 M1 dep M1 A1 M1 M1 A1 **[6]**	1.4 + 50×1.5 (= 76.4) $\frac{'76.4'}{50}$ ($\Sigma x^2 - 2\times1.5\times'76.4' + 50\times1.5^2 = 0.05$) ($\Rightarrow \Sigma x^2 = 116.75$; no marks yet) $\frac{0.05 + 2\times1.5\times'76.4' - 50\times1.5^2}{50}$ – '1.528'^{2} all correct fully correct method, ie nothing added etc cao not isw	eg $\frac{1.4+1.5}{50}$ M0M0A0 not $\frac{0.05}{50}$ – '1.528'^{2}
3	(i)		23	B1 **[1]**	Allow 22.5	NOT 22 (ie 3.5th no) Correct ans is the 4th or 3.75th no.
3	(ii)		0 0	B1 B1 **[2]**	B1 for 30, 30	
3	(iii)		38 or 40 39 40.75	B2 **[2]**	B1 for 38 or 39 seen B2 for 38 & 39 seen alone, not in a range Mixture, eg 38, 40.75 B1B0 3/8 and 3/9 (both): B1B0 8 and 9(both): B1B0 40, 40.75: similar scheme as for 38, 39	eg 38, 38.5, 39 B1B0 (ie UQ = $^3/_4 \times 14 = 10.5^{th}$ no.) 'Between 39 & 46' B1B0 38 ≤ any letter < 40 B1B0 SC 42, 42.5 only B1B0 (ie UQ = 11.5th no.) Correct ans are the poss 11th or 11.25th nos

Question			Answer	Marks	Guidance	
3	(iv)		Shows all the data or you can see all the values oe You can see the actual/exact/indiv numbers/values/results No data is lost oe Shows the shape of the distribution oe Can perform calculations of your choice (eg mean) Shows which group (or class, NOT value) has the highest frequency (or is the mode) oe	 B1 **[1]**	any implication of <u>all</u> the data or the <u>actual</u> numbers/values/results or similar eg Can compare each indiv result Easier to see the numbers eg can <u>find</u> frequencies No mks for ans to (v) given in (iv) unless labelled as (v)	NOT Shows the spread/skew/trend Any comment on skew You can <u>see</u> the actual frequ's Easier to compare sets of data Shows more info or more data Easier to read off the data Ignore all other
3	(v)		Shows the median or it's easier to see the median (or quartiles or IQR) It can measure the middle 50% easily	 B1 **[1]**	eg Shows mean and quartiles B1 Shows range and median B1 No mks for ans to (v) given in (iv) unless labelled as (v) Ignore all other	NOT Shows the spread/skew/trend Can see data in diag form Shows max or min or range Easier to compare sets of data Not affected by outliers Easy to see outliers Shows s.d. or shows mean Can see important data items/measures
4	(i)		Top: 2 branches $\frac{4}{5}$, $\frac{1}{5}$ & R, B shown Bottom: 1st branch: prob = 1 or $\frac{5}{5}$, & R shown no 2nd branch OR branch with prob = 0 or $\frac{0}{5}$	B1 B1 **[2]**	consistent allow eg $\frac{4}{4}$ ignore any 3rd layer branches	 Any missing label(s) on first three branches, subtr B1 once No label needed on zero branch, if drawn.

Question			Answer	Marks	Guidance	
4	(ii)		$\frac{5}{6}\times\frac{1}{5}$ or $\frac{1}{6}(\times 1)$ or $\frac{1}{6}$ seen $\frac{5}{6}\times\frac{1}{5} + \frac{1}{6}(\times 1)$ $= \frac{1}{3}$ oe	M1 M1 A1 **[3]**	 all correct cao	or $1 - \frac{5}{6}\times\frac{4}{5}$ or $1 - \frac{2}{3}$ M2 ft incorrect tree dep probs ≤ 1 if 3rd tree prob = 1, (ii)M1M1A0 if 3rd tree prob $\neq 1$, (ii)M1M0A0 NB!! $2\times\frac{5}{6}\times\frac{1}{5} = \frac{1}{3}$ M1M0A0
4	(iii)		$\frac{4}{5}\times\frac{3}{4} + \frac{1}{5}(\times 1)$ or $1 - \frac{4}{5}\times\frac{1}{4}$ or 1 – 0.2 all correct $= \frac{4}{5}$ or 0.8 oe	M1 A1 **[2]**	or $(\frac{5}{6}\times\frac{4}{5}\times\frac{3}{4}+\frac{5}{6}\times\frac{1}{5})\div\frac{5}{6}$ all correct May be seen without working M1A1 cao	but $\frac{5}{6}\times(\frac{4}{5}\times\frac{3}{4}+\frac{1}{5})$ M0 ft incorrect tree: (iii) M1A0
5	(i)	(a)	1	B1 **[1]**		NOT close to 1
5	(i)	(b)	–1	B1 **[1]**		NOT close to –1
5	(ii)		Σd^2 attempted (= 10) $1 - \frac{6\times\Sigma d^2}{4(4^2-1)}$ = 0	M1 M1 A1 **[3]**	if Σd^2 = 10, may be implied by next line if $\Sigma d^2 \neq 10$, must see working dep M1 Use of $(\Sigma d)^2$ M0M0A0	S_{xx} or $S_{yy} = 30 - \frac{100}{4}$ (= 5) or $S_{xy} = 25 - \frac{100}{4}$ (= 0) M1 $\frac{0}{\sqrt{5\times 5}}$ M1

Question			Answer	Marks	Guidance	
5	(iii)		No ft from (i)(a), (i)(b) & (ii)			NOT:
			ia: Total (or perfect or max or complete)agreement They have the same opinions/ranks/numbers etc They were identical	B1	Identical opinions/views/marks/ranks/ decisions/results/numbers oe Agree on all the ranks	They agree or Strongly agree They agree most ranks Similar rankings As A's ranks increase so do B's Perfect relnship
			ib: Opposite/reverse opinions/views/marks/ranks/ decisions/results oe	B1	Total (or max or complete or perfect) disagreement A's highest is B's lowest oe "Opposite" seen is sufficient	NOT: Don't agree any ranks Disagree or Strongly disagree Disagree on all ranks Perfect neg relnship
			ii: For $r = 0$ must state or imply:			NOT:
			either NO relationship or similar		No relationship/pattern/link/similarity between opinions/views/marks/ranks/ decisions/results oe opinions/etc… not related scoring appears random	Different views Don't agree but some rel'nshp Ranks all different No corr'n betw judges' views Don't agree nothing in common at all
			or indicate BOTH agreement & disagreement or NEITHER agree nor disagree		Neither agree nor disagree oe Both agree & disagree oe Agree for some, disagree for others oe mixed/varied opinions on the ranks	not much in common completely different orders opinions completely different half way between (a) and (b)
			or DIFFERENT but NOT OPPOSITE	B1		
				[3]	All three parts: Must refer to (or imply) opinions/views/marks/ranks/scores or (dis)agreement, or relationship or pattern oe, NOT just corr'n	Ignore all other

Question			Answer	Marks	Guidance	
6			$(1 - 0.1) \div 5$ (= 0.18) 3×0.18 or 2×0.18 or 7×0.1 (or result of these)(poss × 100) (3 × 0.18 only scores if using £3, not score of 3. Similarly for 2 × 0.18) . $4 \times 3 \times 0.18$ AND $2 \times 0.18 + 7\times0.1$ (poss × 100) (or 2.16 AND 1.06 or 216 AND 106) '2.16' – '1.06' or '216' – '106' must be attempt gain on 1,2,3,4 – loss on 5,6 E(profit for 100 rolls) = (£)110	M1 M1 M1 M1 dep any M1 A1 **[5]**	can be implied, eg by 18 5×0.18 or 10×0.1(or result of these)(poss × 100) 3 AND 5 × 0.18+10 × 0.1 (poss × 100) (or 3 AND 1.9 or 300 AND 190) 3 – '1.9' or 300 – '190' must be attempt receipt – payout on 5,6 E(profit for 100 rolls) = (£)110 NB 300–(0.1×300+0.18×300) = 300–84 =216 M1M1M0M0A0	or, using exp no. of 5's & 6's 18 × 5 or 10 × 10 300 AND 18 × 5 + 10 ×10 (NB 300+ 100×0.18 +100×0.1 is insuff) Eg: 300–100×(5×0.18+**6**×0.1)=150 M1M1M0M1A0 Mark one method only Must be matched pair eg 300–106 or 216–190: M1M1M0M0A0
7	(i)	(a)	$^7P_5$ or $\frac{7!}{2!}$ or $7\times6\times5\times4\times3$ or $^7C_5 \times 5!$ alone = 2520	M1 A1 **[2]**	$^7P_2$ or $\frac{7!}{2!}$ M0A0	$^7C_5 = 21$ or 5! = 120 M0A0 but see (i)(b)
7	(i)	(b)	$^6P_4$ or $\frac{6!}{2!}$ or $6\times5\times4\times3$ or $^6C_4 \times 4!$ or 360 × 2 (see middle column) = 720	M1 M1 A1 **[3]**	alone or ×2 only $^6P_4 \times 2$ or 6! alone M2 $^6C_4 \times 2$ or 6! × 2 alone M0M1 only any other × 2 M0M0 or '2520'× $\frac{2}{7}$ M2A0 (eg (ia)21 (ib) $21\times\frac{2}{7} = 6$ M2A0 but if ans is 6, must see wking) cao	or '2520' – 5 × $^6P_4$ M2 SC ONLY on ft from (i)(a): if (i)(a) 5! = 120, then (i)(b)4!×2=48 alone M1M0A0 Other SC $^5P_3 \times 2$ M2 (from a vowel at <u>each</u> end, ie treat as MR) NOT isw eg $\frac{720}{'2520'} = \frac{2}{7}$ M1M1A0
7	(ii)	(a)	21	B1 **[1]**		

Question			Answer	Marks	Guidance	
7	(ii)	(b)	${}^5C_3$ or $\frac{5!}{3!2!}$ or ${}^5C_5$ seen or 10 seen in num	M1	$\frac{5}{7}\times\frac{4}{6}$ oe seen	Allow ${}^5C_2$ seen BOD
			$\frac{{}^5C_3}{{}^5C_3+{}^5C_5}$ oe	M1	$\frac{5}{7}\times\frac{4}{6}\div(\frac{5}{7}\times\frac{4}{6}+\frac{2}{7}\times\frac{1}{6})$	
			$\frac{10}{11}$ or 0.909 (3 sf)	A1		
				[3]		
8	(i)		1 – 0.1754 alone = 0.825 (3 sfs)	M1 A1	Allow 1– 0.2855 or 0.7145 or 0.715 alone	
				[2]		
8	(ii)	(a)	${}^4C_2 \times 0.7^2 \times 0.3^2$ = $\frac{1323}{5000}$ or 0.265 (3 sf)	M1 A1	All correct	
				[2]		
8	(ii)	(b)	4,4,2 & 4,3,3 only, seen or implied	B1	Both needed	
			$P(Y = 4) = 0.7^4$ (or $\frac{2401}{10000}$ or 0.2401)	M1		
			$P(Y = 3) = 4 \times 0.3 \times 0.7^3$ (or $\frac{1029}{2500}$ or 0.4116)	M1		
			P(4,3,3) = 3 × "0.2401" × "0.4116"^{2} (or 0.122)	M1	ie 3 × their P(4) × (their P(3))^2	if "3×" omitted twice or "3!×" used twice allow M1M0
			P(4,4,2) = 3 × 0.2401"^{2} × "0.265" (or 0.0458)	M1	ie 3 × (their P(4))^2 × their P(2) ft (ii)(a) For M mks ignore extra combs eg P(4,4,3)	eg ans 0.0560, 0.0559,0.336, probably B1M1M1M1M0A0 but must see method
			P(Tot = 10) = 0.168 (3 sfs)	A1	If B(30, 0.6) <u>clearly</u> being used: Any 5 combs adding to 10 seen B1 $P(8) = {}^{30}C_8\times0.4^{22}\times0.6^8$ or 0.0002 $P(9) = {}^{30}C_9\times0.4^{21}\times0.6^9$ or 0.0007 $P(10) = {}^{30}C_{10}\times0.4^{20}\times0.6^{10}$ or 0.0020 all three correct M2 or two correct M1 No more marks	
				[6]		

Question			Answer	Marks	Guidance	
9	(i)	(a)	Geo stated or implied $0.9^5 \times 0.1$ alone = 0.059(0…) (2 sfs)	M1 M1 A1 **[3]**	eg by $0.9^p \times 0.1$ or $0.1^p \times 0.9$ alone, $p>1$ all correct	
9	(i)	(b)	0.9^5 or 0.59… (NB cf ans to (i)(a)!!) $1 - 0.9^5$ = 0.4095 or 0.410 (3 sfs)	M1 M1 A1 **[3]**	$0.1 + 0.9\times0.1 + \ldots 0.9^4\times0.1$: M2 1 term wrong or omit or extra or 1 – (all terms correct): M1 or $1 - 0.9^6$: M1	M0M0A0 for $0.9^p \times 0.1$
9	(ii)	(a)	$0.05 + 0.95^2\times 0.05$ = $\frac{761}{8000}$ or 0.0951 (3 sfs)	M1 A1 **[2]**	All correct	NB!! $2 \times 0.95 \times 0.05 = 0.095$ M0A0
9	(ii)	(b)	0.05, $0.95^2\times 0.05$, … or $\frac{1}{20}$, $\frac{361}{8000}$, … oe $\frac{0.05}{1-0.95^2}$ or $\frac{0.05}{1-0.9025}$ oe = $\frac{20}{39}$ or 0.513 (3 sfs)	M1 M1 A1 **[3]**	$\geq$ 2 terms. Not nec'y added May be implied by next line or $\frac{0.05}{1-(1-0.5)^2}$ or $\frac{0.05}{2\times0.05-0.05^2}$ or $\frac{1}{1.95}$ oe	or $r = 0.95^2$ stated or implied NB $\frac{0.05}{1-0.5\times0.05} = 0.0513$ M0A0

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