Paper Reference(s)

6683/01

Edexcel GCE

Statistics S1

Advanced/Advanced Subsidiary

Tuesday 10 June 2014 – Morning

Time: 1 hour 30 minutes

Materials required for examination
Mathematical Formulae (Pink)

Items included with question papers
Nil

Candidates may use any calculator allowed by the regulations of the Joint Council for Qualifications. Calculators must not have the facility for symbolic algebra manipulation, differentiation and integration, or have retrievable mathematical formulas stored in them.

Instructions to Candidates

In the boxes on the answer book, write the name of the examining body (Edexcel), your centre number, candidate number, the unit title (Statistics S1), the paper reference (6683), your surname, other name and signature.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
This paper has 8 questions.
The total mark for this paper is 75.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You must show sufficient working to make your methods clear to the Examiner.
Answers without working may not gain full credit.

1. A random sample of 35 homeowners was taken from each of the villages Greenslax and Penville and their ages were recorded. The results are summarised in the back-to-back stem and leaf diagram below.

Totals	Greenslax								Penville											Totals
(2)							8	7	2	5	5	6	7	8	8	9				(7)
(3)						9	8	7	3	1	1	1	2	3	4	4	5	6	7	9 (11)
(4)					4	4	4	0	4	0	1	2	4	7						(5)
(5)				6	6	5	2	2	5	0	0	5	5	5						(5)
(7)		8	6	5	4	2	1	1	6	2	5	6	6							(4)
(8)	8	6	6	6	4	3	1	1	7	0	5									(2)
(5)				9	8	4	3	2	8											(0)
(1)								4	9	9										(1)

Key: 7 | 3 | 1 means 37 years for Greenslax and 31 years for Penville

Some of the quartiles for these two distributions are given in the table below.

	Greenslax	**Penville**
Lower quartile, Q_1	a	31
Median, Q_2	64	39
Upper quartile, Q_3	b	55

(*a*) Find the value of a and the value of b.

(2)

An outlier is a value that falls either

more than $1.5 \times (Q3 - Q1)$ above $Q3$

or more than $1.5 \times (Q3 - Q1)$ below $Q1$

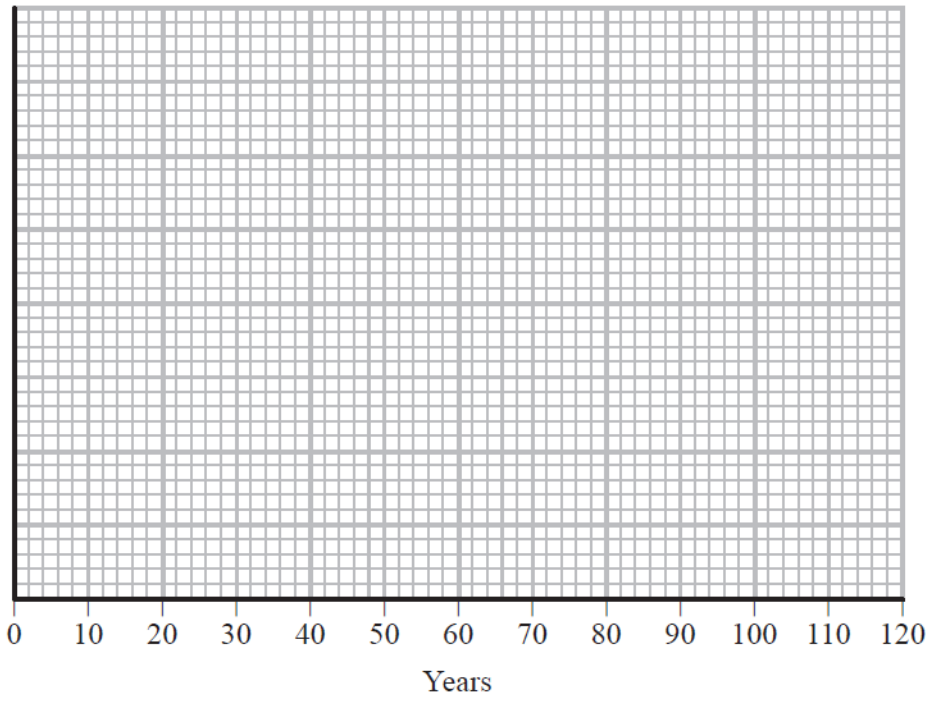
(*b*) On the graph paper on the next page draw a box plot to represent the data from Penville. Show clearly any outliers.

(4)

(*c*) State the skewness of each distribution. Justify your answers.

(3)

Question 1(*b*) graph paper



---

2. The mark, x, scored by each student who sat a statistics examination is coded using

$$y = 1.4x - 20$$

The coded marks have mean 60.8 and standard deviation 6.60.

Find the mean and the standard deviation of x.

(4)

---

3. The table shows data on the number of visitors to the UK in a month, v (1000s), and the amount of money they spent, m (£ millions), for each of 8 months.

Number of visitors v (1000s)	2450	2480	2540	2420	2350	2290	2400	2460
Amount of money spent m (£ millions)	1370	1350	1400	1330	1270	1210	1330	1350

You may use

$S_{vv} = 42587.5$ $S_{vm} = 31512.5$ $S_{mm} = 25187.5$ $\Sigma v = 19390$ $\Sigma m = 10610$

(*a*) Find the product moment correlation coefficient between m and v.

(2)

(*b*) Give a reason to support fitting a regression model of the form $m = a + bv$ to these data.

(1)

(*c*) Find the value of b correct to 3 decimal places.

(2)

(*d*) Find the equation of the regression line of m on v.

(2)

(*e*) Interpret your value of b.

(2)

(*f*) Use your answer to part (*d*) to estimate the amount of money spent when the number of visitors to the UK in a month is 2 500 000.

(2)

(*g*) Comment on the reliability of your estimate in part (*f*). Give a reason for your answer.

(2)

---

4. In a factory, three machines, *J*, *K* and *L*, are used to make biscuits.

Machine *J* makes 25% of the biscuits.

Machine *K* makes 45% of the biscuits.

The rest of the biscuits are made by machine *L*.

It is known that 2% of the biscuits made by machine *J* are broken, 3% of the biscuits made by machine *K* are broken and 5% of the biscuits made by machine *L* are broken.

(*a*) Draw a tree diagram to illustrate all the possible outcomes and associated probabilities.

(2)

A biscuit is selected at random.

(*b*) Calculate the probability that the biscuit is made by machine *J* and is not broken.

(2)

(*c*) Calculate the probability that the biscuit is broken.

(2)

(*d*) Given that the biscuit is broken, find the probability that it was not made by machine *K*.

(3)

---

5. The discrete random variable *X* has the probability function

$$P(X = x) = \begin{cases} kx & x = 2, 4, 6 \\ k(x-2) & x = 8 \\ 0 & \text{otherwise} \end{cases}$$

where *k* is a constant.

(*a*) Show that $k = \frac{1}{18}$.

(2)

(*b*) Find the exact value of F(5).

(1)

(*c*) Find the exact value of E(*X*).

(2)

(*d*) Find the exact value of E(X^2).

(2)

(*e*) Calculate Var(3 – 4*X*) giving your answer to 3 significant figures.

(3)

---

6. The times, in seconds, spent in a queue at a supermarket by 85 randomly selected customers, are summarised in the table below.

Time (seconds)	Number of customers, f
0 – 30	2
30 – 60	10
60 – 70	17
70 – 80	25
80 – 100	25
100 – 150	6

A histogram was drawn to represent these data. The 30 – 60 group was represented by a bar of width 1.5 cm and height 1 cm.

(*a*) Find the width and the height of the 70 – 80 group.

(3)

(*b*) Use linear interpolation to estimate the median of this distribution.

(2)

Given that x denotes the midpoint of each group in the table and

$$\Sigma fx = 6460 \qquad \Sigma fx^2 = 529\,400$$

(*c*) calculate an estimate for

(i) the mean,

(ii) the standard deviation,

for the above data.

(3)

One measure of skewness is given by

$$\text{coefficient of skewness} = \frac{3(\text{mean} - \text{median})}{\text{standard deviation}}$$

(*d*) Evaluate this coefficient and comment on the skewness of these data.

(3)

---

7. The heights of adult females are normally distributed with mean 160 cm and standard deviation 8 cm.

(*a*) Find the probability that a randomly selected adult female has a height greater than 170 cm.

(3)

Any adult female whose height is greater than 170 cm is defined as tall.

An adult female is chosen at random. Given that she is tall,

(*b*) find the probability that she has a height greater than 180 cm.

(4)

Half of tall adult females have a height greater than *h* cm.

(*c*) Find the value of *h*.

(5)

---

8. For the events *A* and *B*,

$$\mathrm{P}(A' \cap B) = 0.22 \quad \text{and} \quad \mathrm{P}(A' \cap B') = 0.18$$

(*a*) Find P(*A*).

(1)

(*b*) Find $\mathrm{P}(A \cup B)$.

(1)

Given that P(*A* | *B*) = 0.6,

(*c*) find $\mathrm{P}(A \cap B)$.

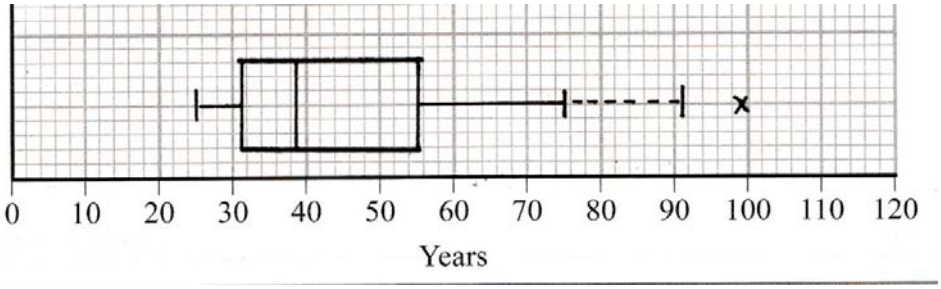
(3)

(*d*) Determine whether or not *A* and *B* are independent.

(2)

TOTAL FOR PAPER: 75 MARKS

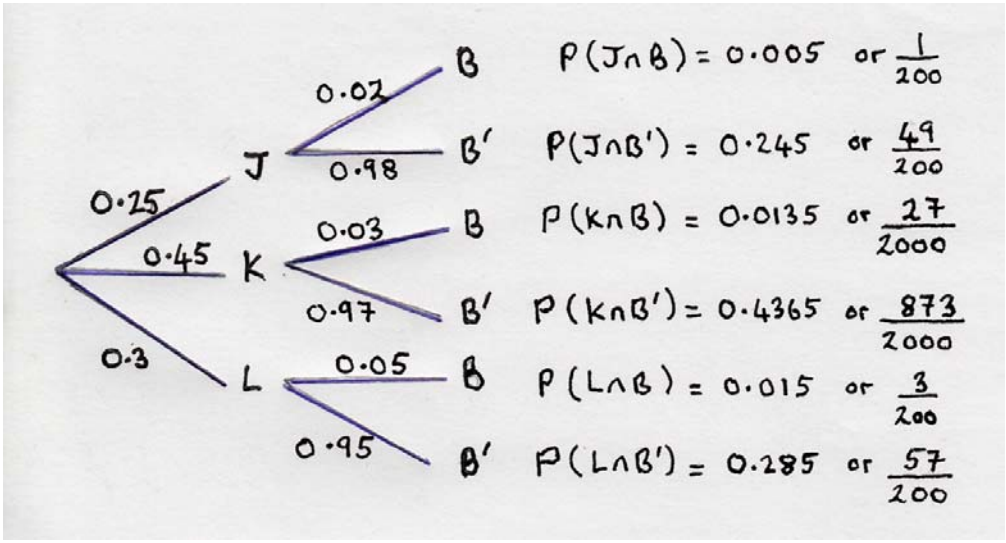
END

Question Number	Scheme		Marks
1. (a)	$a = 44$ $b = 76$	These answers may be in or near the table	B1 B1 **(2)**
(b)	$55+1.5(55-31)=91$ [and $31-1.5(55-31)=-5$] **Penville**  		M1 B1 B1 A1 **(4)**
(c)	Greenslax : $[Q_2-Q_1=20,\ Q_3-Q_2=12$ or $(Q_2-Q_1)>(Q_3-Q_2)]\Rightarrow$ −ve(skew) Penville: $[Q_2-Q_1=8,\ Q_3-Q_2=16$ or $(Q_3-Q_2)>(Q_2-Q_1)]\Rightarrow$ +ve (skew) Don't insist on seeing "skew" so just –ve and +ve will do. Treat "correlation" as ISW Justification that is consistent		B1 B1 ddB1 **(3)** **Total 9**
	Notes		
(b)	**A fully correct box plot scores 4/4. If <u>not</u> fully correct apply scheme and need evidence for M1** **If two box plots are seen ignore the one for Greenslax. If not on graph paper M1 max for (b)** M1 for sight of $55+1.5(55-31)$ <u>or</u> 91 seen (possibly implied by RH whisker of box plot) May be implied by a fully correct box plot 1st B1 box with whiskers (condone missing median) 2nd B1 25, 31, 39, 55, RH whisker to end at 75 or 91. Two RH whiskers is B0 Accuracy must be to within 0.5 of a square so e.g. lower quartile at 30 or 32 is OK		
(c)	A1 only one outlier plotted at 99. Allow cross to be vertically displaced If the RH whisker goes to 99 (2nd B0) <u>and</u> A0 even if outlier is identified since we require a horizontal "gap" between RH whisker and outlier. 1st B1 Greenslax − ve (skew) 2nd B1 Penville + ve (skew). We must be able to tell which is which but labels may be implied by their <u>values</u> but not simply from $Q_3-Q_2>Q_2-Q_1$ If there is just <u>one</u>, unlabelled comment assume Penville. 3rd ddB1 dependent on 1st and 2nd B marks being scored. Justification for <u>both</u> based on: quartiles, median relative to quartiles, or "tail" If only values for Q_3-Q_2 etc are <u>given</u> they should be correct ft for Greenslax and correct for Penville If values for Greenslax imply +ve skew then 1st B0 and 3rd B0		

Question Number	Scheme	Marks
2	mean $= \frac{60.8+20}{1.4}$ <u>or</u> $60.8 = 1.4x - 20$ (o.e.)	M1
	$= 57.7142\ldots$ awrt **57.7**	A1
	standard deviation $= \frac{6.60}{1.4}$ <u>or</u> $6.60 = 1.4x$	M1
	$= 4.7142\ldots$ awrt **4.71**	A1
		(4)
		Total 4
	Notes	
	1st M1 sub. 60.8 for *y* into a correct equation. Allow use of *x* or any other letter or expression for mean 1st A1 for awrt 57.7 or $\frac{404}{7}$ (o.e.). Correct answer only is 2/2 2nd M1 sub. 6.60 or 6.6 for *y* and ignoring the 20 Allow use of *x* or any other letter or expression for st. dev. $6.60^2 = 1.4^2 x^2$ is M0 until we see them take a square root. 2nd A1 for awrt 4.71 or $\frac{33}{7}$ (o.e.). Correct answer only is 2/2	

Question Number		Scheme	Marks
3	(a)	$r = \dfrac{31512.5}{\sqrt{42587.5 \times 25187.5}} = 0.962$ awrt **0.962**	M1 A1 **(2)**
	(b)	r is close to 1 <u>or</u> a **strong correlation**. ["points are close to a straight line" isB0] [Just "positive" correlation is B0] [Use of "relationship" or "skew" not "correlation" is B0]	B1 **(1)**
	(c)	$b = \dfrac{31512.5}{42587.5} = 0.739947... = 0.740$ (3 dp) **0.740** (only)	M1 A1cao **(2)**
	(d)	$a = 1326.25 - (0.7399\ldots \times 2423.75)$ [$= -467.2$ or awrt -467]	M1
		So $\boldsymbol{m = -467 + 0.74v}$	A1 **(2)**
	(e)	b is the <u>money (spent)</u> per <u>visitor</u>. (i.e. definition of a rate in words.)[ignore values]	B1
		So each 1000 visitors generates an extra £0.74 million <u>or</u> each visitor spends £740 oe	B1ft **(2)**
	(f)	$m = -467 + 0.74 \times 2500$	M1
		$m = 1383$ (£ million) awrt **1380**	A1 **(2)**
	(g)	As 2500 is within the range of the data set <u>or</u> it involves <u>interpolation</u>.	B1
		The value of money spent is reliable	dB1 **(2)**
			Total 13

	Notes
(a)	M1 for a correct expression for r. Ans only of 0.96 or awrt 0.96 is M1A0 Ans only of 0.962 or awrt 0.962 is M1A1. Do not allow fractions for A1
(b)	B1 for comment implying strong correlation. (e.g. big/high/clear etc) B0 if $\lvert r \rvert > 1$
(c)	M1 for a correct expression for b (may be implied by 0.74 or better in regression equation)
	A1 A1 for 0.740 only in (c) or $b = 0.740$ seen elsewhere (M1A0 for $\dfrac{2521}{3407}$ or awrt 0.74 here)
(d)	M1 for $1326.25 -$ ('their b' $\times$ 2423.75) Condone fractions or awrt 1330 for $\bar{m}$ and awrt 2420 for $\bar{v}$ A1 for a correct equation in m and v with a = awrt -467 and b = awrt 0.74 Condone $\frac{2521}{3407}$ for b and $\frac{-1591740}{3407}$ for a. [Equation in y and x is A0]
(e)	1st B1 for a correct definition of the rate in words. Must state or imply "money per visitor" Allow alternative words or symbols e.g. £ or "pounds" for money, "people" for visitors etc 2nd B1ft for a correct numerical rate (ft their value of b) e.g. "each <u>visitor</u> <u>spends</u> £740" is B1B1, "b is the extra <u>money</u> spent per <u>visitor</u>" is B1B0 [no values] "b is increase of <u>£0.74 million</u> in m as v increases <u>by 1000</u>" is B0B1[£ for money but no "visitors"] "increase in <u>m</u> as <u>v</u> increases" is B0B0 [Idea of rate but letters not words and no numerical value of rate]
(f)	M1 sub. $v = 2500$ into <u>their</u> equation. Simply substituting 2 500 000 is M0 (unless adjusted eqn) A1 awrt 1380 units (£ and million not required)
(g)	1st B1 for 2500 <u>or</u> 2 500 000 <u>or</u> visitors <u>or</u> v is in range. "it" is B0 unless v clearly implied 2nd dB1 for stating it <u>is</u> reliable. Dependent on previous B mark being awarded "both v and m in range" <u>or</u> "1380 in range" is B0 but use ISW so "interpolation since both in range" scores B1 for the "interpolation". "Not extrapolation" counts as "interpolation"

Question Number	Scheme	Marks
4 (a)		M1 A1 **(2)**
(b)	0.25×0.98, = **0.245** (or exact equiv. e.g. $\frac{49}{200}$)	M1A1 **(2)**
(c)	$0.25\times0.02+0.45\times0.03+0.3\times0.05$, = **0.0335** (or exact equiv. e.g. $\frac{67}{2000}$)	M1A1 **(2)**
(d)	$[\mathrm{P}(J\cup L\mid B)] = \frac{0.25\times0.02+0.3\times0.05}{0.0335}$ or $\frac{0.0335-0.45\times0.03}{0.0335}$ $=0.5970...$ awrt **0.597** (or $\frac{40}{67}$ or exact equiv.)	M1A1ft A1 **(3)**
	Notes	**Total 9**
	Allow fractions or percentages throughout this question	
(a)	Allow 3+6 tree diagram with the 6 correct "end" probs and labels to get 2/2 (1st, 3rd, 5th gets M1) M1 for (3+6) tree drawn with 0.25, 0.45, 0.02, 0.03, 0.05 on correct branches A1 for 0.3, 0.98, 0.97, 0.95 on the correct branches and labels, condone missing B's **Correct answer only scores full marks for parts (b), (c) and (d)** **When using "their probability p" for M1 and A1ft they must have $0 < p < 1$**	
(b)	M1 for $0.25\times$ 'their 0.98' o.e.	
(c)	M1 for $0.25\times$their $0.02+0.45\times$their $0.03+$their $0.3\times$their 0.05 Condone 1 transcription error. Or $1-(0.25\times\text{their } 0.98+0.45\times\text{their } 0.97+\text{their } 0.3\times\text{their } 0.95)$	
(d)	M1 for use of conditional probability with their (c) as denominator. Also exactly 2 products on num' and at least one correct (or correct ft) or their (c) – one of the products from their (c). Ignore an incorrect expression inside their probability statement A1ft for $\frac{0.25\times\text{their } 0.02+\text{their } 0.3\times\text{their } 0.05}{\text{their(c)}}$ or $\frac{\text{their (c)}-0.45\times\text{their } 0.03}{\text{their (c)}}$ or $\frac{0.02}{\text{their (c)}}$ A1 awrt 0.597 or exact fraction e.g. $\frac{40}{67}$	

Question Number		Scheme	Marks
5	(a)	$2k + 4k + 6k + k(8 - 2) = 1$ (commas instead of + or a table OK if $18k = 1$ seen later)	M1
		$k = \frac{1}{18}$ **(*)**	A1cso **(2)**
	(b)	$[2k + 4k] = \frac{6}{18} = \frac{1}{3}$ ($\frac{1}{3}$ or any exact <u>numerical</u> equivalent)	B1 **(1)**
	(c)	$E(X) = \left(2 \times \frac{1}{9}\right) + \left(4 \times \frac{2}{9}\right) + \left(6 \times \frac{1}{3}\right) + \left(8 \times \frac{1}{3}\right)$ <u>or</u> $(2 \times 2k) + (4 \times 4k) + (6 \times 6k) + (8 \times 6k)$	M1
		$= 5\frac{7}{9}$ (or exact equivalent e.g. $\frac{52}{9}$)	A1 **(2)**
	(d)	$E(X^2) = \left(4 \times \frac{1}{9}\right) + \left(16 \times \frac{2}{9}\right) + \left(36 \times \frac{1}{3}\right) + \left(64 \times \frac{1}{3}\right)$ <u>or</u> $(4 \times 2k) + (16 \times 4k) + (36 \times 6k) + (64 \times 6k)$	M1
		$= 37\frac{1}{3}$ (or exact equivalent e.g. $\frac{112}{3}$)	A1 **(2)**
	(e)	$\text{Var}(X) = 37\frac{1}{3} - \left(5\frac{7}{9}\right)^2$ [= 3.95… or $\frac{320}{81}$]	M1
		$\text{Var}(3 - 4X) = 16 \times 3.95\ldots$	M1
		= awrt **63.2** (allow $\frac{5120}{81}$)	A1 **(3)**
			Total 10

	Notes
(a)	M1 for $2k + 4k + 6k + k(8 - 2) = 1$ A1 for $k = \frac{1}{18}$ NB cso so no incorrect working seen <u>or</u> M1 for $2 \times \frac{1}{18} + 4 \times \frac{1}{18} + 6 \times \frac{1}{18} + \frac{1}{18}(8 - 2)$ A1 for =1 and "therefore $k = \frac{1}{18}$"
	If in parts (c), (d) and (e) there is a correct expression worthy of M1 but later they incorrectly go on and multiply or divide by some number *n*, then allow the M1 but mark their <u>final</u> answer (A0) Answers only in (b), (c), (d) and (e) score all the marks.
(c)	M1 for an expression for E(X) with at least 3 correct terms (products) allow use of k e.g. $104k$
(d)	M1 for an expression for E(X^2) with at least 3 correct terms (products) allow use of k e.g. $672k$ A1 for any exact equivalent only. E.g. 37.3 is A0 but, of course, $37.\dot{3}$ is OK
(e)	1st M1 for $E(X^2) - [E(X)]^2$ ft their answers to (c) and (d). Must see values <u>used</u> correctly. 2nd M1 for statement "$4^2 \times \text{Var}(X)$"seen <u>or</u> for $4^2 \times$ their Var(X) provided their Var(X) > 0 Do not allow for $16 \times E(X^2)$ but can score M0M1 NB condone $-4^2 \times \text{Var}(X)$ if the answer later becomes positive. A1 for exact fraction ($\frac{5120}{81}$ o.e.) or decimal approximation that is awrt 63.2 **Beware:** rounding to 3sf in (c) (5.78) and (d) (37.3) gives 62.3 which could be misread as 63.2

Question Number		Scheme	Marks
6	(a)	70 – 80 group - width **0.5 (**cm)	B1
		1.5 cm^2 is 10 customers <u>or</u> 3.75cm^2 is 25 customers <u>or</u> $0.5c = 3.75$ <u>or</u> $\frac{2.5}{\frac{1}{3}}$	M1
		70 – 80 group - height **7.5 (**cm)	A1
			(3)
	(b)	Median $= (70) + \frac{13.5}{25} \times 10$ allow $(n+1) = (70) + \frac{14}{25} \times 10$	M1
		$=$ **75.4** (or if using $(n+1)$ allow 75.6)	A1
			(2)
	(c)	$\left[\text{Mean} = \frac{6460}{85}\right] =$ **76**	B1
		$\sigma = \sqrt{\frac{529400}{85} - 76^2}$	M1
		$= 21.2658\ldots\ldots$ $(s = 21.3920)$ **awrt 21.3**	A1
			(3)
	(d)	Coeff' of skewness $= \frac{3(76 - 75.4)}{21.2658...} = 0.08464\ldots$ **awrt 0.08** (awrt 0.06 for 75.6)	M1 A1
		There is (very slight) positive skew or the data is almost symmetrical (or both) <u>Any</u> mention of "correlation" is B0	B1ft
			(3)
			Total 11

	Notes
(a)	B1 for 0.5
	M1 for one of the given statements <u>or</u> any method where "their width" × "their height" = 3.75 Correct height scores M1A1 independent of width so B0M1A1 is possible.
(b)	M1 for a correct fraction: $+\frac{k}{25} \times 10$ where $k = 13.5$ or 14 for $(n+1)$ case.
	NB may work down so look out for $(80) - \frac{11.5}{25} \times 10$ etc **Beware:** $69.5 + \frac{13.5}{25} \times 11 = 75.44$ (but M0)
(c)	M1 for a correct expression with square root, ft their mean A1 for awrt 21.3 or, if clearly using *s* allow awrt 21.4. Must be evaluated...no surds.
(d)	M1 sub. their values into formula allow use of *s* but their σ or *s* must be > 0 A1 for awrt 0.08 but accept 0.085 No fraction B1ft for a correct comment compatible with their coefficient. Allow "symmetrical" for |coeff'| < 0.25 They may say it is "slightly skew" so omit "positive" but do not allow "negative" if coef' +ve Condone "strongly" positive skew.

Question Number		Scheme	Marks
7	(a)	The random variable H ~ height of females	
		$\mathrm{P}(H>170) = \mathrm{P}\left(Z>\frac{170-160}{8}\right)$ $[=\mathrm{P}(Z>1.25)]$	M1
		$=1-0.8944$	M1
		$=0.1056$ (calc 0.1056498…) **awrt 0.106** (accept 10.6%)	A1 **(3)**
	(b)	$\mathrm{P}(H>180) = \mathrm{P}\left(Z>\frac{180-160}{8}\right)$ $[=1-0.9938]$	M1
		$=0.0062$ (calc 0.006209…) awrt 0.0062 or $\frac{31}{5000}$	A1
		$[\mathrm{P}(H>180 \mid H>170)] = \frac{0.0062}{0.1056}$	M1
		$=0.0587$ (calc 0.0587760…) **awrt 0.0587 or 0.0588**	A1 **(4)**
	(c)	$\mathrm{P}(H>h \mid H>170)$ $(=0.5)$ or $\frac{\mathrm{P}(H>h)}{\mathrm{P}(H>170)}(=0.5)$	M1
		$[\mathrm{P}(H>h)] = 0.5\times\text{"0.1056"} = 0.0528$ (calc 0.0528249…) or $[\mathrm{P}(H<h)] = 0.9472$	A1ft
		$\frac{h-160}{8}=1.62$ (calc 1.6180592…)	M1 B1
		h = awrt 173 cm **awrt 173**	A1 **(5)**
			Total 12

	Notes
(a)	1st M1 for attempt at standardising with 170, 160 and 8. Allow ± i.e. for $\pm\frac{170-160}{8}$
	2nd M1 for attempting $1-p$ where $0.8<p<1$. Correct answer only 3/3
(b)	1st M1 for standardising with 180, 160 and 8
	1st A1 for 0.0062 seen, maybe seen as part of another expression/calculation.
	2nd M1 using conditional probability with denom = their (a) and num < their denom. Values needed.
	2nd A1 for awrt 0.0587 or 0.0588. Condone 5.87% or 5.88% or $\frac{31}{528}$
	Correct answer only 4/4
(c)	1st M1 for a correct conditional probability statement. Either line and don't insist on 0.5, ft (a)
	1st A1ft for $[\mathrm{P}(H>h)] = 0.5\times\text{their}(a)$
	Award M1A1ft for correct evaluation of $0.5\times\text{their}(a)$ or sight of 0.0528 or better
	2nd M1 for attempt to standardise (±) with 160 and 8 and set equal to ± z value $(1.56<\lvert z\rvert<1.68)$
	B1 for (z =) awrt ± 1.62 (seen)
	2nd A1 for awrt 173 but dependent on both M marks.

Question Number	Scheme	Marks
8 (a)	$[\,P(A) = 1 - 0.18 - 0.22\,]$ = **0.6** (or exact equivalent)	B1 **(1)**
(b)	$P(A \cup B) = "0.6" + 0.22$ = **0.82** (or exact equivalent)	B1ft **(1)**
(c)	$x = P(A \cap B)$ | Use $P(B)P(A' \mid B) = P(A' \cap B)$ | **Establish** independence before or after 1st M1and score marks for (d) (RH ver)	M1
	$\frac{x}{x+0.22} = 0.6$ | $P(B) \times [1 - 0.6] = 0.22$ | **Find** $P(B)$	
	$x = 0.6x + 0.132$ | Use $P(A \cap B) = P(A \mid B)P(B)$ | Use $P(B)P(A) = P(A \cap B)$	dM1
	$0.4x = 0.132$ | $P(A \cap B) = 0.6 \times 0.55$ | $P(A \cap B) = 0.6 \times 0.55$	
	$x =$ **0.33** (or exact equivalent)	A1cso **(3)**
(d)	$P(B) = 0.55$	
	$P(B) \times P(A) = 0.55 \times 0.6$ $= 0.33$	**or** stating $P(A) = P(A\|B)$ [= 0.6]	M1
	$P(B) \times P(A) = P(A \cap B)$ therefore (statistically) independent	**or** $P(A) = P(A\|B)$ therefore (statistically) independent	A1cso **(2)**
		Total 7

	Notes
(b)	B1ft for their (a) + 0.22 or $1 - P(A' \cap B')$ Do not ft their (a) if it is > 0.78
	NB 3 versions for (c). Check carefully that Ms are genuinely scored.
	Look out for <u>assuming independence</u> and if you see P(*B*) = 0.55 check it is <u>derived</u> properly
(c)	1st M1 for a correct equation for x e.g. $\frac{x}{x+0.22} = 0.6$ <u>or</u> a correctly derived equation for $P(B)$
	2nd dM1 for solving to get in form $kx = L$ <u>or</u> <u>correct</u> use of $P(B)$ to find $P(A \cap B)$ [2nd or 3rd ver] <u>or</u> $P(A \cap B) = P(B) - 0.22$
	A1cso for 0.33 Dep. on <u>both</u> Ms and no incorrect working seen.
(d)	M1 for finding $P(B) \times P(A) = 0.33$ (values needed) <u>or</u> stating $P(A) = P(A\|B)$ (= 0.6 not needed)
	A1cso for a correct statement: $P(B) \times P(A) = P(A \cap B)$ or $P(A) = P(A\|B)$ <u>and</u> stating independent
	NB The M1 in (d) using $P(A \cap B)$ requires $P(B) = 0.55$ There is no ft of an incorrect $P(B)$ Full marks in (d) is OK even if 0/3 in (c)
	{This Venn diagram may be helpful.}