Paper Reference(s)

6691/01

Edexcel GCE

Statistics S3

Advanced/Advanced Subsidiary

Thursday 13 June 2013 – Morning

Time: 1 hour 30 minutes

Materials required for examination	Items included with question papers
Mathematical Formulae (Pink)	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 or symbolic differentiation/integration, or have retrievable mathematical formulae stored in them.

Instructions to Candidates

In the boxes above, write your centre number, candidate number, your surname, initials and signature.
Check that you have the correct question paper.
Answer ALL the questions.
You must write your answer for each question in the space following the question.
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.
The marks for the parts of questions are shown in round brackets, e.g. **(2)**.
There are 7 questions in this question paper. The total mark for this paper is 75.
There are 20 pages in this question paper. Any blank pages are indicated.

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.

P41817A

1. A doctor takes a random sample of 100 patients and measures their intake of saturated fats in their food and the level of cholesterol in their blood. The results are summarised in the table below.

Cholesterol level / Intake of saturated fats	High	Low
High	12	8
Low	26	54

Using a 5% level of significance, test whether or not there is an association between cholesterol level and intake of saturated fats. State your hypotheses and show your working clearly.

(10)

---

2. The table below shows the number of students per member of staff and the student satisfaction scores for 7 universities.

University	*A*	*B*	*C*	*D*	*E*	*F*	*G*
Number of students per member of staff	14.2	13.1	13.3	11.7	10.5	15.9	10.8
Student satisfaction score	4.1	4.2	3.8	4.0	3.9	4.3	3.7

(*a*) Calculate Spearman's rank correlation coefficient for these data.

(5)

(*b*) Stating your hypotheses clearly test, at the 5% level of significance, whether or not there is evidence of a correlation between the number of students per member of staff and the student satisfaction score.

(3)

---

3. A college manager wants to survey students' opinions of enrichment activities.
She decides to survey the students on the courses summarised in the table below.

Course	Number of students enrolled
Leisure and Sport	420
Information Technology	337
Health and Social Care	200
Media Studies	43

Each student takes only one course.

The manager has access to the college's information system that holds full details of each of the enrolled students including name, address, telephone number and their course of study. She wants to compare the opinions of students on each course and has a generous budget to pay for the cost of the survey.

(*a*) Give one advantage and one disadvantage of carrying out this survey using

(i) quota sampling,

(ii) stratified sampling.

(2)

The manager decides to take a stratified sample of 100 students.

(*b*) Calculate the number of students to be sampled from each course.

(3)

(*c*) Describe how to choose students for the stratified sample.

(2)

4. Customers at a post office are timed to see how long they wait until being served at the counter. A random sample of 50 customers is chosen and their waiting times, x minutes, are summarised in Table 1.

Waiting time in minutes (x)	Frequency
0–3	8
3–5	12
5–6	13
6–8	9
8–12	8

Table 1

(a) Show that an estimate of $\bar{x} = 5.49$ and an estimate of $s_x^2 = 6.88$.

(3)

The post office manager believes that the customers' waiting times can be modelled by a normal distribution.

Assuming the data is normally distributed, she calculates the expected frequencies for these data and some of these frequencies are shown in Table 2.

Waiting Time	$x < 3$	3–5	5–6	6–8	$x > 8$
Expected Frequency	8.56	12.73	7.56	a	b

Table 2

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

(3)

(c) Test, at the 5% level of significance, the manager's belief. State your hypotheses clearly.

(8)

5. Blumen is a perfume sold in bottles. The amount of perfume in each bottle is normally distributed. The amount of perfume in a large bottle has mean 50ml and standard deviation 5ml. The amount of perfume in a small bottle has mean 15ml and standard deviation 3ml.

One large and 3 small bottles of Blumen are chosen at random.

(*a*) Find the probability that the amount in the large bottle is less than the total amount in the 3 small bottles.

(6)

A large bottle and a small bottle of Blumen are chosen at random.

(*b*) Find the probability that the large bottle contains more than 3 times the amount in the small bottle.

(6)

---

6. Fruit-n-Veg4U Market Gardens grow tomatoes. They want to improve their yield of tomatoes by at least 1 kg per plant by buying a new variety. The variance of the yield of the old variety of plant is 0.5 kg^2 and the variance of the yield for the new variety of plant is 0.75 kg^2. A random sample of 60 plants of the old variety has a mean yield of 5.5 kg. A random sample of 70 of the new variety has a mean yield of 7 kg.

(*a*) Stating your hypotheses clearly test, at the 5% level of significance, whether or not there is evidence that the mean yield of the new variety is more than 1 kg greater than the mean yield of the old variety.

(9)

(*b*) Explain the relevance of the Central Limit Theorem to the test in part (*a*).

(2)

---

7. Lambs are born in a shed on Mill Farm. The birth weights, x kg, of a random sample of 8 newborn lambs are given below.

4.12 5.12 4.84 4.65 3.55 3.65 3.96 3.40

(*a*) Calculate unbiased estimates of the mean and variance of the birth weight of lambs born on Mill Farm.

(3)

A further random sample of 32 lambs is chosen and the unbiased estimates of the mean and variance of the birth weight of lambs from this sample are 4.55 and 0.25 respectively.

(*b*) Treating the combined sample of 40 lambs as a single sample, estimate the standard error of the mean.

(7)

The owner of Mill Farm researches the breed of lamb and discovers that the population of birth weights is normally distributed with standard deviation 0.67 kg.

(*c*) Calculate a 95% confidence interval for the mean birth weight of this breed of lamb using your combined sample mean.

(5)

TOTAL FOR PAPER: 75 MARKS

END

Question Number	Scheme	Marks
1.	<table><tr><td>Cholesterol Level</td><td>High</td><td>Low</td><td></td></tr><tr><td>High</td><td>7.6</td><td>12.4</td><td>20</td></tr><tr><td>Low</td><td>30.4</td><td>49.6</td><td>80</td></tr><tr><td></td><td>38</td><td>62</td><td>100</td></tr></table>	M1A1 (2)
	H_0 : Cholesterol level is independent of intake of saturated fats(no association)	B1
	H_1 : Cholesterol level is not independent of intake of saturated fats (association)	(1)
	<table><tr><td>O</td><td>E</td><td>$\frac{(O-E)^2}{E}$</td><td>$\frac{O^2}{E}$</td></tr><tr><td>12</td><td>7.6</td><td>2.547… or $\frac{242}{95}$</td><td>18.947… or $\frac{360}{19}$</td></tr><tr><td>8</td><td>12.4</td><td>1.56129… or $\frac{242}{155}$</td><td>5.161… or $\frac{160}{31}$</td></tr><tr><td>26</td><td>30.4</td><td>0.6368… or $\frac{121}{190}$</td><td>22.236… or $\frac{845}{38}$</td></tr><tr><td>54</td><td>49.6</td><td>0.3903… or $\frac{121}{310}$</td><td>58.790… or $\frac{3645}{62}$</td></tr></table>	dM1 A1
	$\sum\frac{(O-E)^2}{E}$=5.1358234.. or $\frac{1.2^2}{7.6}+\frac{8^2}{12.4}+\frac{26^2}{30.4}+\frac{54^2}{49.6}-100=5.14$ (awrt **5.14**)	A1 (3)
	$\nu=(2-1)(2-1)=1$	B1
	$\chi^2_1(0.05)=3.841$	B1 (2)
	5.14 > 3.841 so sufficient evidence to reject H_0 [Condone "accept H_1"]	M1
	Association between cholesterol level and saturated fat intake	A1 (2)
		Total 10
	Notes	
	Minimum working use part marks: E_i (2), Hyp (1), 5.14 (3), 3.841 (2), Conclusion (2) 1st M1 for some use of $\frac{\text{Row Total}\times\text{Col.Total}}{\text{Grand Total}}$. May be implied by correct E_i 1st A1 for all expected frequencies correct. Allow M1A0 for E_i rounded to integers 1st B1 for both hypotheses. Must mention "cholesterol" and "fats" at least once Use of "relationship" or "correlation" or "connection" is B0 2nd dM1 for at least 2 correct terms (as in 3rd or 4th column) or correct expressions with their E_i Dependent on 1st M1 Accept 2sf accuracy for the M mark 2nd A1 for all correct terms. May be implied by a correct ans.(2 dp or better) Allow truncation eg 2.54… 3rd A1 for awrt 5.14 2nd B1 for correct degrees of freedom (may be implied by a cv of 3.841) 3rd M1 for a correct statement linking their test statistic and their cv(cv could be 2.705 or > 3.5) Contradictory statements score M0 e.g. "significant, do not reject H_0" 4th A1 for a correct comment in context - must mention "cholestorol" and "fats" condone "relationship" or "connection" here but **not** "correlation". e.g. "There is evidence of a relationship between cholesterol level and fat intake" No follow through. If e.g hypotheses are the wrong way round A0 here.	

Question Number	Scheme	Marks
2(a)	<table><tr><td>Uni</td><td>*A*</td><td>*B*</td><td>*C*</td><td>*D*</td><td>*E*</td><td>*F*</td><td>*G*</td><td></td></tr><tr><td>Staff-Stu</td><td>2</td><td>4</td><td>3</td><td>5</td><td>7</td><td>1</td><td>6</td><td></td></tr><tr><td>Satisfaction</td><td>3</td><td>2</td><td>6</td><td>4</td><td>5</td><td>1</td><td>7</td><td></td></tr><tr><td>[*d*]</td><td>–1</td><td>2</td><td>–3</td><td>1</td><td>2</td><td>0</td><td>–1</td><td></td></tr><tr><td>d^2</td><td>1</td><td>4</td><td>9</td><td>1</td><td>4</td><td>0</td><td>1</td><td>20</td></tr></table>	M1A1A1
	$r_s = 1 - \frac{6\times 20}{7(49-1)} = 0.642857....$ (accept $\frac{9}{14}$) (awrt **0.643)**	dM1A1
		(5)
(b)	H_0: $\rho = 0$	
	$H_{1:}$ $\rho \neq 0$ ($\rho > 0$)	B1
	Critical value is ± 0.7857(± 0.7143 for a one tailed test)	B1
	0.643<cv so insufficient evidence to reject H_0	
	There is insufficient evidence to suggest a (positive) correlation between staff-student ratio and satisfaction.	B1ft
		(3)
		Total 8

	Notes
(a)	1st M1 for an attempt to rank the staff-students ratio or satisfaction (at least 4 correct)
	1st A1 for correct rankings for both (one or both may be reversed)
	2nd A1 for $\sum d^2$ =20 or correct d^2 row (NB $\sum d^2 = 92$ for one set of reversed ranks)
	2nd dM1 for use of the correct formula, follow through their $\sum d^2$ (Dependent on 1st M1)
	If answer is not correct, a correct expression is required.
	3rd A1 If $\sum d^2$ = 20 for awrt 0.643 or if $\sum d^2$ =92 for awrt – 0.643 (accept $\pm\frac{9}{14}$)
(b)	1st B1 for both hypotheses in terms of ρ, one tail H_1 must be compatible with their ranking
	Hypotheses just in words e.g. "no correlation" score B0
	2nd B1 for cv of 0.7857 or 0.7143 for one-tailed test (accept ±)
	Their cv must be compatible with their H_1 which may be in words
	If hypotheses are the wrong way around this must be B0 but 3rd B1 is possible.
	3rd B1ft for a correct contextualised comment. Must mention "ratio" or "no. of students per member of staff" and "satisfaction"
	Follow through their r_s and their cv (provided it is |cv| <1)
	Don't insist on the word "positive" for a one-tailed test
	Use of "association" is B0
	Independent of 1st B1 so if $\|r_s\|$ > |cv| must say there is sufficient evidence of …….(o.e.)
	and if $\|r_s\|$ < |cv| must say insufficient evidence of … (o.e.) regardless of their hypotheses
	Contradictory statements score B0
	(This mark is just testing interpretation of comparison of their r_s and their cv)

Question Number	Scheme	Marks
3(a)i e.g e.g	**Quota Sampling:** Advantages: Fieldwork can be done **quick**ly, or administering the test is **easy**, or costs are kept to a minimum (**cheap**), or gives estimates for each course. or OK for large populations or sampling frame not required (o.e.) Disadvantages: **Non-random** process or not possible to estimate the sampling errors, or non response not recorded, or interviewer can introduce **bias** in sample choice. (o.e.)	B1
3(a)ii e.g. e.g.	**Stratified Sampling:** Advantages: Can give accurate estimates as it is a **random** process, or gives estimates for each course or **representative** of [BUT not "proportional" to] the whole population. (o.e.) Disadvantages: Sampling frame required, or strata may not be clear as some students overlap courses or not suitable for large populations. (o.e.)	B1 (2)
3(b)	Total enrolments=1000 (may be implied by calculations) Leisure and Sport= $\frac{420}{1000}\times 100 = 42$ Information Technology= $\frac{337}{1000}\times 100 = 33.7 = 34$ Health and Social Care= $\frac{200}{1000}\times 100 = 20$ Media Studies= $\frac{43}{1000}\times 100 = 4.3 = 4$	B1 M1 A1 (3)
3(c)	The college's information system would be used to identify each student and which course they are enrolled on. i.e. idea of **sampling frame** or **list** for **each course**. Use of **random numbers** to select required number of students **from each course**	B1 B1 (2) **Total 7**
	Notes	
(a)	Do not penalise for lack of context in part (a) 1st B1 for an advantage and a disadvantage for quota sampling (must be 1st or labelled (i)) 2nd B1 for an advantage and a disadvantage for stratified sampling (2nd or labelled (ii)) Do not allow opposite pairs e.g. "quicker/easier" for quota sampling and "takes a long time/more difficult" for stratified or quota "easy to use" but strat. "hard for large populations" Do not allow same reason for both e.g. "gives estimates for each course"	
(b)	M1 for one correct calculation, ft their "1000" A1 for 42, 34, 20 and 4 only	
(c)	1st B1 for some mention of a suitable sampling frame. Need not give the specific term but a suitable source of list is required for all students in each course. 2nd B1 for mentioning use of random numbers or some random selection process for each course. If they are describing systematic sampling score B0 here	

Question Number	Scheme	Marks
4 (a)	$\bar{x} = \dfrac{8\times1.5+12\times4+13\times5.5+9\times7+8\times10}{50} = \dfrac{274.5}{50} = 5.49$ **(*)**	B1cso
	$s^2 = \dfrac{8\times1.5^2+12\times4^2+13\times5.5^2+9\times7^2+8\times10^2}{49} - \dfrac{50}{49}5.49^2, = 6.88$ **(*)**	M1,A1cso
		(3)
(b)	$a = \quad 50\times \mathrm{P}(6 < X < 8) = 50\times \mathrm{P}(0.194.. < Z < 0.956..)$	M1
	a = 12.81 (tables) <u>or</u> 12.68 (calc)	A1
	$b = 50 - (28.85 + a)$ = 8.34 (tables) <u>or</u> 8.47 (calc)	A1ft
		(3)
(c)	H_0: Normal distribution is a good fit H_1: Normal distribution is not a good fit	B1
		M1
		A1
	$\sum \dfrac{O^2}{E} - N = 5.087\ldots \sim 5.1400\ldots$ awrt **(5.09 ~ 5.14)**	A1
	$\nu = 5-3 = 2$ (for 5 – 3 or 2 can be implied by 5.991 seen)	B1
	$\chi^2_2(0.05) = 5.991$	B1
	5.09<5.991 so insufficient evidence to reject H_0	M1
	Normal distribution is a good fit.	A1 (8)
		Total 14

Class	O	E	$\dfrac{O^2}{E}$	$\dfrac{(O-E)^2}{E}$
0-3	8	8.56	7.4766…	0.0366…
3-5	12	12.73	11.31186….	0.0418…
5-6	13	7.56	22.354497…	3.9144…
6-8	9	12.68 or (12.81)	(6.32) ~ 6.38801..	1.0680...~ (1.13)
8-12	8	(8.34) or 8.47	7.556080…~ (7.67)	(0.013) ~ 0.0260..

Notes

(a)

B1cso for denominator of 50 and at least 3 products on num <u>or</u> 274.5 on num

M1 for a correct expression with at least 3 correct products on num or $\dfrac{1844.25}{49} - \dfrac{1507.005}{49}$

<u>or</u> $\dfrac{337.245}{49}$ <u>or</u> $\left(\dfrac{7377}{200} - 5.49^2\right)\times\dfrac{50}{49}$ etc Allow 3sf accuracy

A1cso for 6.88 with M1 scored and no incorrect working seen

(b)

M1 a full method for a or b using the normal dist. Correct use of (6), 8, 5.49 and $\sqrt{6.88}$ seen

1st A1 for a in range 12.68 ~ 12.81 or b in range 8.34~ 8.47 or awrt these values

2nd A1ft for 50 – 28.85 – their a (or b) (but requires M1). Allow awrt 3sf. Must add up to 50

(c)

1st B1 for both hypotheses. B0 if they include 5.49 or 6.88. Condone X~N(μ, σ^2) etc

1st M1 for attempting $\dfrac{(O-E)^2}{E}$ or $\dfrac{O^2}{E}$, at least 3 correct expressions or values.

1st A1 for at least 4 correct calcs - 3rd or 4th column. (2 dp or better and allow e.g. 7.47) Allow any value in the ranges for the last two rows.

2nd A1 for a test statistic that is awrt 5.09 ~ 5.14. Award M1A1A1 if this is obtained.

2nd M1 for a correct statement based on their test statistic (> 1) and their cv (> 3.8) Contradictory statements score M0 e.g. "significant" do not reject H_0.

3rd A1 for a correct comment suggesting that normal model is suitable <u>or</u> manager's belief is correct. **No f t** . Condone mention of 5.49 or 6.88 here. Hypotheses wrong way round scores A0

Question Number	Scheme	Marks
5 (a)	Let $L \sim \mathrm{N}(50, 25)$ and $S \sim \mathrm{N}(15, 9)$	
	Let $X = L - (S_1 + S_2 + S_3)$	B1
	$\mathrm{E}(X) = 50 - 3\times 15 = 5$	B1
	$\mathrm{Var}(X) = 25 + 3\times 9 = 52$	M1A1
	$\mathrm{P}(X < 0) = \mathrm{P}\left(Z < \frac{-5}{\sqrt{52}}\right)$	dM1
	$= \mathrm{P}(Z < -0.693..)$	
	=0.244 or 0.2451 (tables) (awrt **0.244 ~ 0.245)**	A1
		(6)
(b)	Let $Y = L - 3S$	B1
	$\mathrm{E}(Y) = 50 - 3\times 15 = 5$	B1
	$\mathrm{Var}(Y) = 25 + 3^2 \times 9 = 106$	M1A1
	$\mathrm{P}(Y > 0) = \mathrm{P}\left(Z > \frac{-5}{\sqrt{106}}\right)$	dM1
	$= \mathrm{P}(Z > -0.4856..)$	
	=0.686 or 0.6879 (tables) (awrt **0.686 ~ 0.688)**	A1
		(6)
		Total 12
	Notes	
(a)	1st B1 for forming a suitable variable X explicitly seen. Do not give for $L - 3S$ but allow $L - (S + S + S)$	
	2nd B1 for E(X) = 5 (or – 5 if their X is defined the other way around)	
	1st M1 for an attempt at Var(X) = Var(L) + 3Var(S). Do not condone 5 for "25" or 3 for "9"	
	1st A1 for 52	
	2nd dM1 for attempting the correct probability and standardising with their mean and sd. This mark is dependent on 1st M1 so if X is not being used or wrong variance score M0 If their method is not crystal clear then they must be attempting P(Z < -ve value) or P(Z > +ve value) i.e. their probability after standardisation should lead to a prob. < 0.5	
	2nd A1 for awrt 0.244 ~ 0.245	
	Correct ans. only scores 5/6 (or 6/6 if 1st B1) but must be clearly labelled as (a) or the first answer.	
(b)	1st B1 for defining a new variable [Y =]± (L – 3S). May be implied by a correct variance.	
	2nd B1 for E(Y) = 5 (or – 5 if their Y is defined as Y = 3S – L)	
	1st M1 for an attempt at Var(Y) = Var(L) + 3^2 Var(S). Do not condone 5 for "25" or 3 for "9"	
	1st A1 for 106 only	
	2nd dM1 for attempting the correct probability and standardising with their mean and sd. This mark is dependent on 1st M1 so if Y is not being used or wrong variance score M0 If their method is not crystal clear then they must be attempting P(Z > -ve value) or P(Z < +ve value) i.e. their probability after standardisation should lead to a prob. > 0.5	
	2nd A1 for an awrt 0.686 ~ 0.688	
	Correct answer only scores 6/6 but must be clearly labelled as (b) or the second answer.	

Question Number	Scheme	Marks
6 (a)	$H_0 : \mu_{new} - \mu_{old} = 1$	B1
	$H_1 : \mu_{new} - \mu_{old} > 1$	B1
	$z = \dfrac{7 - 5.5 - 1}{\sqrt{\dfrac{0.5}{60} + \dfrac{0.75}{70}}} = 3.62254...$ (awrt **3.62**)	M1 A1A1 A1
	Critical value $z = 1.6449$ (allow $\pm$)	B1
	[3.62 >1.6449] so sufficient evidence to reject H_0	dM1
	Evidence that the mean yield of new variety is more than 1 kg greater than the old variety.	A1
		(9)
(b)	Mean yield is normally distributed	B1
	Sample size is large. Must state or imply that **in this case** sample size is large	B1
		(2)
		Total 11
	Notes	
(a)	1st & 2nd B1 for hypotheses. Accept μ_1, μ_2 or μ_A, μ_B etc if there is some indication of which is which e.g. $A \sim N(\mu_A, 0.5)$	
	1st M1 for an attempt at se. Condone switching 0.5 and 0.75 $\sqrt{\dfrac{0.5 \text{ or } 0.75}{60} + \dfrac{0.75 \text{ or } 0.5}{70}}$	
	1st A1 for a correct expression for denominator of test statistic or 0.138… or $\sqrt{0.0190...}$	
	2nd A1 for a correct numerator of test statistic (must have the -1)	
	3rd A1 for awrt 3.62	
	[Allow – 3.62 from numerator of 5.5 – 7 − − 1 and compatible H_1]	
	3rd B1 for $\pm$ 1.6449 seen or	
	probability of 0.0002 (tables) or 0.000145…(calc) [allow 0.0001]	
	2nd dM1 dep. on 1st M1 for a correct statement based on their normal cv and their test statistic	
	2nd A1 for correct comment in context. Must mention "yield" and "varieties" or "old" and "new" and "1"	
	If second B mark is B0 award A0 here	
ALT	**Pooled estimate:** If they calculate $s_p = \sqrt{0.41845...} = 0.64688...$ allow 1st M1, 1st A1 for expression (or awrt 0.114) and 2nd A1 if numerator correct but A0 for test statistic (4.39)	
(b)	1st B1 for mention of mean (yield) and normal (distribution)	
	2nd B1 for mention of sample (size) being large in this case	

Question Number		Scheme	Marks
7	**(a)**	$\hat{\mu} = \bar{x} = \frac{33.29}{8} = 4.16125$ (awrt **4.16**)	B1
		$\hat{\sigma}^2 = s^2 = \frac{4.12^2 + 5.12^2 + ... - 8 \times \bar{x}^2}{7}$	M1
		$\hat{\sigma}^2 = s^2 = \frac{141.4035 - 138.528013}{7} = 0.41078...$ (awrt **0.411**)	A1
			(3)
	(b)	$\sum x = 33.29 + 32 \times 4.55 = 178.89,$ (awrt **179**)	B1
		$\sum x^2 = "141.4035" + 31 \times 0.25 + 32 \times 4.55^2 (= 811.6335)$ (awrt **812**)	M1A1
		Combined sample: $s^2 = \frac{811.6335 - \frac{178.89^2}{40}}{39} = 0.29724865...$ (awrt **0.297**)	M1A1
		$\frac{s}{\sqrt{n}} = \frac{\sqrt{0.297....}}{\sqrt{40}} = 0.0862$ (awrt **0.0862)**	M1A1
			(7)
	(c)	$\bar{x} \pm 1.96\frac{\sigma}{\sqrt{n}} = \frac{178.89}{40} \pm 1.96\frac{0.67}{\sqrt{40}}$	M1B1
		= (4.2646…, 4.67988…) awrt (**4.26**[or 4.265]**, 4.68**)	A1
			(3)
			Total 13

	Notes
(a)	M1 for an attempt at s^2: correct denom, clear attempt at $\sum x^2$ and ft their $\bar{x}$ Ans only 2/2
(b)	B1 for correct sum or mean or fully correct expression (accept mean = awrt 4.47) **May be in (c)** 1st M1 for their $141.4035 + 31 \times 0.25 + 32 \times 4.55^2$ or "141.4035" + 7.75+ 662.48 (accept 3sf) **Beware:** $32(0.25 + 4.55^2)$ + "141.4035" = awrt 812 but scores M0A0. 1st A1 for a fully correct expression (all to 3sf or better) or answer only = awrt 812 2nd M1 for a correct expression using their values 3rd M1 dependent on using a changed s^2 (not their 0.411 or 0.25) for $\frac{\sqrt{"0.297"}}{\sqrt{40}}$ This s^2 must be based on a combination of their 0.411 and 0.25 e.g. 0.661
(c)	M1 for $\bar{x} \pm z \times \frac{\sigma}{\sqrt{n}}$ for any z (> 1.5) and ft their $\bar{x}$ based on combining their 4.16 and 4.55, do not award for simply using 4.55 or their 4.16. Condone $\sigma = \sqrt{\text{their } 0.297}$ or their (b) B1 for z = 1.96 used in an attempt at a CI, may for example miss $\sqrt{n}$ A1 for both limits awrt 3sf. Allow lower limit of 4.265