



GuideMyGrades

— YOUR GOALS. OUR EXPERTISE. —

EDEXCEL INTERNATIONAL A LEVEL MATHS: QUESTIONS BY TOPIC

Estimation, Confidence Intervals & Tests (S3.3)

International A Level Mathematics · S3: Statistics 3 (WST03)

35 questions · 365 marks · Mark schemes included

Past papers October 2020 - June 2025 · 2 sub-topics

 guidemygrades.com

 @Guidemygrades

 @Guidemygrades

 @Guidemygrades

 info@guidemygrades.com

 Piccadilly Business Centre, Unit C, Aldow Enterprise Park
Blackett Street, Manchester, M12 6AE

HOW TO USE THIS BOOKLET

Questions are grouped by sub-topic and run from the lowest to the highest mark tariff, then oldest paper first.

Click any row below to jump straight to that sub-topic or question.

On every question page, click ANSWER in the top-right corner to jump to its mark scheme; < QUESTION brings you back.

A gold bar in the left margin marks exactly where the question sits on the original page.

Diagrams, graphs, grids and tables are printed exactly as they appear in the original papers.

Each question is filed under its main topic; longer questions may also test skills from other topics.

QUESTION	SOURCE	MARKS	PAGE
S3.3a CENTRAL LIMIT THEOREM & CONFIDENCE INTERVALS		26 Qs · 272 marks	3
1. Q1	IAL S3 (WST03) October 2020	4	4
2. Q6	IAL S3 (WST03) October 2020	8	6
3. Q3	IAL S3 (WST03) June 2021	8	8
4. Q5	IAL S3 (WST03) October 2021	8	10
5. Q2	IAL S3 (WST03) (unused paper) January 2022	8	12
6. Q2	IAL S3 (WST03) January 2022	8	14
7. Q7	IAL S3 (WST03) June 2024	8	18
8. Q6	IAL S3 (WST03) June 2025	8	23
9. Q5	IAL S3 (WST03) June 2022	9	25
10. Q3	IAL S3 (WST03) June 2023	9	29
11. Q6	IAL S3 (WST03) January 2025	9	33
12. Q1	IAL S3 (WST03) January 2022	10	37
13. Q4	IAL S3 (WST03) June 2022	10	41
14. Q6	IAL S3 (WST03) January 2023	10	43
15. Q4	IAL S3 (WST03) June 2025	11	47
16. Q7	IAL S3 (WST03) June 2025	11	51
17. Q5	IAL S3 (WST03) October 2020	12	55
18. Q6	IAL S3 (WST03) October 2021	12	59
19. Q3	IAL S3 (WST03) June 2024	12	63
20. Q4	IAL S3 (WST03) January 2025	12	67
21. Q5	IAL S3 (WST03) June 2023	13	71
22. Q4	IAL S3 (WST03) January 2021	14	75
23. Q3	IAL S3 (WST03) June 2022	14	79
24. Q5	IAL S3 (WST03) June 2024	14	83
25. Q5	IAL S3 (WST03) (unused paper) January 2022	15	87
26. Q6	IAL S3 (WST03) January 2024	15	91
S3.3b HYPOTHESIS TESTS FOR MEANS		9 Qs · 93 marks	95
27. Q1	IAL S3 (WST03) October 2021	5	96
28. Q2	IAL S3 (WST03) June 2021	9	98
29. Q5	IAL S3 (WST03) January 2024	9	100
30. Q5	IAL S3 (WST03) January 2023	10	102
31. Q2	IAL S3 (WST03) June 2022	11	106
32. Q1	IAL S3 (WST03) January 2023	12	110
33. Q3	IAL S3 (WST03) January 2025	12	114



QUESTION	SOURCE	MARKS	PAGE
34. Q5	IAL S3 (WST03) January 2025	12	118
35. Q6	IAL S3 (WST03) June 2023	13	122



QUESTIONS

35 questions · 365 marks · 2 sub-topics

SUB-TOPIC S3.3a

CENTRAL LIMIT THEOREM & CONFIDENCE INTERVALS

Sub-topic S3.3a · 26 questions · 272 marks · lowest tariff first



Q1

Question 3 continued

(Total for Question 3 is 8 marks)



Question 5 continued

Leave
blank

Lined area for writing the answer to Question 5.

(Total 8 marks)

Q5

Marking box for Q5.



Leave
blank

2. Krishni owns a farm on which he keeps chickens.

He selects, at random, 10 of the eggs produced and weighs each of them.

You may assume that these weights are a random sample from a normal distribution with standard deviation 1.9 g

The total weight of these 10 eggs is 537.2 g

- (a) Find a 95% confidence interval for the mean weight of the eggs produced by Krishi's chickens.

(4)

Krishi was hoping to obtain a 99% confidence interval of width at most 1.5 g

- (b) Calculate the minimum sample size necessary to achieve this.

(4)

Question 2 continued

Leave
blank

(Total 8 marks)

Q2







(Total for Question 2 is 8 marks)



7. The continuous random variable D is uniformly distributed over the interval $[x - 1, x + 5]$ where x is a constant.

A random sample of n observations of D is taken, where n is large.

- (a) Use the Central Limit Theorem to find an approximate distribution for $\bar{D}$. Give your answer in terms of n and x where appropriate.

(3)

The n observations of D have a sample mean of 24.6

Given that the lower bound of the 99% confidence interval for x is 22.101 to 3 decimal places,

- (b) find the value of n
Show your working clearly.

(5)

Question 7 continued

Lined area for writing the answer to Question 7 continued.



Question 7 continued



Question 7 continued



(Total for Question 7 is 6 marks)

TOTAL FOR PAPER IS 75 MARKS



Question 6 continued

Lined area for writing the answer to Question 6.

(Total for Question 6 is 8 marks)



Question 5 continued



Question 5 continued

Lined area for writing the answer to Question 5 continued.



Question 5 continued

(Total for Question 5 is 9 marks)

Question 3 continued



Question 3 continued

Lined area for writing the answer to Question 3 continued.



Question 3 continued

Lined area for writing the answer to Question 3.

(Total for Question 3 is 9 marks)







(Total for Question 6 is 7 marks)







(Total for Question 1 is 10 marks)



4. Navtej travels to work by train. A train leaves the station every 7 minutes and Navtej's arrival at the station is independent of when the train is due to leave.

(a) Write down a suitable model for the distribution of the time, T minutes, that he has to wait for a train to leave.

(1)

(b) Find the mean and standard deviation of T

(3)

During a 10-week period, Navtej travels to work by train on 46 occasions.

(c) Estimate the probability that the mean length of time that he has to wait for a train to leave is between 3.4 and 3.6 minutes.

(5)

(d) State a necessary assumption for the calculation in part (c).

(1)

Question 4 continued

Lined area for writing the answer to Question 4.

(Total for Question 4 is 10 marks)



This image shows a single sheet of white paper with horizontal ruling lines. The lines are evenly spaced and run across the width of the page. There are no margins, text, or other markings on the paper.



This image shows a single sheet of white paper with horizontal ruling lines. The lines are evenly spaced and run across the width of the page. There are no margins, text, or other markings on the paper.



(Total for Question 6 is 10 marks)



Question 4 continued

Lined area for writing the answer to Question 4 continued.



Question 4 continued



Question 4 continued

Lined area for writing the answer to Question 4.

(Total for Question 4 is 11 marks)



Question 7 continued

Lined area for writing the answer to Question 7 continued.



Question 7 continued

Lined area for writing the answer to Question 7.

(Total for Question 7 is 11 marks)



Leave
blank



Leave
blank



Q5

Leave
blank

Leave
blank

Question 3 continued

Lined area for writing the answer to Question 3 continued.





(Total for Question 3 is 12 marks)

Question 4 continued





(Total for Question 7 is 12 marks)



Question 5 continued

Lined area for writing the answer to Question 5.



Question 5 continued

Lined area for writing the answer to Question 5 continued.



(Total for Question 5 is 13 marks)

Leave
blank

4. The scores in a national test of seven-year-old children are normally distributed with a standard deviation of 18
A random sample of 25 seven-year-old children from town *A* had a mean score of 52.4

- (a) Calculate a 98% confidence interval for the mean score of the seven-year-old children from town A .

(4)

An independent random sample of 30 seven-year-old children from town B had a mean score of 57.8

A local newspaper claimed that the mean score of seven-year-old children from town B was greater than the mean score of seven-year-old children from town A .

- (b) Stating your hypotheses clearly, use a 5% significance level to test the newspaper's claim. You should show your working clearly.

(6)

The mean score for the national test of seven-year-old children is μ .

Considering the two samples of seven-year-old children separately, at the 5% level of significance, there is insufficient evidence that the mean score for town A is less than μ , and insufficient evidence that the mean score for town B is less than μ .

- (c) Find the largest possible value for μ .

(4)



Leave
blank

Leave
blank



Leave
blank

Q4



Question 3 continued

Lined area for writing the answer to Question 3 continued.



Question 3 continued

Lined area for writing the answer to Question 3.

(Total for Question 3 is 14 marks)



Question 5 continued

Lined area for writing the answer to Question 5 continued.



Question 5 continued

Lined area for writing the answer to Question 5 continued.



(Total for Question 5 is 14 marks)



Question 5 continued

Leave
blank





(Total for Question 6 is 15 marks)



SUB-TOPIC S3.3b

HYPOTHESIS TESTS FOR MEANS

Sub-topic S3.3b · 9 questions · 93 marks · lowest tariff first



Leave
blank

1. A machine makes screws with a mean length of 30 mm and a standard deviation of 2.5 mm.

A manager claims that, following some repairs, the machine is now making screws with a mean length of less than 30 mm. The manager takes a random sample of 80 screws and finds that they have a mean length of 29.5 mm.

Use a suitable test, at the 5% level of significance, to determine whether there is evidence to support the manager's claim. State your hypotheses clearly.

(5)

Question 1 continued

Leave
blank

Lined area for writing the answer to Question 1.

(Total 5 marks)

Q1

Mark box for Question 1.



2. A doctor believes that the diet of her patients and their health are not independent.

She takes a random sample of 200 patients and records whether they are in good health or poor health and whether they have a good diet or a poor diet. The results are summarised in the table below.

	Good health	Poor health
Good diet	86	8
Poor diet	91	15

Stating your hypotheses clearly, test the doctor's belief using a 5% level of significance. Show your working for your test statistic and state your critical value clearly.

(9)

(Total for Question 2 is 9 marks)



(Total for Question 3 is 7 marks)



This image shows a single sheet of white paper with horizontal ruling lines. The lines are evenly spaced and run across the width of the page. There are no margins, text, or other markings on the paper.



This image shows a single sheet of white paper with horizontal ruling lines. The lines are evenly spaced and run across the width of the page. There are no margins, text, or other markings on the paper.



(Total for Question 3 is 10 marks)



Question 2 continued

Lined area for writing the answer to Question 2 continued.



Question 2 continued

Lined area for writing the answer to Question 2.



(Total for Question 2 is 11 marks)



1 A machine fills bottles with mineral water.

The machine is checked every day to ensure that it is working correctly. On a particular day a random sample of 100 bottles is taken. The volume of water, x millilitres, for each bottle is measured and each measurement is coded using

$$y = x - 1000$$

The results are summarised below

$$\sum y = 847 \quad \sum y^2 = 13\,510.09$$

- (a) (i) Show that the value of the unbiased estimate of the mean of x is 1008.47
(ii) Calculate the unbiased estimate of the variance of x

(4)

The machine was initially set so that the volume of water in a bottle had a mean value of 1010 millilitres.

Later, a test at the 5% significance level is used to determine whether or not the mean volume of water in a bottle has changed. If it has changed then the machine is stopped and reset.

- (b) Write down suitable null and alternative hypotheses for a 2-tailed test.

(1)

- (c) Find the critical region for $\bar{X}$ in the above test.

(4)

- (d) Using your answer to part (a) and your critical region found in part (c), comment on whether or not the machine needs to be stopped and reset.
Give a reason for your answer.

(2)

- (e) Explain why the use of $\sigma^2 = s^2$ is reasonable in this situation.

(1)

[illegible]



This image shows a single sheet of white paper with horizontal ruling lines. The lines are evenly spaced and run across the width of the page. There are no margins, text, or other markings on the paper.



(Total for Question 1 is 12 marks)







(Total for Question 9 is 12 marks)

Question 5 continued

Lined area for writing the answer to Question 5 continued.





(Total for Question 3 is 12 marks)



Question 6 continued

Lined area for writing the answer to Question 6 continued.



Question 6 continued

Lined area for writing the answer to Question 6 continued.



(Total for Question 6 is 15 marks)



MARK SCHEMES

official mark schemes, in the same order as the questions



Question Number	Scheme	Marks
1.		
(a)	$E(2\bar{X}) = 2E(\bar{X}) = \frac{2(1+\alpha)}{2}$	M1
	$= 1 + \alpha \neq \alpha$ (therefore $2\bar{X}$ is a biased estimator of α)	A1
		(2)
(b)	$\bar{x} = 6 \quad 2\bar{x} - 1 =$	M1
	<u>11</u>	A1
		(2)
		Total 4
	Notes	
(a)	M1 for use of $2 \times \frac{1+\alpha}{2}$ Also correct comparison with α and no incorrect working seen. $\frac{1+\alpha}{2} \neq \alpha$ is M0A0	
(b)	M1 for attempt at $\bar{x}$ and use of $2\bar{x} - 1$ or ft use of their $E(\bar{X})$ from part (a) to find an estimate for α A1 11 cao	

Question Number	Scheme	Marks
6.		
(a)	$\bar{T} \sim N(4, \frac{4}{35})$	M1 A1
		(2)
(b)	$\bar{K} \sim N(\lambda, \frac{\lambda}{40})$	M1
	$2 \times 2.5758 \times \sqrt{\frac{\lambda}{40}} = 2.6$	B1 M1
	$\lambda = \text{awrt } \underline{10.2}$	A1
		(4)
(c)	$2 \times 0.99 \times 0.01$	M1
	$= \underline{0.0198}$	A1
		(2)
		Total 8
	Notes	
(a)	M1 for Normal distribution A1 for correct mean and variance (allow N(4, awrt 0.114))	
(b)	M1 for use of $\frac{\lambda}{40}$ if $\lambda = \sigma$ is used, then M0 B1 for ± 2.5758 (may be implied by sight of 10.188...) M1 for use of $2 \times z \times \frac{\sigma'}{\sqrt{40}} = 2.6$ with $ z > 2$ A1 awrt 10.2 (an answer of 10.15...or awrt 10.2 on its own scores M1B0M1A1) SC: Use of $\sqrt{\lambda}$ instead of λ leading to an answer of awrt 3.19... scores M0B1M1A0	
(c)	M1 for $2p(1-p)$ for any p $0 < p < 1$ A1 0.0198	

Question Number	Scheme	Notes	Marks
3.			
3(a)	$\bar{x} = \frac{1}{2}(11.52 + 13.75) = 12.635$	12.635 (may be implied by correct CI)	B1
		Use of 1.96	B1
	$\left(\frac{\sigma}{\sqrt{n}} = \frac{13.75 - 12.635}{1.96} (= 0.56887...)\right)$ $\left(\frac{\sigma}{\sqrt{n}} = \frac{13.75 - 11.52}{2 \times 1.96} (= 0.56887...)\right)$	For attempt at standard error (may be implied by awrt 0.569)	M1
		Use of 1.6449 or better (1.644853... from calc) Use of 1.64 or 1.65 is B0	B1
	$12.635 \pm 1.6449 \times 0.56887...$	For $(\text{their } \bar{x}) \pm (\text{their } 1.6449) \left(\text{their } \frac{\sigma}{\sqrt{n}} \right)$ their $\frac{\sigma}{\sqrt{n}}$ must be numerical	M1
	90% CI is (11.699..., 13.5707...)	awrt (11.7, 13.6) from correct working Correct answer with no working scores B1B1M1B0M1A1	A1
			(6)
3(b)	$4 \times 0.9^3 \times 0.1$ $= 0.2916$	$4p^3(1-p)$ (where $0 < p < 1$) awrt 0.292	M1 A1
			(2)
			Total 8

Question Number	Scheme		Marks
5(a)	$2.977 \pm 2.5758 \times \frac{0.015}{3}$		M1,B1
	$= (2.9641..., 2.9898...)$ awrt (2.964, 2.990)		A1
			(3)
(b)	The CI does not contain the stated weight.		B1
			(1)
(c)	$2.995 - 1.96 \times \frac{0.015}{\sqrt{n}} < 2.991$		M1
	$\sqrt{n} < \frac{1.96 \times 0.015}{2.995 - 2.991}$		M1d
	$\sqrt{n} < \text{awrt } 7.35$		A1
	$n = 54$		A1cao
			(4)
	Notes		Total 8
(a)	M1	$2.977 \pm (z \text{ value}) \times \frac{0.015}{3}$	
	B1	awrt 2.5758	
	A1	awrt (2.964, 2.990 (condone 2.99))	
(b)	B1	cao this must be consistent with their confidence interval	
(c)	M1	Setting up an inequality using z value > 1.5 Condone =	
	M1d	Dep on previous M mark. Correct rearranging to get $\sqrt{n} < \dots$ or $n < \dots$ Condone = or >	
	A1	awrt 7.35 may be implied by awrt 54	
	A1cao	54	

Question Number	Scheme		Marks
2 (a)	Use of $\bar{x} \pm z \times \frac{1.9}{\sqrt{10}}$; $z = 1.96$		M1;B1
	(52.54..., 54.897...)	awrt 52.5 and 54.9	A1 A1
			(4)
(b)	Use of $1.5 > 2 \times z \times \frac{1.9}{\sqrt{n}}$ oe ; $z = 2.5758$ (or better)		M1;B1
	$1.5 > \frac{9.78804}{\sqrt{n}}$		dM1
	$n > 42.58...$ So $n = 43$		A1
			(4)
	Notes		Total 8
2 (a)	M1	for use of correct expression with 1.9, 10 and $1 < z < 3$	
	B1	for $z = 1.96$	
	A1	for awrt 52.5	
	A1	for awrt 54.9	
(b)	M1	use of $z \times \frac{1.9}{\sqrt{n}}$ in a correct inequality with 0.75 or 1.5 and $2 < z < 3$ (allow written as an equation)	
	B1	for $z = 2.5758$ (or better)	
	dM1	dependent on 1 st M1, for solving a correct inequality for the width of the 99% CI (allow an equation rather than an inequality)	
	A1	cao	

Question Number	Scheme		Marks
2(a)	$H_0 : \mu_{year7} = \mu_{year8} \quad H_1 : \mu_{year7} \neq \mu_{year8}$		B1
	$SE = \sqrt{\frac{38}{240} + \frac{42}{240}}$		M1
	$z = \frac{103 - 101}{SE}$		M1
	$= (\pm)3.464.... \quad (2\sqrt{3})$ awrt $(\pm) 3.46$		A1
	$Z_{critical} = 2.5758$		B1
	In CR/Significant/Reject H_0		M1
	There is sufficient evidence to suggest that the regional education <u>officer</u> 's claim is not correct/ There is a difference between the <u>mean scores</u> of the two year groups.		A1
(b)	CLT allows us to use <u>sample means</u> (oe) being normally distributed		(7) B1
			(1)
	Notes		Total 8
(a)	B1	both hypotheses correct. Allow equivalent rearrangements. Must be in terms of μ If using e.g. $\mu_A = \mu_B$ A and B must be clearly identified with year groups	
	M1	for use of SE with 38 and 42 (may be implied by SE = awrt 0.577)	
	M1	for a correct standardisation expression using 103, 101 (in either order) and SE = awrt.0577 or ft their stated SE or if not stated (i.e. only seen in standardisation) only allow $\sqrt{\frac{38^2}{240} + \frac{42^2}{240}}$ or $\sqrt{\frac{\sqrt{38}}{240} + \frac{\sqrt{42}}{240}}$	
	A1	awrt 3.46 or awrt -3.46 allow p value of awrt 0.000266	
	B1	$ CV = 2.5758$ or better (seen)	
	M1	a correct statement linking their test statistic and their CV – need not be contextual but do not allow contradicting non contextual comments.	
	A1	do not allow a ft conclusion here. a correct contextual statement (dependent on 2 nd M1) which must be consistent with their test statistics and CV and which also must reject H_0 . It must mention the officer or mean scores.	
(b)	B1	a correct explanation which must mention sample means oe (population means are normally distributed is B0) ignore extraneous non-contradictory comments	

Question Number	Scheme	Marks
7(a)	$E(D) = x + 2$	M1
	$\text{Var}(D) = \frac{((x+5)-(x-1))^2}{12} [= 3]$	M1
	$\bar{D} \sim N\left(x+2, \frac{3}{n}\right)$	A1
		(3)
(b)	" $x+2$ " = 22.101 + "2" (= 24.101) or " $x+2$ " = 24.6 $\Rightarrow$ 24.6 – "2" (= 22.6)	M1
	$24.6 - "2.5758" \sqrt{\frac{"3"}{n}} = "24.101" \text{ oe}$	B1M1 dM1
	$n = 80$	A1cao
		(5)
	Notes	Total 8
(a)	M1 $E(D)$ correct	
	M1 Correct method to find $\text{Var}(D)$ Must be subtracting the correct way round but condone missing brackets	
	A1 for a fully correct distribution. Either states $N\left(x+2, \frac{3}{n}\right)$ or accept e.g. "normal" with mean = $x+2$ and variance = $\frac{3}{n}$ oe Must be seen in (a)	
(b)	M1 For a correct method to find d using x as 22.101 in their " $x+2$ " from (a) or a correct method to find x by rearranging their " $x+2$ " to $x = 24.6 - "2"$ Implied by 24.101 or 22.6 or ± 0.499 oe $x+2 = 24.6$ on its own is M0	
	B1 for awrt ± 2.5758 may be implied by an unrounded value for n of awrt 79.94	
	M1 for $24.6 \pm z \sqrt{\frac{\sigma^2}{n}}$ or " $22.6 \pm z \sqrt{\frac{\sigma^2}{n}}$ where $2.55 < z < 2.6$ (ft their mean and variance from (a) or may restart). May be part of an equation. Their numerical variance does not have to be substituted in for this mark. This mark can still be scored if it is equated inconsistently to 22.101 eg $24.6 - z \sqrt{\frac{\sigma^2}{n}} = 22.101$ oe eg $z \sqrt{\frac{\sigma^2}{n}} = 2.499$ or $\sqrt{\frac{\sigma^2}{n}} = \frac{4165}{4293} (= 0.970...)$	
	dM1 dep on both the previous M marks awarded. For setting up a valid equation (allow 22.101 instead of "24.101" provided it is correctly paired with "22.6"). Their numerical variance must be substituted in for this mark. $24.6 - "2.5758" \sqrt{\frac{"3"}{n}} = "24.101"$ or " $22.6 - "2.5758" \sqrt{\frac{"3"}{n}} = 22.101$ oe eg " $-2.5758 = \frac{"24.101" - 24.6}{\sqrt{\frac{"3"}{n}}}$ or " $2.5758 \sqrt{\frac{"3"}{n}} = "0.499"$	
	A1 cao dependent on seeing a correct equation but allow use of $z = 2.576$ so M1B0M1dM1A1 is possible. Note awrt 79.94 seen can imply B1	



Question Number	Scheme		Marks
6 (a)	e.g. based solely on known observations or e.g. contains no unknown (population) parameters		B1
			(1)
(b)	$\left[E(R) = \frac{1}{2}E(X_1) + \frac{3}{4}E(X_{20}) \right] = \frac{1}{2}\mu + \frac{3}{4}\mu$		M1
	$E(R) = \frac{1}{2}\mu + \frac{3}{4}\mu = \frac{5}{4}\mu \neq \mu$ e.g. So R is a biased estimator for μ		A1
			(2)
(c)	$\frac{5}{4}\mu - \mu = \frac{1}{4}\mu$		B1ft
			(1)
(d)	One of $a - b = 2$ oe or $14a + 6b = 20$ oe or $\frac{a - b}{2} = \frac{14a + 6b}{20}$ $[-4b = a]$ oe		B1
	Two of $a - b = 2$ oe , $14a + 6b = 20$ oe , $\frac{a - b}{2} = \frac{14a + 6b}{20}$ oe		B1
	Solves simultaneously e.g. $14a + 6(a - 2) = 20 \Rightarrow a = \dots, b = \dots$		M1
	$a = 1.6$ oe, $b = -0.4$ oe		A1
			(4)
Notes			Total 8
(a)	B1	For a correct explanation. Allow a valid explanation which suggests one of the following that it is - based (solely) on observations/calculations/values/information/data oe - contains no unknown (population) parameters - calculated/measured - numerical property of a sample/derived from a sample Do not allow “because it is known” (too vague) but condone “because X_1 and X_{20} are known”	
(b)	M1	For use of $E(R) = aE(X_1) + bE(X_{20})$. May be implied by a correct expression or $\frac{5}{4}\mu$	
	A1	Requires an expression in terms of μ before proceeding to a correct simplified value of $\frac{5}{4}\mu$, a reason (e.g. $\neq \mu$) and a conclusion (e.g. biased estimator). Must see use of expectation notation e.g. $E(R)$ or $\frac{1}{2}E(X_1) + \frac{3}{4}E(X_{20})$	
(c)	B1ft	Follow through their part (b) – μ Their part (b) must be in terms of μ . May be seen in (b)	
(d)	B1	For a correct equation in a and b only. μ must not be present	
	dB1	Dep on 1st B1. For a second different correct equation in a and b only. μ must not be present	
	M1	Dependent on at least one correct equation. For solving simultaneously and proceeding to values for a and b . You do not need to be concerned with the mechanics of the rearrangement as long as values for a and b are achieved. May be implied by a correct value for a or a correct value for b .	
	A1	For correct values of a and b	

Question Number	Scheme		Marks
5 (a)	It is not a statistic as it involves <u>unknown</u> [population parameters]		B1
			(1)
(b)	An estimator for μ is unbiased if its <u>expected</u> value is equal to μ		B1
			(1)
(c)	$E(U_1) = 3E(X_1) - 2E(X_2)$ or $E(U_2) = \frac{1}{4}(E(X_1) + 3E(X_2))$		M1
	$E(U_1) = 3\mu - 2\mu = \mu$ (therefore unbiased)		A1cso
	$E(U_2) = \frac{1}{4}(\mu + 3\mu) = \mu$ (therefore unbiased)		A1cso
			(3)
(d)	$\text{Var}(U_1) = 9\text{Var}(X_1) + 4\text{Var}(X_2)$ or $\text{Var}(U_2) = \frac{1}{16}\text{Var}(X_1) + \frac{9}{16}\text{Var}(X_2)$		M1
	$[\text{Var}(U_1) =]13\sigma^2$		A1
	$[\text{Var}(U_2) =]\frac{5}{8}\sigma^2$		A1
	As $\text{Var}(U_1) > \text{Var}(U_2)$ U_2 is the most efficient estimator for μ		A1
			(4)
Notes			Total 9
(a)	B1	for a correct explanation, must include unknown	
(b)	B1	for a correct explanation that refers to expected X . Allow $\mu - E(X) = 0$, but bias = 0 is B0	
(c)	M1	for use of $aE(X_1) + bE(X_2)$ May be implied by $3\mu - 2\mu$ or $\frac{1}{4}(\mu + 3\mu)$	
	A1cso	for a correct solution for $E(U_1)$ with no incorrect working Condone missing notation. Condone missing subscripts	
	A1cso	for a correct solution for $E(U_2)$ with no incorrect working seen Condone missing notation. Condone missing subscripts	
(d)	M1	for use of $a^2\text{Var}(X_1) + b^2\text{Var}(X_2)$	
	A1	Allow $9\sigma^2 + 4\sigma^2$	
	A1	Allow $\frac{1}{16}\sigma^2 + \frac{9}{16}\sigma^2$ or $\frac{5}{8}\sigma^2$ oe	
	A1	for U_2 with a correct reason	
		NB It is possible to score M1 A0 A0 A1 if $\text{Var}(U_1)$ and $\text{Var}(U_2)$ are correct	

Question Number	Scheme		Marks
3 (a)	It is not a statistic as it involves <u>unknown</u> [population] parameter		B1
			(1)
(b)	$E(S) = E\left(\frac{3}{5}X_1 + \frac{5}{7}X_2\right) = \frac{3}{5}E(X_1) + \frac{5}{7}E(X_2)$		M1
	$= \frac{3}{5}\mu + \frac{5}{7}\mu = \frac{46}{35}\mu \neq \mu$ So S is a biased estimator for μ		A1
(c)			(2)
	$\frac{46}{35}\mu - \mu = \frac{11}{35}\mu$		B1ft
(d)			(1)
	$E(Y) = aE(X_1) + bE(X_2) = \mu$ $\Rightarrow (a + b)\mu = \mu$		M1
(e)	$a + b = 1$		A1
			(2)
(e)	$\text{Var}(Y) = a^2\text{Var}(X_1) + b^2\text{Var}(X_2) = (a^2 + b^2)\sigma^2$		M1
	$\text{Var}(Y) = (a^2 + (1 - a)^2)\sigma^2$		M1
(e)	$\text{Var}(Y) = (2a^2 - 2a + 1)\sigma^2$ *		A1*
			(3)
	Notes		Total 9
(a)	B1	For a correct explanation Allow σ is unknown (Do not allow σ is unknown variance)	
(b)	M1	For writing or using $E(S) = aE(X_1) + bE(X_2)$ Condone missing subscripts	
	A1	cao (Allow $1.31\mu \neq \mu$)	
(c)	B1ft	Follow through their part (a) – μ	
(d)	M1	For writing or using $E(Y) = aE(X_1) + bE(X_2) = \mu$ (May be implied by $a + b = 1$) Condone missing subscripts	
	A1	Cao	
(e)	M1	For writing or using $\text{Var}(Y) = a^2\text{Var}(X_1) + b^2\text{Var}(X_2)$ Condone missing subscripts	
	M1	For substitution of $b = 1 - a$ ft their part (d) into their expression for $\text{Var}(Y)$	
	A1*	Answer is given so no incorrect working must be seen	

Question Number	Scheme		Marks
6 (a)	$\left[E(\bar{X}) \right] = \frac{2a+3+4a+9}{2}$		M1
	$= \frac{6a+12}{2} = 3a+6 \neq a^* \quad (\text{So biased})$		A1*
			(2)
(b)	$'(3a+6)' - a = 2a+6$		B1ft
			(1)
(c)	$c = \frac{1}{3}$		B1ft
	$\frac{1}{3} \times '(3a+6)' + d = a$		M1
	$d = -2$		A1
			(3)
(d)	$\frac{1}{3} \times 7.32 - '2' [= 0.44] \quad \text{or} \quad 3a+6 = 7.32 [\Rightarrow a = 0.44]$		M1
	$4 \times '0.44' + 9$		M1
	$= 10.76$		A1
			(3)
Notes			Total 9
(a)	M1	For using the formula $\left(\frac{a+b}{2} \right)$ May be implied by $\frac{6a+12}{2}$ or $3a+6$	
	A1*	For $\frac{6a+12}{2}$ or $3a+6$ and $\neq a$ (Allow $3a+6-a$ or $2a+6$ and $\neq > 0$)	
(b)	B1ft	For $2a+6$ or ft their part (a)	
(c)	B1	For $\frac{1}{3}$ or $\frac{1}{\text{coefficient of } a \text{ (from part a)}}$	
	M1	For $c \times \text{their } (3a+6) + d = a$ oe written or used May be implied by $d = -2$	
	A1	Cao	
(d)	M1	For their $c \times 7.32 - \text{their } d$ oe or $7.32 = '3a+6'$	
	M1	For $4 \times \text{their } 0.44 + 9$	
	A1	cao Do not ISW but condone rounding	

Question Number	Scheme		Marks
1 (a)	$\bar{x} = 11.42$		B1
	$s^2 = \frac{1310.464 - 10 \times 11.42^2}{9}$		M1
	$= 0.7$		A1
			(3)
(b)	z value for 95% CI is 1.96		B1
	$11.42 \pm 1.96 \times \frac{0.8}{\sqrt{10}}$		M1
	(10.924..., 11.915...)	awrt (10.92, 11.92)	A1 A1
			(4)
(c)	$Y \sim N(11.92, 0.8^2)$		M1
	$P(Y < 10.5) = P\left(Z < \frac{10.5 - 11.92}{0.8}\right) = P(Z < -1.775)$		M1
	$= 0.03837...$	awrt 0.038	A1
			(3)
	Notes		Total 10
1(a)	B1	for 11.42 cao	
	M1	for use of $s^2 = \frac{\sum x^2 - n\bar{x}^2}{n-1}$	
	A1	for 0.7 cao	
(b)	B1	for writing or using 1.96 (or better from calculator 1.9599...)	
	M1	For use of $\bar{x} \pm z \text{ value} \times \frac{\sigma}{\sqrt{n}}$ ft their z value, $1 < z < 2$ and their 11.42	
	A1	for awrt 10.9 or awrt 11.9	
	A1	for awrt 10.92 and awrt 11.92	
(c)	M1	for identifying the normal distribution with the upper confidence interval value as mean and 0.8 as standard deviation (may be seen in standardisation)	
	M1	for standardising with 10.5, their mean (which must be in their confidence interval (including limits) from part (b)) and standard deviation = 0.8	
	A1	awrt 0.038 (tables = 0.0375)	



Question Number	Scheme		Marks
4 (a)	[Continuous] uniform on the interval [0, 7]		B1
			(1)
(b)	mean = 3.5		B1
	standard deviation = $\sqrt{\frac{(7-0)^2}{12}}$		M1
	$= \frac{7}{\sqrt{12}} = 2.0207...$ awrt 2.02		A1
			(3)
(c)	By the CLT $\bar{T} \sim N\left(3.5, \frac{49}{552}\right)$		M1
	$P\left(3.4 < \bar{T} < 3.6\right) = P\left(\frac{3.4 - "3.5"}{"\sqrt{\frac{49}{552}}"} < Z < \frac{3.6 - "3.5"}{"\sqrt{\frac{49}{552}}"}\right) = \left[P\left(-0.34 < Z < 0.34\right)\right]$		M1 A1
	$= 0.6331 - (1 - 0.6331)$ (Calculator gives 0.6314...)		M1
	$= 0.2662$ (Calculator gives 0.2628...) awrt 0.263 to 0.266		A1
			(5)
(d)	Large/ independent/ random sample allows use of CLT		B1
			(1)
	Notes		Total 10
(a)	B1	For the correct distribution stated (need uniform and correct interval) Allow U[0, 7] A fully correct pdf implies B1 e.g. $f(x) = \begin{cases} \frac{1}{7} & 0 \leq x \leq 7 \\ 0 & \text{otherwise} \end{cases}$	
(b)	B1	For 3.5	
	M1	For a correct method for finding the standard deviation	
	A1	awrt 2.02 (Allow $\frac{7}{\sqrt{12}}$ or $\frac{7\sqrt{3}}{6}$ oe)	
(c)	M1	For writing or using $N\left(3.5, \frac{49}{552}\right)$ oe Allow $N\left(3.5, \frac{2.02^2}{46}\right)$ or ft from part (b) e.g. if Po(7) given in part (a) allow $N\left(7, \frac{7}{46}\right)$	
	M1	For standardising using either 3.4 or 3.6 and their mean and standard deviation	
	A1	For a fully correct expression for either 3.4 or 3.6. May be implied by $\pm$ awrt 0.34	
	M1	For $p - (1 - p)$ or $2(p - 0.5)$ oe	
	A1	awrt 0.263 to 0.266	
(d)	B1	Any suitable assumption	

Question	Scheme		Marks
6 (a)	$\left[\bar{x} = \frac{806.4}{36} = \right] 22.4$		B1
	$"22.4" \pm 2.3263 \times \frac{0.4}{\sqrt{36}}$		M1 B1
	$(22.24\dots, 22.55\dots)$ awrt (22.2, 22.6)		A1
	NB answers which are awrt (22.2, 22.6) gain full marks		
			(4)
(b)	[The Central Limit Theorem is not required as] the original population is normally distributed		B1
			(1)
(c)	22.5 is within the confidence interval		B1 ft
	So no reason to doubt the manufacturers claim		dB1 ft
			(2)
(d)	$\bar{Y} \sim N\left(850, \left(\frac{5}{\sqrt{10}}\right)^2\right)$		B1
	$P(\bar{Y} < 848) = P\left(Z < \frac{848 - 850}{\frac{5}{\sqrt{10}}}\right) = [P(Z < -1.26)]$		M1
	$= 0.1038$ (Calculator gives 0.10295...) awrt 0.103 / 0.104		A1
			(3)
ALT	N(8500, 250)		B1
	$P(\bar{Y} < 848) = P\left(Z < \frac{8480 - 8500}{\sqrt{250}}\right) = [P(Z < -1.26)]$		M1
	$= 0.1038$		A1
	Notes		Total 10
(a)	B1	For 22.4	
	M1	For use of $\bar{x} \pm z$ value $\times \frac{\sigma}{\sqrt{n}}$ with $1.2 < z < 2.6$	
	B1	For z value = 2.3263 or better seen (Calculator gives 2.3263479...)	
	A1	awrt (22.2, 22.6) This does not imply the B1	
(b)	B1	For reference to the data is modelled by normal distribution	
(c)	B1 ft	ft their CI For a comment on whether 22.5 (or it) is or is not in their CI allow eg range for CI Allow "22.24" < 22.5 < " 22.6" Answer must be compatible with their CI	
	dB1 ft	Dependent on B1 ft. For a correct comment ft their CI eg claim is correct oe	
(d)	B1	for $\bar{Y} \sim N(850, \dots)$ or $\bar{Y} < \frac{848 - 850}{5}$ Must have $\bar{Y}$ or $N\left(850, \left(\frac{5}{\sqrt{10}}\right)^2\right)$ or $N(850, 2.5)$ seen or used or $N(8500, 250)$ seen or used. Both implied by a correct standardisation.	
	M1	For $\pm$ (a correct standardisation) implied by a correct answer	
	A1	awrt 0.103 to 0.104	

Question Number	Scheme		Marks
4 (a)	$\left[\bar{y} = \frac{23200}{80} \Rightarrow \right] \bar{y} = 290$		B1
	$\left[s_y^2 = \right] \frac{6741351 - 80(290)^2}{80 - 1}$ or $\left[s_y^2 = \right] \frac{80 \left(\frac{6741351}{80} - \left(\frac{23200}{80} \right)^2 \right)}{79}$		M1
	$= 169$		A1
			(3)
(b)	$H_0: \mu_x - \mu_y = 200$ $H_1: \mu_x - \mu_y > 200$ oe		B1
	$z = \pm \frac{500 - '290' - 200}{\sqrt{\frac{'169'}{80} + \frac{1156}{80}}}$		M1 A1ft
	$= 2.45718...$ awrt 2.46		A1
	One tailed c.v. $Z = 1.6449$ or CR: $Z \dots 1.6449$		B1
	e.g. In CR/Significant/Reject H_0		M1
	Sufficient evidence to support the research student's belief		A1
			(7)
(c)	Assume that $s^2 = \sigma^2$		B1
			(1)
	Notes		Total 11
(a)	B1	For $\bar{y} = 290$	
	M1	A correct method for s_y^2 . May be implied by 169 on its own (not within a calculation)	
	A1	169 do not isw	
(b)	B1	Both hypotheses correct. Allow equivalent hypotheses. Must be in terms of μ . Use of $\bar{X} - \bar{Y}$ is B0. Allow g for x and r for y but use of any other letters must be defined.	
	M1	An attempt at $\pm \frac{500 - a - 200}{\sqrt{\frac{b}{80} + \frac{1156}{80}}}$ oe with at least 1 of a or b correct using their 290 or their 169 May be implied by awrt ± 2.46 or may state e.g. $N\left(\pm 200, \frac{'169'}{80} + \frac{1156}{80}\right)$ (μ may be 0)	
	A1ft	A correct expression [se $\sqrt{\frac{265}{16}} = 4.0697...]$ or imp by awrt ± 2.46 or p -value of awrt 0.0072	
	A1	awrt $z = 2.46$ (possibly $\pm$) or $P(\bar{X} - \bar{Y} > 210) = \text{awrt } 0.00716$ Correct answer scores M1A1ftA1 but $200 - (500 - 290) \rightarrow -10$ loses this 2nd A mark	
	B1	awrt ± 1.6449 or better seen or a comparison of p -value awrt $0.007 < 0.05$ oe	
	M1	Dep. on $1 < z < 3$ or $0 < P(\bar{X} - \bar{Y} > '210') < 0.5$ A correct statement ft on their z value and CV, with compatible signs, or prob. Context not needed but do not accept contradicting non contextual comments	
	A1	Dep on all previous M and A marks but independent of hypotheses. A correct contextual statement e.g. the mean weight of grey squirrels is more than 200g greater than (the mean weight) of red squirrels o.e. Must mention mean weight at least once if they mention weight or weights. Condone the difference (in mean weights) is greater than 200g	
(c)	B1	Assume that sample variance = population variance oe (but not e.g. $s_x = s_y$).	

Question Number	Scheme	Marks
7 (a)	$\left[P(\bar{X}_n < 45.2) = 0.3446 \right] \Rightarrow \frac{45.2 - \mu}{\frac{\sigma}{\sqrt{n}}} = -0.4$ oe	M1
	$45.2 + \frac{0.4\sigma}{\sqrt{n}} = \mu$ oe	A1
	$\left[P(\bar{X}_n > 75.2) = 0.0179 \right] \Rightarrow \frac{75.2 - \mu}{\frac{\sigma}{\sqrt{n}}} = 2.1$ oe	M1
	$75.2 - \frac{2.1\sigma}{\sqrt{n}} = \mu$ oe	A1
	e.g. $45.2 + \frac{0.4\sigma}{\sqrt{n}} = 75.2 - \frac{2.1\sigma}{\sqrt{n}}$ or e.g. $\frac{30}{\frac{\sigma}{\sqrt{n}}} = 2.5$ oe	M1
	$\frac{2.5\sigma}{\sqrt{n}} = 30 \Rightarrow [\sigma = 12\sqrt{n}]^*$	A1*
		(6)
(b)	$\mu = 45.2 + \frac{0.4 \times 12\sqrt{n}}{\sqrt{n}}$ or $\mu = 75.2 - \frac{2.1 \times 12\sqrt{n}}{\sqrt{n}}$	M1
	$\mu = 50$	A1
		(2)
(c)	$\frac{\sigma}{\sqrt{n}} = \left[\frac{12\sqrt{n}}{\sqrt{n}} \right] = 12$	B1
	$P(\bar{X}_n > 59) = P\left(Z > \frac{59 - '50'}{12}\right)$	M1
	$= 0.2266$ awrt 0.227	A1
		(3)
Notes		Total 11
(a)	Note	If they have not used the tables then maximum score M1A1M1A1M1A0* (5/6)
	M1	For standardising with μ and $\frac{\sigma}{\sqrt{n}}$ and setting = $\pm$ awrt 0.4
	A1	For a correct equation with compatible signs. Condone use of e.g. awrt 0.4
	M1	For standardising with μ and $\frac{\sigma}{\sqrt{n}}$ and setting = $\pm$ awrt 2.1
	A1	For a correct equation with compatible signs. Condone use of e.g. awrt 2.1
	M1	Dependent on at least one correct equation. For solving their 2 equations simultaneously to eliminate μ . Note they may eliminate $\frac{\sigma}{\sqrt{n}}$ first to find $\mu (= 50)$ and then substitute this back into one of their simultaneous equations which can score this mark.
	A1*	Answer is given so no incorrect working must be seen. They must have an intermediate stage where the equation without μ has the constant terms collected and does not have any fractions within fractions before proceeding to the given answer.

		e.g. $\frac{30}{\frac{\sigma}{\sqrt{n}}} = 2.5 \Rightarrow \frac{25\sigma}{\sqrt{n}} = 300 \Rightarrow \sigma = 12\sqrt{n}$
(b)	Note	If they do not have an equation in (a) involving the standard error $\frac{\sigma}{\sqrt{n}}$ then M0A0
	M1	For substitution of $\sigma = 12\sqrt{n}$ into an equation for μ . May be implied by 50
	A1	50
(c)	B1	Correct SE seen or may be implied e.g. N(...,144) or awrt 0.227
	M1	For standardising using their μ and standard error = 12. Can be implied by awrt 0.227
	A1	awrt 0.227 (correct answer provided no incorrect working seen in (c) scores 3/3)

Question Number	Scheme	Marks
5.		
(a)	$s_A^2 = \frac{1}{39}(790258 - 140.4^2 \times 40)$ $s_B^2 = \frac{1}{31}(581430 - 134.7^2 \times 32)$ $= 45.4256....$ awrt 45.4 $= 26.4232...$ awrt 26.4	M1 A1 A1 (3)
(b)	$H_0: \mu_A - \mu_B = 5$ $H_1: \mu_A - \mu_B > 5$ $s.e. = \sqrt{\frac{45.4256...}{40} + \frac{26.4232...}{32}}$ (= awrt 1.4) $z = \frac{\pm(140.4 - 134.7 - 5)}{s.e.}$ = awrt 0.50 c.v. = 1.6449 (Do not reject H_0) Insufficient evidence to support the greengrocer's belief . (Insufficient evidence that the difference in weight between type A oranges and type B oranges is over 5 grams).	B1 B1 M1 dM1 A1 B1 A1ft (7)
(c)	Large sample sizes so... Sample means are normally distributed (CLT) $s_A^2 = \sigma_A^2$ and $s_B^2 = \sigma_B^2$	B1 B1 (2)
	Notes	Total 12
(a)	M1 one correct expression A1 either awrt 45.4 or awrt 26.4 A1 both awrt 45.4 and awrt 26.4	
(b)	B1 Allow equivalent rearrangements. Must use μ B1 Allow equivalent rearrangements. Must use μ . For both hypotheses do not allow e.g. μ_1 and μ_2 unless each has been clearly defined M1 attempt at standard error (allow one slip) ft their (a) dM1 standardising with $(140.4 - 134.7 - 5)$ and their s.e. (dep on previous M1) A1 for 0.5 or awrt 0.50 B1 correct c.v. 1.6449 or better. Allow -1.6449 or better with use of $\mu_B - \mu_A$ Allow $p = \text{awrt } 0.309$ A1ft correct ft conclusion in context with either greengrocer's belief (oe) or difference in weights (oe) dependent on all B and M marks scored.	
(c)	B1 must comment on both sample means, $\bar{A}$ and $\bar{B}$ B1 must comment on both variances/standard deviations	

Question Number	Scheme	Marks
6(a)	$\bar{h} = 65.4$	B1
	$s^2 = \frac{214676 - 50 \times ("65.4")^2}{49}$	M1
	$= 16.693...$ awrt 16.7	A1
		(3)
(b)	$H_0: \mu_{do} = \mu_{do \text{ not}}$ $H_1: \mu_{do} < \mu_{do \text{ not}}$	B1
	$z = \pm \frac{"65.4" - 70.8}{\sqrt{\frac{"16.693..."}{50} + \frac{29.6}{40}}}$	M1M1
	$= \pm 5.21...$ awrt 5.21	A1
	CV 1.6449	B1
	Amala's <u>belief</u> is supported	A1 ft
		(6)
(c)	CLT enables you to assume that (the sampling distribution of the sample mean of) resting heart rate is normally distributed for <u>both</u> groups	B1
		(1)
(d)	Each population/sample is independent or each male is independent of the other males.	B1
	Assume the $\sigma_{do}^2 = s_{do}^2$ and $\sigma_{do \text{ not}}^2 = s_{do \text{ not}}^2$	B1
		(2)
Notes		Total 12
(a)	B1 65.4 only M1 Correct method to find s^2 using their $\bar{h}$ A1 awrt 16.7	
(b)	B1 Both hypotheses correct - must be clear which is exercise and which is not M1 For the denominator. Ft their 16.693... M1 Correct ft their 65.4 and 16.693... A1 awrt 5.21 allow $ z = 5.21...$ B1 $ CV = 1.6449$ or better ft their z value and CV if the hypotheses are the correct way round. Correct conclusion in context need belief. May be in words with heart and exercise e.g. resting heart rate is lower in men who exercise regularly A1	
(c)	B1 For the idea both groups normally distributed	
(d)	B1 For identifying the need for the groups or males to be independent. B1 Realising the $\sigma^2 = s^2$ Allow sample sizes big enough for CLT to hold	

Question Number	Scheme	Marks
3(a)	$H_0 : \mu = 328$ $H_1 : \mu \neq 328$ oe	B1
	Significance level is 5%	dB1
	(328 is within the interval therefore) no evidence to support that μ is not 328	B1
		(3)
(b)	$1.96 \times \frac{\sigma}{\sqrt{150}} = \frac{329.76 - 327.84}{2} (= 0.96)$ oe or e.g. $328.8 + 1.96 \times \frac{\sigma}{\sqrt{150}} = 329.76$ or $328.8 - 1.96 \times \frac{\sigma}{\sqrt{150}} = 327.84$ [$\sigma = 5.9987$]	M1 B1
	$328 - 2.3263 \times \frac{"5.9987"}{\sqrt{200}}$ oe or $328 + 2.3263 \times \frac{"5.9987"}{\sqrt{200}}$ oe	M1 B1
	$328 \pm 2.3263 \times \frac{"5.9987"}{\sqrt{200}}$ oe	dM1
	(327.0132... , 328.986...) (awrt 327, awrt 329)	A1
		(6)
(c)	$X \sim B(5, 0.98)$ or $Y \sim B(5, 0.02)$	M1
	$P(X \leq 4)$ or $P(Y = 1) = {}^5C_4 (0.98)^4 (0.02) + (0.98)^5$ oe	A1
	= 0.9961... awrt 0.996	A1
		(3)
	Notes	Total 12
(a)	Note that many are carrying out a full hypothesis test by finding the distribution of the sample mean. These three marks can still be scored. You do not need to check the calculations to score in (a)	
	B1 for both hypotheses correct in terms of μ (do not accept use of $\bar{x}$)	
	dB1 correct significance level given. Allow ($\alpha =$) 0.05 Do not allow if more than one given. This mark is dependent on a two tail test being indicated with their hypotheses condoning slips in notation e.g. use of $\bar{x}$ Do not allow $p = 0.05$ on its own.	
	B1 idea that $\mu = 328$ is supported. Allow e.g. "do not reject H_0 " provided $H_0 : \mu = 328$ but condone poor notation for their null hypothesis. Does not need to be in context.	
(b)	M1 forms an equation to find σ Allow $1.6 < z < 2$ may be seen in (a) or implied by awrt 6	
	B1 use of awrt 1.96 in the calculation may be seen in (a) or implied by their value for σ	
	M1 correct method to find one end of the confidence interval using their σ Allow $2.3 < z < 3.1$ condone use of 150 rather than 200 (may be implied by awrt 327 or awrt 329 if no incorrect working is seen)	
	B1 awrt 2.3263 seen	
	dM1 dependent on both the previous method marks being awarded. Fully correct method using their σ and using $n = 200$ to find both ends of the confidence interval. A correct confidence interval with no working – send to review	
	A1 (awrt 327, awrt 329) Condone missing brackets	
(c)	M1 use of Binomial e.g. $B(5, 0.98)$ or $B(5, 0.02)$ or 0.98^5 or ${}^5C_x \times 0.98^x \times 0.02^{5-x}$ (may be implied by awrt 0.0922 or awrt 0.904 within a calculation or awrt 0.996)	
	A1 for ${}^5C_4 \times 0.98^4 \times 0.02 + 0.98^5$ oe eg awrt 0.0922 + awrt 0.904 or implied by awrt 0.996	
	A1 awrt 0.996	

Question Number	Scheme		Marks
4 (a)	$2 \times \text{awrt } 2.5758 \times \text{SE} = 0.964 - 0.9$ or $\text{awrt } 2.5758 \times x = 0.032$		M1 B1
	$\Rightarrow \frac{0.964 - 0.9}{2 \times \text{awrt } 2.5758} [= 0.0124]^*$ or $x = \frac{0.032}{\text{awrt } 2.5758} [= 0.0124]^*$		A1*
			(3)
(b)	$[\bar{x}] = \frac{0.964 + 0.9}{2} [= 0.932]$ or $[\bar{x}] = 0.964 - '2.5758' \times 0.0124 [= \text{awrt } 0.932]$ or $[\bar{x}] = 0.9 + '2.5758' \times 0.0124 [= \text{awrt } 0.932]$		M1
	'0.932' $\pm$ 1.96×0.0124		M1 B1
	$(0.9076..., 0.9563...)$ awrt (0.908, 0.956)		A1
			(4)
(c)	$2 \times z \times 0.0124 = 0.04$		M1
	$z = 1.612...$ awrt 1.61		A1
	$P(Z > '1.61') = P(Z < -'1.61') = 1 - '0.9463'$		M1
	$= 0.0537$ (Calculator gives 0.05371....) awrt 0.0537		
	Confidence level = $[100 \times](1 - 2 \times '0.0537')$ or $[100 \times]('0.9463' \times 2 - 1)$		M1
	$= 89.26$ awrt 89.3		A1
			(5)
Notes			Total 12
(a)	M1	For $2 \times z \text{ value} \times \text{SE} = 0.964 - 0.9$ oe or $z \text{ value} \times x = 0.032$ oe where $2 < z < 3$ May be implied by $\frac{0.964 - 0.9}{2 \times \text{awrt } 2.5758}$ or $\frac{0.032}{\text{awrt } 2.5758}$	
	B1	awrt 2.5758	
	A1*	Answer is given so no incorrect working must be seen. Must be at least one line of correct working between M1 and the final answer. Must use awrt 2.5758 May be implied by awrt 0.01242...	
(b)	M1	Accept awrt 0.932 to imply a correct method. If using a z value, then this must be awrt 2.5758 or consistent with the z value used in part (a)	
	M1	For $\bar{x} \pm z \text{ value} \times 0.0124$ ft their $\bar{x}$ and where $1.5 < z < 2$	
	B1	awrt 1.96	
	A1	for (awrt 0.908, awrt 0.956) Allow awrt $0.908 < \mu < \text{awrt } 0.956$	
(c)	M1	For $2 \times z \times 0.0124 = 0.04$ oe May be implied by awrt 1.61	
	A1	For $z = \text{awrt } 1.61$	
	M1	For awrt 0.946 or awrt 0.947 or awrt 0.053 or awrt 0.054 NB awrt 0.946 or or awrt 0.947 or awrt 0.053 or awrt 0.054 scores M1A1M1	
	M1	For $[100 \times](1 - 2 \times '0.0537')$ or $[100 \times]('0.9463' \times 2 - 1)$ ft their $P(Z > '1.61')$ (May be implied by 89.26 or awrt 89.2 or awrt 89.3 or 0.8926 or awrt 0.892 or awrt 0.893)	
	A1	For awrt 89.3 NB An answer of 89.2 or 89 can score M1A1M1M1A0	

Question Number	Scheme		Marks
5 (a)	$\bar{x} \pm 1.6449 \times \frac{5}{\sqrt{10}}$		M1 B1
	$\bar{x} \pm 2.60 \Rightarrow (\bar{x} - 2.60, \bar{x} + 2.60)^*$		A1*
			(3)
(b)	$\bar{y} \pm 1.96 \times \frac{3}{\sqrt{20}}$		M1 B1
	$\bar{y} \pm 1.31 \Rightarrow (\bar{y} - 1.31, \bar{y} + 1.31)$		A1
			(3)
(c)(i)	$\bar{X} - \bar{Y} \sim N\left(\mu - \mu, \frac{5^2}{10} + \frac{3^2}{20}\right) \Rightarrow \bar{X} - \bar{Y} \sim N(0, 2.95)$		M1 A1
(ii)	Do not overlap when either		
	$\bar{x} - 2.60 > \bar{y} + '1.31'$ or $\bar{x} + 2.60 < \bar{y} - '1.31'$		M1
	$\bar{x} - \bar{y} > 3.91$ or $\bar{x} - \bar{y} < -3.91$		A1ft
	$2 \times P(\bar{X} - \bar{Y} > 3.91) = 2 \times P\left(Z > \frac{'3.91' - '0'}{\sqrt{2.95}}\right) = [2 \times P(Z > 2.276...)]$		M1 M1
	$[2 \times 0.0113] = 0.0226$ (calculator gives $[2 \times 0.0114...] = 0.0228$)		A1
			(7)
	Notes		Total 3
(a)	M1	For use of $\bar{x} \pm z$ value $\times \frac{5}{\sqrt{10}}$	
	B1	For use of $z = 1.6449$ or better	
	A1*	Answer is given so no incorrect working should be seen (condone use of 1.645)	
(b)	M1	For use of $\bar{y} \pm z$ value $\times \frac{3}{\sqrt{20}}$	
	B1	For use of $z = 1.96$ or better	
	A1	For $(\bar{y} - \text{awrt}1.31, \bar{y} + \text{awrt}1.31)$ Allow 1.315	
(c)(i)	M1	For a correct method to find the variance (May be seen in a standardisation expression)	
	A1	For $N(0, 2.95)$ (May be seen in a standardisation expression) Allow $N\left(0, \frac{5^2}{10} + \frac{3^2}{20}\right)$ oe	
(ii)	M1	For $\bar{x} - 2.60 > \bar{y} + 1.31$ oe or $\bar{x} + 2.60 > \bar{y} - 1.31$ oe ft part (b)	
	A1ft	For $\bar{x} - \bar{y} > '3.91'$ or $\bar{x} - \bar{y} < -'3.91'$ ft part (b)	
	M1	For multiplying by 2 (may be seen at any stage of their working)	
	M1	For standardising ft their 3.91, their mean and their standard deviation (Do not allow use of 2.6 or 1.31 as their 3.91)	
	A1	For answers in the range awrt 0.0226 – awrt 0.0228	

Question Number	Scheme	Marks
4. (a)	Use of $\bar{x} \pm z \times \frac{18}{\sqrt{25}}$; $z = 2.3263$ (or better) $= (44.0253..., 60.7746...)$ awrt (44.0, 60.8)	M1;B1 A1, A1 (4)
(b)	$H_0: \mu_A = \mu_B$ $H_1: \mu_B > \mu_A$ $z = (\pm) \frac{57.8 - 52.4}{18\sqrt{\frac{1}{25} + \frac{1}{30}}}$ $= (\pm) 1.1078...$ awrt (±) 1.11 5% one-tail critical value is 1.6449 (or p -value = 0.13396... i.e. awrt 0.134) (not sig') so insufficient evidence (in these data) to support newspaper's claim	B1 M1dM1 A1 B1 A1 (6)
(c)	Require $\frac{\bar{x} - \mu}{\frac{18}{\sqrt{n}}} > z$ where $z = -1.6449$ (o.e.) $\mu < 52.4 + 1.64(49) \times \frac{18}{5}$ or $\mu < 57.8 + 1.64(49) \times \frac{18}{\sqrt{30}}$ i.e. $\mu < 58.3216...$ and $\mu < 63.2056...$ So $\mu = \mathbf{58.3}$	M1 A1 M1 A1 (4)
[14 marks]		
Notes		
(a)	M1 for use of correct expression with 18, 25 and $1 < z < 3$ (Ignore $\bar{x}$ for this mark) B1 for $z = 2.3263$ or better (calc: 2.32634787...) 1 st A1 for awrt 44.0 (ans only of 44.02...or awrt 44.03 scores M1B1 implied) 2 nd A1 for awrt 60.8 (ans only of 60.77... or awrt 60.77 scores M1B1 implied)	
(b)	1 st B1 for both hypotheses in terms of μ s (If using μ_1 etc they must define which is which) 1 st M1 for a correct denominator (18 needn't be outside square root) [4.87(44...)] 2 nd dM1 for a correct expression for test statistic 1 st A1 for awrt (±) 1.11 2 nd B1 for critical value of 1.6449 or better (If B0 in (a) for 2.33 allow 1.64 or 1.645 here) [Allow p -value of awrt 0.134 and condone awrt 0.866 if compared with 0.95] 2 nd A1 Correct contextual conclusion, ft comparing their "1.11" with 1.64 (or their cv) but must be not significant and mention "claim" or "score in town A" and "score in town B"	
(c)	1 st M1 for a correct starting <u>inequality</u> with any z such that $ z > 1$ (Allow $\geq$) 1 st A1 for either correct <u>inequality</u> for μ , allow $z = 1.64$ or better 2 nd M1 for both cases of $\bar{x} \pm z \frac{18}{\sqrt{n}}$ ($z > 1$) can allow "=" or inequality, may be in CI 2 nd A1 (dep on both Ms) for sight of both awrt 58.3 and awrt 63.2 and selecting awrt 58.3	

Question Number	Scheme		Marks
3 (a)	$\left[\frac{26.624 + 28.976}{2} \right] = 27.8$		B1
			(1)
(b)	$28.976 - 26.624 = 2 \times 1.96 \times \frac{\sigma}{\sqrt{25}}$ or $26.624 = '27.8' - 1.96 \times \frac{\sigma}{\sqrt{25}}$		M1 B1
	or $28.976 = '27.8' + 1.96 \times \frac{\sigma}{\sqrt{25}}$		
	$\sigma = 3^*$		A1* cso
			(3)
(c)	$2 \times z \times \frac{3}{\sqrt{25}} = 2.1$ So $z = 1.75$		M1 A1
	$P(Z > '1.75') = P(Z < -'1.75') = 1 - '0.9599' = '0.0401'$		M1 A1ft
	Confidence level = $100 \times (1 - 2 \times '0.0401') = 91.98\%$		M1 A1
			(6)
(d)	$2 \times 1.96 \times \frac{3}{\sqrt{n}} < 1.5$		M1
	$\sqrt{n} > \frac{6 \times 1.96}{1.5}$		dM1
	$\sqrt{n} > \text{awrt } 7.84$ So $n = 62$		A1 A1
			(4)
	Notes		Total 14
(a)	B1	for 27.8	
(b)	M1	for $28.976 - 26.624 = 2 \times z \text{ value} \times \frac{\sigma}{\sqrt{25}}$ or $26.624 = '27.8' - z \text{ value} \times \frac{\sigma}{\sqrt{25}}$ or $28.976 = '27.8' + z \text{ value} \times \frac{\sigma}{\sqrt{25}}$ where $1.5 < z < 2.4$	
	B1	awrt 1.96	
	A1* cso	answer is given so no incorrect working must be seen	
(c)	M1	for $2 \times z \times \frac{3}{\sqrt{25}} = 2.1$	
	A1	for $z = 1.75$	
	M1	for $1 - p$, where p is a probability	
	A1ft	for 0.0401 or ft their z value (Allow 0.04)	
	M1	for $100 \times (1 - 2 \times 0.0401)$ ft their $P(Z < -1.75)$	
	A1	awrt 92.0 (allow 92)	
(d)	M1	for $2 \times z \text{ value} \times \frac{3}{\sqrt{n}} < 1.5$ oe z value must either be correct or consistent with part (b) Allow $\leq$ or $=$ Condone $>$ or $\geq$	
	dM1	Dependent on previous M mark. Correct rearrangement to get $\sqrt{n} > \dots$ or $n > \dots$ oe Allow $\geq$ or $=$ Condone $<$ or $\leq$	
	A1	awrt 7.84 may be implied by awrt 61.5	
	A1	for $n = 62$	

Question Number	Scheme	Marks
5(a)	$H_0: \mu_p - \mu_f = 1$ oe	B1
	$H_1: \mu_p - \mu_f > 1$ oe	B1
	$s.e. = \sqrt{\frac{9}{605} + \frac{4}{45}} = [\sqrt{0.10376...}] = [0.322...]$	M1
	$z = \pm \frac{7.0 - 5.6 - 1}{\sqrt{\frac{9}{605} + \frac{4}{45}}}$	dM1
	$= 1.24175...$ awrt 1.24	A1
	CV 5% one tailed $= \pm 1.6449$ (see notes)	B1
	Not significant, do not reject H_0	dM1
	Insufficient evidence that full-time staff are more than one minute faster than part-time staff or manager's claim is not supported	A1ft
		(8)
(b)	Assume both samples are normal or both large enough for CLT oe	B1
	Assume $s^2 = \sigma^2$ for both samples	B1
	Assume individual results are independent	
		(2)
(c)	$\bar{a} = \frac{45 \times 7 + 8}{46} [= 7.0217...]$	M1
	$\sum a^2 = 44 \times 4 + 45 \times 7^2 + 8^2 [= 2445]$	M1
	$s^2 = \frac{"2445" - 46 \times "7.0217..."^2}{45}$	M1
	$= 3.93285...$ awrt 3.93	A1
		(4)
	Notes	Total 14
(a)	B1 H_0 correct oe e.g. $H_0: \mu_f - \mu_p = -1$ Must be in terms of μ Use of $\bar{t}$ is B0. May use other letters to p and f but must be defined	
	B1 H_1 correct oe e.g. $H_1: \mu_f - \mu_p < -1$ SC B0B1 for $H_0: \mu_f - \mu_p = 1$ and $H_1: \mu_f - \mu_p > 1$ Must be in terms of μ Use of $\bar{t}$ is B0 May use other letters to p and f but must be defined	
	M1 Correct method to find s.e. may be seen within formula to standardise (implied by awrt ± 1.24)	
	dM1 dep on previous M being awarded. Correct method to find z value (implied by awrt ± 1.24)	
	A1 awrt ± 1.24	
	B1 awrt ± 1.6449 or p -value awrt $0.107 > 0.05$ oe	
	dM1 all previous method marks awarded for a correct conclusion ft their z value and CV	
	A1ft dependent on all previous method marks but independent of hypotheses. A correct statement in context with the words in bold oe which does not reject H_0 . ft their z value and CV Note B0B0M1dM1A1B0dM1A1ft is possible	
	Candidates who incorrectly test $H_0: \mu_f - \mu_p = 1$ and $H_1: \mu_f - \mu_p > 1$ will score maximum B0B1M1dM0A0B1dM0A0	
(b)	B1 One correct assumption. Accept e.g. (times taken by) employees chosen follow a normal distribution, times taken follow a normal distribution, employees are selected	

	independently. Samples are independent is B0
	B1 2 correct assumptions. Need to see reference to both for each assumption but condone if written as one sentence or statement e.g. both samples are normally distributed and $s^2 = \sigma^2$
(c)	M1 correct method to find $\bar{a}$ may be implied by awrt 7.02 or may be seen within a calculation e.g. $(s^2 =) \frac{2445 - \frac{(45 \times 7 + 8)^2}{46}}{45}$
	M1 correct method to find $\sum a^2$ may see e.g. $2381 + 8^2$ which may be embedded within their calculation to find the sample variance. Sight of 2445 implies this mark
	M1 correct method to find s^2 ft their $\bar{a}$ (which cannot be 315) and $\sum a^2$ (which cannot be 2381)
	A1 awrt 3.93

Question Number	Scheme		Marks
5 (a)(i)	$\left[\bar{x} = \frac{3610}{50} \Rightarrow \bar{x} = 72.2 \quad s_x^2 = \frac{260955.6 - 50(72.2)^2}{50 - 1} = 6.4 \right]$		B1; M1 A1
5(a)(ii)	$\left[\bar{y} = \frac{2585}{50} \Rightarrow \bar{y} = 51.7 \quad s_y^2 = \frac{133757.2 - 50(51.7)^2}{50 - 1} = 2.3 \right]$		B1 A1
			(5)
(b)	$H_0 : \mu_x - \mu_y = 20$ $H_1 : \mu_x - \mu_y > 20$		B1
	$z = \frac{'72.2' - '51.7' - 20}{\sqrt{\frac{'6.4'}{50} + \frac{'2.3'}{50}}}$		M1 M1
	$= 1.1986...$ awrt 1.20		A1
	One tailed c.v. $Z = 1.6449$ or CR: $Z \dots 1.6449$		B1
	Not in CR/Not significant/Do not reject H_0		M1
	No significant evidence to support Tammy's belief		A1
			(7)
(c)	Since the sample is large the CLT applies.		M1
	No need to assume (the weights) are normally distributed.		A1
			(2)
(d)	Assumed that $s^2 = \sigma^2$		B1
			(1)
Notes			Total 15
5 (a)(i)	B1	$\bar{x} = 72.2$	
	M1	A correct method for finding an unbiased estimate of the variance e.g. $\frac{\sum x^2 - n(\bar{x})^2}{n - 1}$ (May be seen in (i) or (ii))	
	A1	6.4	
5(a)(ii)	B1	$\bar{y} = 51.7$	
	A1	2.3	
(b)	B1	Both hypotheses correct. Allow equivalent hypotheses. Must be in terms of μ	
	M1	For correct standard error. Follow through their values from (a)	
	M1	An attempt at $\frac{a - b - 20}{\sqrt{\frac{c}{50} + \frac{d}{50}}}$ with at least 2 of a, b, c or d correct. Allow $\pm$	
	A1	awrt 1.20 Allow 1.2 if no incorrect working shown	
	B1	1.6449 or better (seen)	
	M1	A correct statement – need not be contextual but do not allow contradicting non contextual comments.	
	A1	A correct contextual statement. Allow the difference in mean weights is not greater than 20 kg	
(c)	M1	A suitable comment that mentions large and CLT	
	A1	A correct answer, context not required.	
(d)	B1	for the assumption that sample variance = population variance	

Question Number	Scheme	Marks
6. (a)	$[\bar{x} = 49.8]$ $2 \times 1.96 \left(\frac{\sigma}{\sqrt{8}} \right) = 53.88 - 45.72 = 8.16$ $2 \times 2.5758 \left(\frac{\sigma}{\sqrt{8}} \right) = \frac{8.16 \times 2.5758}{1.96} = 10.7238...$ $99\% \text{ CI} = 49.8 \pm \frac{10.7238}{2}$	M1 B1 M1 M1 A1 (5)
(b)	$\hat{\mu} = \bar{x} = \frac{91.2}{8} = 11.4$ $\hat{\sigma}^2 = s^2 = \frac{1145.16 - 8 \times "11.4^2"}{7} = 15.06857...$	B1 M1 A1 (3)
(c)	Combined $\Sigma x = 10.8 \times 24 + 91.2 = 350.4$ Combined $\Sigma x^2 = 1145.16 + 23 \times 17.64 + 24 \times 10.8^2 = 4350.24$ $\text{Combined } s^2 = \frac{"4350.24" - 32 \times \left(\frac{"350.4"}{32} \right)^2}{31} = 16.56$ $\frac{s}{\sqrt{n}} = \frac{\sqrt{16.56}}{\sqrt{32}} = 0.719374...$	M1A1 M1 A1 M1 A1 (7) [15]
Notes		
(a)	1st M1 use of $2z \frac{\sigma}{\sqrt{n}}$ or $z \frac{\sigma}{\sqrt{n}}$ with $1.5 < z < 2$. Allow σ_m for $\frac{\sigma}{\sqrt{n}}$ B1 1.96 (or better) and 2.5758 (or better) 2nd M1 attempt to find width or semi-width of 99% CI with $z > 2$ Allow $\sigma = \frac{4.08 \times \sqrt{8}}{1.96} [= 5.887...]$ 3rd M1 Use of $49.8 \pm$ awrt 5.36 or $49.8 \pm 2.5758 \left(\frac{"5.887..."}{\sqrt{8}} \right)$ If σ is incorrect then working must be shown. A1 correct interval with (awrt 44.4, awrt 55.2) Correct answer from less accurate z-values scores M1B0M1M1A1	
(b)	B1 11.4 cao M1 full attempt at s^2 ft their $\bar{x}$ A1 awrt 15.1	
(c)	M1 for correct combined sum (may be implied by combined mean of 10.95) 2nd M1 for attempt at combined sum of squares $1145.16 + (n-1) \times 17.64 + n \times 10.8^2$ (allow 1 error) 1st A1 fully correct expression or awrt 4350 3rd M1 using their values in a complete expression for combined s^2 oe 2nd A1 $s^2 = 16.56$ or $s =$ awrt 4.07 (either of these implies M1M1A1M1A1) 4th M1 use of $\frac{s}{\sqrt{n}}$ with combined values 3rd A1 awrt 0.719	

Question Number	Scheme		Marks
1.	$H_0: \mu = 30 \quad H_1: \mu < 30$		B1
	$z = \frac{29.5 - 30}{\frac{2.5}{\sqrt{80}}}$		M1
	$z = -1.7888\dots$	awrt -1.79	A1
	$-1.7888 < -1.6449$		B1
	Reject H_0 or significant result or in the critical region		
	There is evidence to support the <u>manager's</u> claim.		A1
			(5)
	Notes		Total 5
	B1	Both hypotheses correct in terms of μ	
	M1	for attempting test statistic, allow $\pm$, Condone $\sqrt{\frac{2.5}{80}}$	
	A1	awrt -1.79 allow $ z = 1.7888\dots$ Allow p value of 0.0367 or awrt 0.0368 or $CR \leq 29.54$	
	B1	$ CV = 1.6449$ or better (Ignore any comparisons) Allow $CR \leq 29.54$ SC If p value of 0.0367 or awrt 0.0368 award B1 if 2 nd A1 is awarded	
	A1	For correct conclusion. Allow the manager's claim in words if it includes screws and less (oe)	

Question Number	Scheme	Notes	Marks																								
2	H_0 : Diet and health are independent (or not associated). H_1 : Diet and health are not independent (or associated).	“diet” and “health” mentioned at least once. Use of correlation is B0.	B1																								
	<table border="1"> <thead> <tr> <th>EXPECTED</th><th>Good health</th><th>Poor health</th><th>Totals</th></tr> </thead> <tbody> <tr> <td>Good diet</td><td>83.19</td><td>10.81</td><td>94</td></tr> <tr> <td>Poor diet</td><td>93.81</td><td>12.19</td><td>106</td></tr> <tr> <td>Totals</td><td>177</td><td>23</td><td>200</td></tr> </tbody> </table>	EXPECTED	Good health	Poor health	Totals	Good diet	83.19	10.81	94	Poor diet	93.81	12.19	106	Totals	177	23	200	Attempt $\frac{RT \times CT}{GT}$ with at least one correct to 1dp; all correct to 1dp.	M1; A1								
EXPECTED	Good health	Poor health	Totals																								
Good diet	83.19	10.81	94																								
Poor diet	93.81	12.19	106																								
Totals	177	23	200																								
	<table border="1"> <thead> <tr> <th>Observed</th><th>Expected</th><th>$\frac{(O-E)^2}{E}$</th><th>$\frac{O^2}{E}$</th></tr> </thead> <tbody> <tr> <td>86</td><td>83.19</td><td>0.095</td><td>88.905</td></tr> <tr> <td>8</td><td>10.81</td><td>0.730</td><td>5.920</td></tr> <tr> <td>91</td><td>93.81</td><td>0.084</td><td>88.274</td></tr> <tr> <td>15</td><td>12.19</td><td>0.648</td><td>18.458</td></tr> <tr> <td colspan="2">Totals</td><td>1.557</td><td>201.557</td></tr> </tbody> </table>	Observed	Expected	$\frac{(O-E)^2}{E}$	$\frac{O^2}{E}$	86	83.19	0.095	88.905	8	10.81	0.730	5.920	91	93.81	0.084	88.274	15	12.19	0.648	18.458	Totals		1.557	201.557	Attempt at $\frac{(O-E)^2}{E}$ or $\frac{O^2}{E}$ with their values with 2 correct or 2 correct f.t. Allow 2sf for this mark.	M1
Observed	Expected	$\frac{(O-E)^2}{E}$	$\frac{O^2}{E}$																								
86	83.19	0.095	88.905																								
8	10.81	0.730	5.920																								
91	93.81	0.084	88.274																								
15	12.19	0.648	18.458																								
Totals		1.557	201.557																								
	$\sum \frac{(O-E)^2}{E}$ or $\sum \frac{O^2}{E} - 200 = 1.557$	awrt 1.6	A1																								
	$\nu = (2-1)(2-1) = 1$ $\chi^2_1(5\%) = 3.841$	1 (may be implied) 3.841 NB: may see $\chi^2_3(5\%) = 7.815$ for f.t. NB: p -value 0.212 but scores B0B0 on its own	B1 B1f.t.																								
	$\chi^2 = 1.557$ does not lie in cr so insufficient evidence to reject H_0	For correct non-contextual statement linking their test statistic and their cv.	M1																								
	E.g. Diet and health are independent or There is no association between diet and health or The doctor's belief is not supported by this data.	Dependent on a cv of 3.841 and 3 rd M1 Correct conclusion in context with “diet” and “health” or “doctor”. Condone “connection” or “relationship” but not “correlation”.	A1ft																								
			(9)																								
			Total 9																								

Question Number	Scheme	Marks
5. (a)	$H_0 : \mu_H - \mu_M = 15$ $H_1 : \mu_H - \mu_M > 15$ $z = \frac{56.3 - 39.8 - 15}{\sqrt{\frac{27.2}{38} + \frac{18.5}{45}}}$ $= 1.4130\dots$ $\text{CV} = 1.6449 \text{ (or better) or } p = \text{awrt } 0.0788$ Do not reject H_0 /Not significant There is not sufficient evidence to support the <u>professor's claim</u> /there is not sufficient evidence to suggest that undergraduates studying <u>History type</u> more than 15 words/minute faster than undergraduates studying <u>Maths</u> .	B1 M1 M1 awrt 1.41 A1 B1 M1 A1 (7)
(b)	$s^2 \approx \sigma^2$ for both History and Maths Assume sample sizes are large enough so that CLT applies or $\bar{X}$ is normally distributed for both	B1 B1 (2) [9]
Notes		
(a)	1 st B1 both hypotheses correct in terms of μ_H and μ_M . Allow equivalent rearrangements. Allow other letters as long it is clear which is History and which is Maths Must be attached to H_0 and H_1 1 st M1 for $z = \frac{a - b - 15}{\sqrt{\frac{c}{38} + \frac{d}{45}}}$ with at least 2 of a, b, c or d correct (allow $\pm$) 2 nd M1 for $z = \frac{56.3 - 39.8 - 15}{\sqrt{\frac{27.2}{38} + \frac{18.5}{45}}}$ (allow $\pm$) 1 st A1 awrt 1.41 2 nd B1 for CV = ± 1.6449 and compatible sign with their test statistic (allow $p = \text{awrt } 0.0788$) 3 rd M1 correct statement consistent with their test statistic and CV (no contradictory non-contextual comments) May be implied by correct contextual comment. 2 nd A1 contextual conclusion that is consistent with their test statistic and their CV. Must mention professor's claim or History, Maths and typing (oe).	
(b)	1 st B1 must mention both. Allow $s \approx \sigma$ for both History and Maths 2 nd B1 either correct assumption	

Question	Scheme		Marks
5 (a)	$H_0: \mu_A = \mu_B$ $H_1: \mu_A > \mu_B$ oe		B1
	$se = \sqrt{\frac{17.8^2}{50} + \frac{18.4^2}{40}}$		M1
	$z = \pm \frac{1377 - 1368}{\sqrt{\frac{17.8^2}{50} + \frac{18.4^2}{40}}}$		M1
	$= \pm 2.339...$	awrt ± 2.34	A1
	One tailed c.v. $ Z = 2.3263$ or CR: $Z \leq -2.3263$ or $Z \geq 2.3263$		B1
	In CR/Significant/Reject H_0		dM1
	Sufficient evidence to support that the mean yield from plants using fertiliser A is greater than the mean yield from plants using fertiliser B		A1ft
			(7)
	finding the CI can get B1M1M1A0B1M1A1 unless test statistic given		
	award M1 for $z = \pm \frac{D}{\sqrt{\frac{17.8^2}{50} + \frac{18.4^2}{40}}}$ dep on first M1 where $2.3 \leq z \leq 2.4$		
(b)	Expected profit per plant		
	A: $3 \times 1.377 - \frac{75}{50}$ B: $3 \times 1.368 - \frac{50}{40}$		M1
	A: £2.63(1) B: £2.85(4)		A1
	Claire should use fertiliser B		dA1 (3)
Notes			Total 10
(a)	B1	Both hypotheses correct. Allow equivalent hypotheses. Must be in terms of μ If A and B not used the letter must be defined	
	M1	For a correct attempt to find the se or se^2 Condone slip in sample sizes May be implied by $se = \text{awrt } 3.85$ or $se^2 = \text{awrt } 14.8$. Allow for a p -value of 0.0096 or awrt 0.0097	
	M1	For an attempt to find z value. Allow slip in sample sizes and/or use of 17.8 and 18.4 rather than 17.8^2 and 18.4^2 Allow for a p -value of 0.0096 or awrt 0.0097	
	A1	awrt $= \pm 2.34$ Allow for a p -value of 0.0096 or awrt 0.0097	
	B1	± 2.3263 or better seen (Calculator gives 2.3263479...) must be compatible with their test statistic	
	dM1	dep on previous dM1 awarded, ft their test statistic and CV only. A correct statement compatible with their test statistic and CV only – need not be contextual but do not allow contradicting non contextual comments.	
	A1ft	ft their z value and CR only. A correct contextual statement compatible with their test statistic and CV with context of yield (at least once) and A and B	
		NB id they give a p -value of awrt 0.0096/7 they could get B1M1dM1A1B0dM1A1	
(b)	M1	A correct method to find the profit per n plants or m kg for either fertiliser A or fertiliser B $n(3 \times 1.377 - \frac{75}{50})$ or $n(3 \times 1.368 - \frac{50}{40})$ or $m(3 - \frac{75}{50} \times 1.377)$ or $m(3 - \frac{50}{40 \times 1.368})$ where n and $m \neq 0$ Implied by one correct value for A or B	
	A1	must have 2 values which can be compared, ie using same n or m . Profit per n plant £2.63(1) n and £2.85(4) n or profit per m kg awrt £1.91 m and awrt £2.09 m (2dp) or cost per m kg awrt £1.09 m and awrt £0.91 m or number plants per £y awrt 0.38...y and awrt 0.35...y Useful numbers ($n = 50$ gives profit 131.55, 142.7) or ($n = 40$ gives profits 105.24 and 114.16) gain M1A1	
	dA1	dependent on 1 st A1 being awarded. For a correct statement.	

Question Number	Scheme		Marks
2 (a)	$\bar{x} = \left[\frac{7690}{100} \right] = 76.9$		B1
	$s_x^2 = \frac{669.24}{99} = 6.76$		M1 A1
			(3)
(b)	$H_0 : \mu_x = \mu_y \quad H_1 : \mu_x \neq \mu_y$		B1
	$Z = \frac{"76.9" - 75.9}{\sqrt{\frac{"6.76"}{100} + \frac{2.2^2}{80}}} = 2.793...$	awrt ± 2.79	M1 M1 A1
	2 tailed critical value $z = \pm 2.5758$		B1
	Reject H_0 /Significant/In the critical region		M1
	There is sufficient evidence to suggest that the mean <u>water temperature</u> after 4 hours for brand <u>A</u> is different to brand <u>B</u>		A1ft
			(7)
(c)	(It is reasonable) since both samples are (reasonably) large		B1
			(1)
Notes			Total 11
(a)	B1	for 76.9	
	M1	for use of $\frac{1}{n-1} \sum (x - \bar{x})^2$ oe	
	A1	for 6.76	
(b)	B1	for both hypotheses correct. Must be attached to H_0 and H_1 . Allow equivalent hypotheses. Must be in terms of μ . Allow any letter for the subscripts	
	M1	for a correct method to find the standard error. Follow through their values from (a)	
	M1	an attempt at $\pm \frac{a-b}{\sqrt{\frac{c}{100} + \frac{d^2}{80}}}$ with at least 3 of a, b, c or d correct.	
	A1	awrt ± 2.79	
	B1	$z = \text{awrt} \pm 2.5758$ seen (Allow $z = \text{awrt} \pm 2.3263$ if a one tailed test is used)	
	M1	a correct statement consistent with their CV and Z value – need not be contextual but do not allow contradicting non contextual comments. This may be implied by a correct contextual conclusion.	
	A1ft	This mark is dependent on the 2 nd M mark being awarded. A correct contextual statement ft their CV and their Z value	
(c)	B1	a correct explanation, which makes reference to both samples. e.g. Do not allow the sample is large enough	

Question	Scheme		Marks
1 (a)(i)	Method 1	Method 2	
	$[\bar{y}] = \frac{847}{100} [= 8.47]$	$847 + 100 \times 1000 [= 100847]$	M1
	So $\bar{x} = 1000 + \frac{847}{100} = 1008.47$ *	$\bar{x} = \frac{847 + 1000 \times 100}{100} = 1008.47^*$	A1*
(ii)	$[s_x^2 = s_y^2 =] \frac{13510.09 - 100 \times "8.47"'^2}{99}$	$[s_x^2 =] \frac{101707510.1 - \frac{"100847"'^2}{100}}{99}$	M1
	$= 64$		A1
			(4)
(b)	$H_0 : \mu_x = 1010$ $H_1 : \mu_x \neq 1010$		B1
			(1)
(c)	$\frac{\bar{X} - 1010}{\frac{"8"}{\sqrt{100}}} = -1.96$ oe $\frac{\bar{X} - 1010}{\frac{"8"}{\sqrt{100}}} = 1.96$ oe		M1 B1
	$\bar{X} = 1008.432$ $\bar{X} = 1011.568$ awrt 1008 and 1012(or 1011)		A1
	$\bar{X} \leq "1008.432"$ $\bar{X} \geq "1011.568"$		A1ft
			(4)
(d)	1008.47 is not in the critical region		M1
	The machine does not need to be stopped /reset		A1ft
			(2)
(e)	It is reasonable since the sample size is (reasonably) large		B1
			(1)
Notes			Total 12
(a)(i)	M1	For 8.47 or $\frac{847}{100}$ or $847 + 100 \times 1000$ or $847 = \sum x - 100 \times 1000$ or 100847 seen	
	A1*	cso correct solution including $\bar{x} = ..$ and $... = 1008.47$ allow alt notation for $\bar{x}$ but must refer to x not y and must not be just x eg $E(X)$, μ_x , mean of x	
(ii)	M1	For a correct expression ft their 100847 Allow for answer of 1064	
	A1	Cao do not ISW Allow 64.00...	
(b)	B1	Both hypotheses correct. Must be in terms of μ . (Allow $H_0 : \mu_y = 10$ $H_1 : \mu_y \neq 10$)	
		Mark (c) and (d) together	
(c)	M1	For $\pm$ standardisation with 1010 and their sd Allow equivalent eg $1010 \pm n \times "8" / \sqrt{100}$	
	B1	For c.v. = ± 1.96 or better seen (Calculator gives 1.95996...) Condone 1.6449 or better if they have a one tail hypotheses in (b)	
	A1	For both limits 1008 or better and 1012 or better seen. (condone 1011 from correct working)	
	A1	For selecting the correct region ft their figures(not z value). Allow use of $<$ and $>$ also allow other letters(condone μ) Allow other notation eg $[1012, \infty]$, $(\infty, 1008]$ allow $[1012, \infty]$, $[\infty, 1008]$	
(d)	M1	ft their CR if the final A mark in part (c) is awarded. For a correct comment compatible with their CR. Must refer to 1008.47 (allow mean of x) is in or out of their CR	
	A1ft	Allow writing in the form $"1008.432" < 1008.47 < "1011.568"$ etc but if in middle it must have both ends. If no clear CR it is M0A0	
		dep on M1 awarded. Correct conclusion consistent with comparing 1008.47 with their CR(allow interval/ range etc). If it is in the CR they must say it needs to be reset/stopped. If it is not in the CR it must say it does not need to be stopped/reset. (allow equivalent wording)	

	SC	If the CR in (c) is of the form " $1008.432 < \bar{X} < 1011.568$ " oe (not z values) then award M0A1 for concluding the machine does not need to be stopped/reset.
(e)	B1	Any suitable comment about the sample being large eg n is large



Question Number	Scheme		Marks
2 (a)	$\frac{[0 \times 5] + 1 \times 38 + 2 \times 32 + 3 \times 17 + 4 \times 7 + 5 \times 1}{100} [= 1.86]^*$		B1*
			(1)
(b)	$[r = 1.203]$ because total expected frequency must equal 100		B1
			(1)
(c)	[The manager needed to do this] to ensure that [all] expected frequencies were greater than 5		B1
			(1)
(d)	H_0 : Poisson (distribution) is [a] suitable/ sensible (model)		B1
	H_1 : Poisson (distribution) is not [a] suitable/ sensible (model)		
	$\nu = [5 - 1 - 1] = 3$		B1
	$c_3^2(0.01) = 11.345 \Rightarrow \text{CR: } X^2 \dots 11.345$		M1
	[Lies in the CR/Reject H_0] Sufficient evidence to say that Poisson is not a suitable model		A1ft
		(4)	
Notes			Total 7
(a)	B1*	For a correct method to show the mean is 1.86 (Ignore use of 6×0) Allow $\frac{[0] + 38 + 64 + 51 + 28 + 5}{100}$	
(b)	B1	A correct explanation referring to the fact that total/sum expected frequency/ E_i must equal total observed frequency e.g. $100 - (15.567 + 28.955 + 26.928 + 16.696 + 7.763 + 2.888) = r$	
(c)	B1	A correct explanation referring to the fact that [all] E_i / expected frequencies/values need to be greater than 5 e.g because expected 5 customers and [expected] 6 or more customers are both less than 5 Allow $2.88 < 5$ and $1.203/r < 5$ or $4.091 < 5$	
(d)	B1	Both hypotheses correct. Must mention Poisson/Po at least once.	
	B1	$\nu = 3$ This mark can be implied by a correct critical value of 11.345 if no DoF given	
	M1	For 11.345 or ft their degrees of freedom $\left[c_4^2(0.01) = 13.277 \right]$	
	A1ft	A correct conclusion based on their χ^2 critical value. Must mention Poisson	

Question Number	Scheme		Marks
3 (a)	$\left[p = \frac{118}{40} = \right] 2.95$		B1
	$\left[q = \right] \frac{350.05 - 40('2.95')^2}{39} = 0.05$		M1 A1
			(3)
(b)	$H_0 : \mu_A = \mu_B \quad H_1 : \mu_A < \mu_B$		B1
	$z = \pm \frac{2.65 - '2.95'}{\sqrt{\frac{0.07}{50} + \frac{'0.05'}{40}}}$		M1 M1
	$= 5.827... \text{ or } = -5.827...$ awrt ± 5.83		A1
	CV = 1.6449		B1
	Reject H_0 There is significant evidence to support the biologist's belief		M1 A1ft
			(7)
(c)	Large sample sizes so ...		
	both sample means are normally distributed (CLT)		B1
	$s_A^2 = \sigma_A^2$ and $s_B^2 = \sigma_B^2$		B1
			(2)
	Notes		Total 12
(a)	B1	2.95 only	
	M1	For use of $\frac{\sum x^2 - n\bar{x}^2}{n-1}$ oe ft their $\bar{x}$ May be implied 0.05 provided no incorrect working seen	
	A1	cao	
(b)	B1	Both hypotheses correct. Allow equivalent hypotheses. Must be in terms of μ	
	M1	For the denominator. Ft their 0.05	
	M1	Fully correct. Ft their 2.95 and their 0.05	
	A1	awrt 5.83 allow $ z = 5.827...$ accept $p = 2.8(1) \times 10^{-9}$	
	B1	$ CV = 1.6449$ or better	
	M1	A correct conclusion not in context ft their z value and CV or a correct p value (2 sf)	
	A1ft	ft their z value and their CV (NB their CV must be consistent with their z value) or a correct p value (2 sf). Correct conclusion in context, need belief/claim . May be in words with weights and region e.g. the weights in region A are smaller	
(c)	B1	Must comment on both sample means e.g. the sample means are normally distributed	
	B1	Must comment on both variances/standard deviations e.g. sample variances can be used as values for the population variances	

Question Number	Scheme			Marks
5 (a)(i) (ii)	Quota sampling would remove the need for a sampling frame oe			B1
	Quota sampling [can be/introduce] bias			B1
				(2)
(b)(i)	$\frac{(66+40) \times 120}{200} = 63.6$			M1 A1
(ii)	$(66+40) - 63.6 = 42.4$ or $\frac{(66+40) \times 80}{200} = 42.4$			A1
				(3)
(c)	H ₀ : Students favourite science subject and place lived are independent/not associated H ₁ : Students favourite science subject and placed lived are not independent/associated			B1
	Observed	Expected	$\frac{(O-E)^2}{E}$	M1
	66	63.6	$\frac{(66-63.6)^2}{63.6} \left[= \frac{24}{265} = 0.09056... \right]$	
	40	'42.4'	$\frac{(40-'42.4')^2}{'42.4'} \left[= \frac{36}{265} = 0.13584... \right]$	
	$\sum \frac{(O-E)^2}{E} = 4.549 + '0.09056...' + '0.13584...'$			M1
	$= 4.775...$ awrt 4.78			A1
	$\nu = (2-1)(3-1) = 2$			B1
	$\chi^2_{0.1}(2) = 4.605 \Rightarrow \text{CR: } \chi^2 \dots 4.605$			B1ft
	[in the CR/significant/Reject H ₀] There is sufficient evidence to suggest that students' favourite science subject is not independent of the place they live.			dA1ft
				(7)
Notes				Total 12
(a)(i)	B1	For a correct advantage. Possible advantages (but not an exhaustive list): includes all key subgroups, effective for small populations (Do not allow quick oe or cheap oe or easy oe)		
(ii)	B1	For a correct disadvantage. Possible disadvantages (but not an exhaustive list): [risk of] non-random [selection], difficulty in setting quotas		
(b)(i)	M1	For a correct method to find either expected frequency May be implied by 63.6 or 42.4		
(ii)	A1	For either 63.6 or 42.4		
	A1	For both 63.6 and 42.4		
(c)	B1	For both hypotheses correct. Must mention subject and place at least once. Do not allow correlation to imply association. Allow dependent to imply not independent		
	M1	A correct method for finding both contributions to the χ^2 value ft their 63.6 and their 42.4		
	M1	Adding their two values to 4.549 (may be implied by a full χ^2 calculation, do not ISW)		
	A1	awrt 4.78 NB This implies M1M1A1		
	B1	$\nu = 2$ This mark can be implied by a correct critical value of 4.605		
	B1ft	4.605 or better ft their degrees of freedom [$\chi^2_{0.1}(2) = 6.251$]		
	dA1ft	Dependent on both M marks being awarded. A correct contextual conclusion, which has the words subject and place (Allow 'where they live' to imply 'the place they live'). Allow an answer in terms of association. Do not allow correlation to imply association. Allow dependent to imply not independent ft their $\sum \frac{(O-E)^2}{E}$ and their χ^2 critical value This mark is independent of hypotheses		

Question Number	Scheme		Marks
6 (a)	$\alpha = 5.1$		B1
	$\beta = \sqrt{\frac{1694.65 - 65 \times ('5.1')^2}{64}}$		M1
	$= 0.25$		A1
			(3)
(b)	$H_0 : \mu_A = \mu_B$ $H_1 : \mu_A < \mu_B$		B1
	$z = \pm \frac{5.0 - '5.1'}{\sqrt{\frac{0.24^2}{70} + \frac{'0.25'^2}{65}}}$		M1 M1
	$= -2.367...$	awrt -2.37	A1
	One tailed c.v. $z = -1.6449$ or CR: $z \leq -1.6449$		B1
	In CR/Significant/Reject H_0		M1
	Sufficient evidence to support Roxane's claim		A1
			(7)
(c)	Since the sample is large the CLT applies.		M1
	No [need to assume that the fat content is normally distributed]		A1
			(2)
(d)	Assumed that $s^2 = \sigma^2$ in both groups		B1
			(1)
	Notes		Total 13
(a)	B1	cao	
	M1	For a correct method to find β using their α	
	A1	Cao	
(b)	B1	Both hypotheses correct. Allow equivalent hypotheses. Must be in terms of μ	
	M1	For correct standard error ft their s in part a	
	M1	For an attempt to find the test statistic, ft their SE and their α	
	A1	awrt -2.37 (Allow 2.37)	
	B1	-1.6449 or better (seen) (Allow 1.6449 or better if comparing to their 2.37)	
	M1	A correct statement – need not be contextual but do not allow contradicting non contextual comments. ft their CV and test statistic	
	A1	A correct contextual statement e.g sufficient evidence to support that crisps from brand A have a lower fat content than the crisps from brand B (must include the words in bold)	
(c)	M1	A suitable comment that mentions large and CLT	
	A1	A correct answer, context not required.	
(d)	B1	For the assumption that sample variance = population variance for both groups	