1. A teacher wishes to test whether playing background music enables students to complete a task more quickly. The same task was completed by 15 students, divided at random into two groups. The first group had background music playing during the task and the second group had no background music playing.
The times taken, in minutes, to complete the task are summarised below.

	Sample size n	Standard deviation s	Mean $\bar{x}$
With background music	8	4.1	15.9
Without background music	7	5.2	17.9

You may assume that the times taken to complete the task by the students are two independent random samples from normal distributions.

(a) Stating your hypotheses clearly, test, at the 10% level of significance, whether or not the variances of the times taken to complete the task with and without background music are equal.

(5)

(b) Find a 99% confidence interval for the difference in the mean times taken to complete the task with and without background music.

(7)

Experiments like this are often performed using the same people in each group.

(c) Explain why this would not be appropriate in this case.

(1)
(Total 13 marks)

2. As part of an investigation, a random sample of 10 people had their heart rate, in beats per minute, measured whilst standing up and whilst lying down. The results are summarised below.

Person	1	2	3	4	5	6	7	8	9	10
Heart rate lying down	66	70	59	65	72	66	62	69	56	68
Heart rate standing up	75	76	63	67	80	75	65	74	63	75

(a) State one assumption that needs to be made in order to carry out a paired t-test.

(1)

(b) Test, at the 5% level of significance, whether or not there is any evidence that standing up increases people's mean heart rate by more than 5 beats per minute. State your hypotheses clearly.

(8)
(Total 9 marks)

3. A farmer set up a trial to assess whether adding water to dry feed increases the milk yield of his cows. He randomly selected 22 cows. Thirteen of the cows were given dry feed and the other 9 cows were given the feed with water added. The milk yields, in litres per day, were recorded with the following results.

	Sample size	Mean	s^2
Dry feed	13	25.54	2.45
Feed with water added	9	27.94	1.02

You may assume that the milk yield from cows given the dry feed and the milk yield from cows given the feed with water added are from independent normal distributions.

(a) Test, at the 10% level of significance, whether or not the variances of the populations from which the samples are drawn are the same. State your hypotheses clearly.

(5)

(b) Calculate a 95% confidence interval for the difference between the two mean milk yields.

(7)

(c) Explain the importance of the test in part (a) to the calculation in part (b).

(2)
(Total 14 marks)

4. A large number of students are split into two groups *A* and *B*. The students sit the same test but under different conditions. Group A has music playing in the room during the test, and group B has no music playing during the test. Small samples are then taken from each group and their marks recorded. The marks are normally distributed.

The marks are as follows:

Sample from Group *A*	42	40	35	37	34	43	42	44	49
Sample from Group *B*	40	44	38	47	38	37	33		

(a) Stating your hypotheses clearly, and using a 10% level of significance, test whether or not there is evidence of a difference between the variances of the marks of the two groups.

(8)

(b) State clearly an assumption you have made to enable you to carry out the test in part (a).

(1)

(c) Use a two tailed test, with a 5% level of significance, to determine if the playing of music during the test has made any difference in the mean marks of the two groups. State your hypotheses clearly.

(7)

(d) Write down what you can conclude about the effect of music on a student's performance during the test.

(1)
(Total 17 marks)

5. The weights, in grams, of mice are normally distributed. A biologist takes a random sample of 10 mice. She weighs each mouse and records its weight.

The ten mice are then fed on a special diet. They are weighed again after two weeks.

Their weights in grams are as follows:

Mouse	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*
Weight before diet	50.0	48.3	47.5	54.0	38.9	42.7	50.1	46.8	40.3	41.2
Weight after diet	52.1	47.6	50.1	52.3	42.2	44.3	51.8	48.0	41.9	43.6

Stating your hypotheses clearly, and using a 1% level of significance, test whether or not the diet causes an increase in the mean weight of the mice.

(Total 8 marks)

6. A medical student is investigating two methods of taking a person's blood pressure. He takes a random sample of 10 people and measures their blood pressure using an arm cuff and a finger monitor. The table below shows the blood pressure for each person, measured by each method.

Person	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*
Arm cuff	140	110	138	127	142	112	122	128	132	160
Finger monitor	154	112	156	152	142	104	126	132	144	180

(a) Use a paired t-test to determine, at the 10% level of significance, whether or not there is a difference in the mean blood pressure measured using the two methods. State your hypotheses clearly.

(8)

(b) State an assumption about the underlying distribution of measured blood pressure required for this test.

(1)

(Total 9 marks)

7. As part of an investigation into the effectiveness of solar heating, a pair of houses was identified where the mean weekly fuel consumption was the same. One of the houses was then fitted with solar heating and the other was not. Following the fitting of the solar heating, a random sample of 9 weeks was taken and the table below shows the weekly fuel consumption for each house.

Week	1	2	3	4	5	6	7	8	9
Without Solar Heating	19	19	18	14	6	7	5	31	43
With Solar Heating	13	22	11	16	14	1	0	20	38

Units of fuel used per week

(a) Stating your hypotheses clearly test, at the 5% level of significance, whether or not there is evidence that the solar heating reduces the mean weekly fuel consumption.

(8)

(b) State an assumption about weekly fuel consumption that is required to carry out this test.

(1)
(Total 9 marks)

8. Seven pipes of equal length are selected at random. Each pipe is cut in half. One piece of each pipe is coated with protective paint and the other is left uncoated. All of the pieces of pipe are buried to the same depth in various soils for 6 months.

The table gives the percentage area of the pieces of pipe in the various soils that are subject to corrosion.

Soil	A	B	C	D	E	F	G
% Corrosion coated pipe	39	40	43	32	42	33	36
% Corrosion uncoated pipe	41	36	61	48	42	48	45

(a) Stating your hypotheses clearly and using a 5% significance level, carry out a paired *t*-test to assess whether or not there is a difference between the mean percentage of corrosion on the coated pipes and the mean percentage of corrosion on the uncoated pipes.

(9)

(b) (i) State an assumption that has been made in order to carry out this test.

(1)

(ii) Comment on the validity of this assumption.

(1)

(c) State what difference would be made to the conclusion in part (a) if the test had been to determine whether or not the percentage of corrosion on the uncoated pipes was higher than the mean percentage of corrosion on the coated pipes. Justify your answer.

(2)
(Total 13 marks)

9. A psychologist gives a test to students from two different schools, *A* and *B*.
A group of 9 students is randomly selected from school *A* and given instructions on how to do the test.

A group of 7 students is randomly selected from school *B* and given the test without the instructions.

The table shows the time taken, to the nearest second, to complete the test by the two groups.

A	11	12	12	13	14	15	16	17	17
B	8	10	11	13	13	14	14		

Stating your hypotheses clearly,

(a) test at the 10% significance level, whether or not the variance of the times taken to complete the test by students from school *A* is the same as the variance of the times taken to complete the test by students from school *B*. (You may assume that times taken for each school are normally distributed.)

(7)

(b) test at the 5% significance level, whether or not the mean time taken to complete the test by students from school *A* is greater than the mean time taken to complete the test by students from school *B*.

(7)

(c) Why does the result to part (a) enable you to carry out the test in part (b)?

(1)

(d) Give one factor that has not been taken into account in your analysis.

(1)
(Total 16 marks)

10. A farmer set up a trial to assess the effect of two different diets on the increase in the weight of his lambs. He randomly selected 20 lambs. Ten of the lambs were given diet *A* and the other 10 lambs were given diet *B*. The gain in weight, in kg, of each lamb over the period of the trial was recorded.

(a) State why a paired *t*-test is not suitable for use with these data.

(1)

(b) Suggest an alternative method for selecting the sample which would make the use of a paired *t*-test valid.

(1)

(c) Suggest two other factors that the farmer might consider when selecting the sample.

(2)

The following paired data were collected.

Diet *A*	5	6	7	4.6	6.1	5.7	6.2	7.4	5	3
Diet *B*	7	7.2	8	6.4	5.1	7.9	8.2	6.2	6.1	5.8

(d) Using a paired *t*-test, at the 5% significance level, test whether or not there is evidence of a difference in the weight gained by the lambs using diet *A* compared with those using diet *B*.

(8)

(e) State, giving a reason, which diet you would recommend the farmer to use for his lambs.

(1)
(Total 13 marks)

11. A doctor believes that the span of a person's dominant hand is greater than that of the weaker hand. To test this theory, the doctor measures the spans of the dominant and weaker hands of a random sample of 8 people. He subtracts the span of the weaker hand from that of the dominant hand. The spans, in mm, are summarised in the table below.

	A	*B*	*C*	*D*	*E*	*F*	*G*	*H*
Dominant hand	202	251	215	235	210	195	191	230
Weaker hand	195	249	218	234	211	197	181	225

Test, at the 5% significance level, the doctor's belief.

(Total 9 marks)

12. A grocer receives deliveries of cauliflowers from two different growers, *A* and *B*. The grocer takes random samples of cauliflowers from those supplied by each grower. He measures the weight *x*, in grams, of each cauliflower. The results are summarised in the table below.

	Sample size	$\sum x$	$\sum x^2$
A	11	6600	3960540
B	13	9815	7410579

(a) Show, at the 10% significance level, that the variances of the populations from which the samples are drawn can be assumed to be equal by testing the hypothesis H_0: $\sigma_A^2 = \sigma_B^2$ against hypothesis H_1: $\sigma_A^2 \neq \sigma_B^2$.

(You may assume that the two samples come from normal populations.)

(6)

The grocer believes that the mean weight of cauliflowers provided by *B* is at least 150 g more than the mean weight of cauliflowers provided by *A*.

(b) Use a 5% significance level to test the grocer's belief.

(8)

(c) Justify your choice of test.

(2)

(Total 16 marks)

13. Two methods of extracting juice from an orange are to be compared. Eight oranges are halved. One half of each orange is chosen at random and allocated to Method *A* and the other half is allocated to Method *B*. The amounts of juice extracted, in ml, are given in the table.

	Orange							
	1	2	3	4	5	6	7	8
Method *A*	29	30	26	25	26	22	23	28
Method *B*	27	25	28	24	23	26	22	25

One statistician suggests performing a two-sample *t*-test to investigate whether or not there is a difference between the mean amounts of juice extracted by the two methods.

(a) Stating your hypotheses clearly and using a 5% significance level, carry out this test.

(You may assume $\bar{x}_A = 26.125$, $s_A^2 = 7.84$, $\bar{x}_B = 25$, $s_B^2 = 4$ and $\sigma_A^2 = \sigma_B^2$)

(7)

Another statistician suggests analysing these data using a paired *t*-test.

(b) Using a 5% significance level, carry out this test.

(9)

(c) State which of these two tests you consider to be more appropriate. Give a reason for your choice.

(1)
(Total 17 marks)

1. (a) $\mathrm{H}_0 : \sigma_1^2 = \sigma_2^2,\ \mathrm{H}_1 : \sigma_1^2 \neq \sigma_2^2$ B1

critical values $\mathrm{F}_{6,7} = 3.87 \left(\frac{1}{\mathrm{F}_{6,7}} = 0.258 \right)$ B1

$\frac{s_2^2}{s_1^2} = \frac{5.2^2}{4.1^2}; = 1.61 \left(\frac{s_1^2}{s_2^2} = \frac{4.1^2}{5.2^2} = 0.622 \right)$ M1; A1

Since 1.61 (0.622) is not in the critical region we accept H_0 and conclude there is no evidence that the two variances are different A1ft 5

(b) $\mathrm{Sp}^2 = \frac{7 \times 4.1^2 + 6 \times 5.2^2}{7+6}$ = 21.53… M1 A1

t_{13}=3.012 B1

99% CI = (17.9 – 15.9) $\pm$ 3.012 $\times \sqrt{21.53} \times \sqrt{\frac{1}{8} + \frac{1}{7}}$ M1 A1ft

= $\pm$ (9.23, –5.233), [or accept: [0,9.23] or [–9.23,0]]
awrt 9.23, –5.23 A1 A1 7

(c) a person will be quicker at the task second time through/ times not independent/ familiar with the task/groups are not independent B1 1

<u>Note</u>

B1 Allow $\sigma_1 = \sigma_2$ and $\sigma_1 \neq \sigma_2$
B1 must match their F
M1 for $\frac{s_2^2}{s_1^2}$ or other way up
A1 awrt 1.61(0.622)
M1 A1 Sp^2 may be seen in part a
B1 3.012 only

M1 for (17.9 – 15.9) $\pm$ t value $\times \sqrt{\mathrm{S_p}^2} \times \sqrt{\frac{1}{8} + \frac{1}{7}}$
A1ft their Sp^2
A1 awrt 9.23/–9.23 A1 awrt –5.23/5.23
(c) B1 any correct sensible comment

[13]

2. (a) The differences in the mean heart rates are normally distributed. B1 1

(b) D = standing up – lying down

H_0: $\mu_D = 5$ H_1: $\mu_D > 5$ B1

d: 9, 6, 4, 2, 8, 9, 3, 5, 7, 7 M1

$\bar{d} = 6$; $s_d = \sqrt{\dfrac{414 - 10 \times 36}{9}} = 2.45$ M1; M1

$t_9 = \dfrac{6-5}{2.45/\sqrt{10}} = 1.29$ M1 A1

$t_9(5\%) = 1.833$ B1

insignificant. There is no evidence to suggest that heart rate rises by more than 5 beats when standing up. A1 ft 8

Note

must have "The differences in (mean heart rate) are normally distributed)
B1 both correct allow $\mu_D - 5 > 0$ ($\mu_D = -5$ H_1: $\mu_D < -5$)

M1 finding differences

M1 finding $\bar{d}$

M1 $\sqrt{\dfrac{\sum d^2 - 10 \times (\bar{d})^2}{9}}$ o.e.

M1 $\pm\left(\dfrac{6-5}{s_d/\sqrt{10}}\right)$ need to see full expression with numbers in

A1 awrt $\pm$ 1.29.

B1 $\pm$ 1.833 only

A1 ft their CV and t. Need context. Heart rate and 5 beats

[9]

3.	(a)	H_0: $\sigma_A^2 = \sigma_B^2$, $H_1 : \sigma_A^2 \neq \sigma_B^2$	B1	
		critical values $F_{12,8}$ =3.28 and $\frac{1}{F_{8,12}} = 0.35$	B1	
		$\frac{s_B^2}{s_A^2} = 2.40 \left(\frac{s_A^2}{s_B^2} = 0.416 \right)$	M1A1	
		Since 2.40 (0.416) is not in the critical region we accept H_0 and conclude there is no evidence that the two variances are different.	A1ft	5
	(b)	$\text{Sp}^2 = \frac{8 \times 1.02 + 12 \times 2.45}{20}$	M1	
		= 1.878	A1	
		(27.94 – 25.54) ± 2.086 × $\sqrt{1.878} \times \sqrt{\frac{1}{9} + \frac{1}{13}}$	B1M1 A1ft	
		(1.16, 3.64)	A1 A1	7
	(c)	To calculate the confidence interval the variances need to be equal.	B1	
		In part (a) the test showed they are equal.	B1	2

[14]

4.	(a)	$H_1 : \sigma_A^2 = \sigma_B^2$ $H_0 : \sigma_A^2 \neq \sigma_B^2$	B1	
		$S_A^2 = 22.5$ $S_B^2 = 21.6$ awrt	M1A1A1	
		$\frac{s_1^2}{s_2^2} = 1.04$	M1A1	
		$F_{(8, 6)} = 4.15$	B1	
		1.04 < 4.15 do not reject $H_{0.}$ The variances are the same.	B1	8
	(b)	Assume the samples are selected at random, (independent)	B1	1

(c)	$s_p^2 = \frac{8(22.5)+6(21.62)}{14} = 22.12$	awrt 22.1	M1A1	
	H_0: $\mu_A = \mu_B$ H_1: $\mu_A \neq \mu_B$		B1	
	$t = \frac{40.667-39.57}{\sqrt{22.12}\sqrt{\frac{1}{9}+\frac{1}{7}}}$		M1	
	$= 0.462$ $0.42 – 0.47$		A1	
	Critical value = $t_{14}(2.5\%) = 2.145$		B1	
	$0.462 < 2.145$ No evidence to reject H_0. The means are the same		B1	7
(d)	Music has no effect on performance		B1	1

[17]

5.	Differences 2.1 –0.7 2.6 –1.7 3.3 1.6 1.7 1.2 1.6 2.4		M1	
	$\bar{d} = 1.41$		M1	
	$H_0 : \mu_d = 0$ H_1: $\mu_d > 0$		B1	
	$s = \sqrt{\frac{40.65-10\times1.41^2}{9}} = 1.5191...$		M1	
	$t = \frac{1.41}{\left(\frac{1.519...}{\sqrt{10}}\right)} = 2.935$	awrt 2.94 /2.93	M1A1	
	t_9 (1%) = 2.821		B1	
	2.935.. > 2.821 Evidence to reject H_0. There has been an increase in the mean weight of the mice.		B1ft	8

2 sample test can score M0 M0

B1 for $H_0 : \mu_A = \mu_B$ H_1: $\mu_A < \mu_B$

M1 $\frac{9\times24.5+9\times17.16}{18}$

M0 A0

B1 2.552

B1 ft

ie 4/8

[8]

6. (a) d: 14 2 18 25 0 –8 4 4 12 20 M1

$\bar{d} = \pm 9.1$ sd $= \sqrt{106.7} = 10.332...$ A1A1

($\Sigma d = 91$, $\Sigma x^2 = 1789$)

H_0: $\mu_d = 0$ H_1: $\mu_d \neq 0$ B1

$t = \pm \dfrac{9.1\sqrt{10}}{10.332} = \pm 2.785$ awrt ±2.78 or 2.79 M1A1

Critical value $t_9 = \pm 1.833$ B1

Significant. There is a difference between <u>blood pressure</u> measured by arm cuff and finger monitor. A1 8

One tail test
Loses the first B1, CV is 1.383 in this case. Can get 7/8

(b) The <u>difference in measurements</u> of blood pressure is <u>normally</u> distributed B1 1

looking for the difference in measurements.
Not just it is normally distributed.

[9]

7. $\mu_1 = \mu_2$ etc is B0

$$\begin{bmatrix} \bar{D} = 0 & \bar{D} > 0 & \text{B0} \\ D = 0 & D > 0 & \text{B0} \end{bmatrix}$$

(D = Without solar heating – with solar heating)

(a) $H_0: \mu_D = 0$ $H_1: \mu_D > 0$ B1

d: 6,–3, 7, –2, –8, 6, 5, 11, 5 (attempted) M1

$\bar{d} = 3$, $S_d = 6$ $\left(= \sqrt{\dfrac{369 - 9 \times 3^2}{8}} \right)$ M1, M1

$t_8 = \dfrac{3-0}{6/\sqrt{9}}$ (±) M1 A1 c.a.o.

(links to A1ft below)

t_8 (5% 1 tail c.v.) = 1860 B1

Not significant – insufficient evidence of (that solar heating has) decreased weekly fuel consumption. A1ft 8

	(b)	Difference in weekly fuel consumption is normally distributed.	B1	1
				[9]
8.	(a)	d = U – C, 2, –4, 18, 16, 0, 15, 9	M1	
		$\bar{d} = \frac{5B}{7} = 8$	B1	
		$s_d^2 = \frac{906 - 7 \times 8^2}{6} = 76\frac{1}{3}$	M1 A1	
		$H_0 : \mu_d = 0$, $H_1 : \mu_d \neq 0$	B1	
		$t = \frac{8}{\sqrt{\frac{76.\dot{3}}{7}}} = 2.42260$ .. awrt 2.42	M1 A1	
		t (2.5%) = 2.447	B1	
		Insufficient evidence to reject H_0.		
		No evidence of a difference between the mean amount of correction on coated and uncoated pipes.	A1	9
	(b)	(i) Differences are normally distributed	B1	1
		(ii) Values do not appear to be normally distributed	B1	1
	(c)	to (5%) = 1.943. There is evidence to reject H_0.	B1	
		These is evidence to suggest that there is greater correction on coated pipes.	B1	2
				[13]
9.	(a)	$s_A{}^2 = 5.11$, $5_B{}^2 = 5.14$	B1 B1	
		H_0: $\sigma_A{}^2$, H_1: $\sigma_A{}^2 \neq \sigma_B{}^2$	B1	
		Critical value $F_{6,\,8} = 3.58$	B1	
		$\frac{s_B{}^2}{s_A{}^2} = 1.0062112$ ... awrt 1.01	M1 A1	
		No evidence to reject H_0. The variances are equal.	A1	7

	(b)	$s_p^2 = \dfrac{8\times 5.14 + 6\times 5.11}{9+7-2} = 5.1247$ awrt 5.12	M1 A1		
		$\mu_A = 14.11\,..\,,\ \mu_B = 11.857\,..$			
		$H_0 : \mu_A = \mu_B$, $H_1 : \mu_A > \mu_B$	B1		
		Critical value t_{14} (5%) = 1.761	B1		
		$T = \dfrac{14.11..-11.857...}{\sqrt{5.1247...\left(\frac{1}{9}+\frac{1}{7}\right)}} = 1.9757$ awrt 1.98	M1 A1		
		There is evidence to reject H_0. Mean time taken from school A is greater than school B.	A1	7	
	(c)	Equal variances are a condition for the test in part(b)	B1	1	
	(d)	Groups not equal ability	B1	1	
					[16]
10.	(a)	The data use not collected in pairs	B1	1	
	(b)	Use data from twin lambs	B1	1	
	(c)	Age, weight, gender *Any two sensible factors*	B1; B1	2	
	(d)	$d = B - A$			
		d = 2, 1.2, 1, 1.8, –1, 2.2, 2, –1.2, 1.1, 2.8	M1		
		$\sum d = 11.9$; $\sum d^2 = 30.01$			
		$\therefore \bar{d} = 1.19$; $s^2 = 1.761$ (s 1.327)	A1; A1		
		H_0: $\delta = 0$; H_1: $\delta \neq 0$ *Allow μ_D for δ*	both B1		
		$T = \dfrac{1.19 - 0}{\sqrt{1.761/10}} = 2.83574\ldots$	M1		
		Awrt 2.84 for A1	A1		
		$\propto = 9$; CV: $t = 2.262$	B1		
		Since 2.8357… is in the critical region ($t > 2.262$) there is evidence to reject H_0. The (mean) weight gained by the lambs is different for each diet.	A1ft	8	

Using non-paired t-test.

H_0: $\mu_A = \mu_B$; H_1: $\mu_A \neq \mu_B$ B1

$t = \dfrac{\mu_A - \mu_B}{\sqrt{s_p^2\left(\dfrac{1}{10}+\dfrac{1}{10}\right)}} = -2.30$ B1

Awrt –2.30

CV: $|t| = 2.101$ B1

Conclusion: Mean weight gained is different B1 4

<u>NB μ_A = 5.16 μ_B = 6.79 s_p^2 = 1.342722..</u>

(e) Diet B B1 1

[13]

11. d: 7 2 –3 1 –1 –2 10 5 M1

$\Sigma d = 19$; $\Sigma d^2 = 193$

$\therefore\ \bar{d} = \dfrac{19}{8} = \underline{2.375}$; $S_d^2 = \dfrac{1}{7}\left\{193 - \dfrac{19^2}{8}\right\} = \underline{21.125}$ B1; M1 A1

H_0: $\mu_D = 0$; H_1: $\mu_D > 0$ B1

both

$t = \dfrac{2.375-0}{\sqrt{\dfrac{21.125}{8}}} = \underline{1.4615\ldots..}$ M1

AWRT 1.46 A1

$\upsilon = 7 \Rightarrow$ critical region: $t > 1.895$ B1

Since 1.4615… is <u>not</u> in the critical region there is insufficient evidence to reject H_0 and we conclude that there is in sufficient evidence to support the doctors' belief. A1 ft

[9]

Alternative:

Use of 2 sample t-test $\Rightarrow$ 6/9 max

$S_p^2 = \dfrac{7\times 440.125 + 7\times 501.357}{8+8-2} = \underline{470.74}$ M1 A1

$t = \dfrac{216.125 - 213.75}{\sqrt{470.74\left(\frac{1}{8}+\frac{1}{8}\right)}} = \underline{0.2189\ldots}$ M1 A1

CR: $t > 1.761$ B1

Conclusion as above A1 ft

12. (a) $S_A^2 = \dfrac{1}{10}\{3960540 - \dfrac{6600^2}{11}\} = \underline{54.0}$ B1

$S_B^2 = \dfrac{1}{12}\{7410579 - \dfrac{9815^2}{13}\} = \underline{21.1\dot{6}}$ B1

$\text{H}_0\text{: } \sigma_A^2 = \sigma_B^2\text{; H}_1\text{: } \sigma_A^2 \neq \sigma_B^2$ B1

CR: $F_{10,\,12} > 2.75$

$S_A^2 / S_B^2 = \dfrac{54.0}{21.1\dot{6}} = 2.55118\ldots$ M1 A1

Since 2.55118… is not in the critical region, we can assume that the variances of A and B are equal. B1 6

(b) $\text{H}_0\text{: } \mu_B = \mu_A + 150\text{; H}_1\text{: } \mu_B > \mu_A + 150$ B1

both

CR: $t_{22}(0.05) > 1.717$ B1

$S_p^2 = \dfrac{10\times 54.0 + 12\times 21.1\dot{6}}{22} = \underline{36.0\dot{9}\dot{0}}$ M1 A1

$t = \dfrac{1755 - 6001 - 150}{\sqrt{36.0909\ldots\left(\frac{1}{11}+\frac{1}{13}\right)}} = \underline{2.03157}$ M1 A1

AWRT 2.03 A1

Since 2.03… is in the critical region we reject H_0 and conclude that the mean weight of cauliflowers from *B* exceeds that from *A* by at least 150g. A1 ft 8

(c)	Samples from normal populations		B1 B1	2
	Any two sensible verifications			
	Equal variances			
	Independent samples			

[16]

13.	(a)	$s_p^2 = \dfrac{7\times 7.84 + 7\times 4}{7+7} = 5.92$		M1	
		$s_p = 2.433105$	awrt 2.43	A1	
		$H_0 = \mu_A = \mu_B$, H_1: $\mu_A \neq \mu_B$	both	B1	
		$t = \dfrac{26.125 - 25}{2.43\sqrt{\frac{1}{8}+\frac{1}{8}}} = 0.92474$	awrt 0.925	M1 A1	
		$t_{14}(2.5\%) = 2.145$	2.145	B1	
		Insufficient evidence to reject H_0 that there is no difference in the means.		A1∫	7
	(b)	d = M1 –M2			
		2,5,-2,1,3,-4,1,3		M1	
		$\bar{d} = \dfrac{9}{8} = 1.125$	1.125	B1	
		$s_d^2 = \dfrac{69 - 8\times 1.125^2}{7} = 8.410714$	awrt 8.41	M1 A1	
		H_0: $\mu_d = 0$, H_1: $\mu_d \neq 0$	both	B1	
		$t = \dfrac{1.125}{\sqrt{\frac{8.41}{8}}} = 1.0972$	awrt 1.10	M1 A1	
		$t_7(2.5\%) = 2.365$	2.365	B1	
		There is no significant evidence of a difference between method A and method B.		A1∫	9
	(c)	Paired sample as they are two measurements on the same orange		B1	1

[17]

1. This proved to be a good starter question and most candidates gave good solutions.

Part (a) was answered well with many candidates gaining full marks.

In part(b) although a pooled estimate was worked out correctly by many candidates they then failed to use the square root of it in their calculations of the confidence interval or they used $\sqrt{\frac{21.53}{15}}$ rather than $\sqrt{21.53\left(\frac{1}{8}+\frac{1}{7}\right)}$. A few candidates found the confidence intervals for the mean times separately rather than for the difference.

2. In part (a) a minority of candidates realise that it is the "differences" which need to be normally distributed and not the distributions themselves.

In part (b) the most common error was an incorrect standard deviation. The majority of candidates were able to apply the method correctly and draw a conclusion in context.

3. Part (a) was answered well with many candidates gaining full marks. In part (b) although a pooled estimate was worked out correctly by many candidates they then failed to use the square root of it in their calculations of the confidence interval.

In part (c) candidates knew that to find the confidence interval/pooled estimate that the variances needed to be equal but few commented on the fact that this has been established in part (a).

4. This question was answered well with a large proportion of the candidates getting full marks. In part (b) many candidates stated that the marks were normally distributed which was not an assumption they had made as it had been stated in the question.
In part (c) the pooled estimate was worked out correctly by many candidates but they then failed to use the square root of it in their calculations of t. Many candidates were able to draw a conclusion in the context of the question.

5. Full marks were gained on this question by many candidates with a large proportion of the them using a paired t-test. A minority used a 2 sample test. It is important that candidates learn when to use which test as the test required will not always be stated in the question

6. This proved to be a good starter question and most candidates gave good solutions. A minority carried out a two sample test and a few did not interpret their conclusion in terms of the question. In part (b) many mentioned that the blood pressure had to be normally distributed. Whilst this is a sufficient condition the required answer was that the differences in blood pressure was normally distributed. Only a handful of candidates spotted this.

7. Most candidates identified the need to carry out a paired *t* test and the method was well known and clearly demonstrated. In part (b) many mentioned that the weekly fuel consumption had to be normally distributed. Whilst this is a sufficient condition the required answer was that the differences in weekly fuel consumption was normally distributed. Only a handful of candidates spotted this.

8. Hypotheses were often incorrect in part (a), but calculations usually gained most of the available credit. Candidates are still forgetting to state their conclusions in the context of the question set.

9. Candidates found this question very demanding and parts (a) and (b) were usually confused and incoherent due to accuracy errors.

10. The first three parts of this question were not well answered with many not knowing the conditions for the use of a paired *t*-test. Many could not relate to the practical aspects of this question. Apart from poor arithmetic part (d) was usually correct although the conclusion was not always in context. The correct diet was usually stated in part (e) but the reason was not always convincing.

11. The majority of the candidates recognised the need to use a paired *t*-test with only the weaker ones using a test for two means. The hypotheses were not always correctly specified even though the correct test was carried out and some candidates did not give their conclusion relative to the context of the question.

12. Many of the candidates scored full marks in part (a). In part (b) too many of them could not specify the hypotheses correctly since they did not know how to cope with the 150g. Surprisingly many of the candidates did not score both marks in the final part of this question.

13. This was done very well by a large number of candidates. They had clearly prepared well for the examination and it was not unusual to see the majority of the marks being awarded for this question. Occasional accuracy errors were seen, as were some incorrect statements of the hypotheses in part (b) together with missing context in the conclusions. On the whole the responses were very good.