1 A group of 10 school children are asked to estimate the size of an angle $\theta°$ in a given acute angled triangle. These estimates, in degrees, are as follows.

84 85 77 85 84 87 86 88 83 85

(a) Stating any assumptions you make, calculate a 95% confidence interval for θ. [5]

(b) Give a reason why the assumptions made in part **(a)** may not be appropriate in this case. [1]

3 A random sample of 10 newborn baby boys is taken and their masses in kg are recorded. From this sample, the population standard deviation of all newborn baby boys is estimated as 0.6 kg. A random sample of 5 newborn baby girls is taken and their masses in kg are recorded as follows.

3.9 3.1 2.9 3.1 3.6

It is assumed that the masses of newborn baby boys and girls have the same population standard deviation, σ kg.

By pooling the two samples, calculate an estimate of σ. [4]

6 Nine athletes in a club have a new coach. The coach adopts a new training programme which he believes will reduce the race times of these athletes. Each athlete completes a 1500 m time trial before and after completing the new training programme. Their times, in seconds (s), are recorded.

Athlete	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*
Time before training (s)	250	251	252	267	276	291	310	320	335
Time after training (s)	245	251	253	261	275	293	302	313	320

(a) Carry out a paired *t*-test at the 5% significance level to test the coach's belief. [7]

Further research suggests that the effects of the training programme tend to reduce the times of the slower athletes by more than those of the faster athletes.

(b) Suggest a reason why the paired *t*-test used in part **(a)** may not have been an appropriate test in this case. [1]

(c) Suggest a suitable alternative test that could have been used instead of a paired *t*-test. [1]

2 A factory produces packets of biscuits. The total mass of biscuits in a packet has a normal distribution with mean μ. A random sample of 12 packets is taken and the mass of the contents of each packet, x g, is recorded. The results are summarised as follows.

$$\sum x = 2390 \qquad \sum x^2 = 476117$$

(a) Find a 99% confidence interval for μ. [4]

A test of the null hypothesis $\mu = k$ is carried out on this sample using a 5% significance level. The test does not support the alternative hypothesis $\mu < k$.

(b) Find the greatest possible value of k. [3]

5 An engineer is comparing the tensile strengths of steel rods made from two machines, A and B. The engineer randomly selects 8 rods from machine A and 6 rods from machine B. The tensile strengths, in appropriate units, are given in the following table.

Machine A	402	403	415	412	409	407	406	410
Machine B	401	398	395	397	410	405		

You should assume that the two distributions are normal and have the same population variance.

Use a t-test at the 5% significance level to test whether there is any difference in the mean tensile strengths of steel rods from the two machines. [9]

1 Local residents are concerned about the speed of cars travelling through their village. They record the speed, in km h^{-1}, of a random sample of 13 cars travelling through their village. The recorded speeds are as follows.

40 53 59 42 43 48 62 67 46 82 66 45 70

(a) Construct a 90% confidence interval for the population mean speed of cars passing through the village. [5]

(b) State an assumption that is necessary for the confidence interval found in **1(a)** to be valid. [1]

3 Two different types of juice extractor, machine X and machine Y, are being compared. The manufacturer claims that machine Y extracts more juice per orange on average than machine X.

A random sample of 20 oranges is selected and each carefully measured. Oranges of a similar size and mass are then paired and numbered from 1 to 10. One orange from each pair is randomly allocated to machine X with the other allocated to machine Y. The amount of juice, in ml, extracted from each orange is recorded in the following table.

	pair									
	1	2	3	4	5	6	7	8	9	10
machine X	65	73	58	61	72	79	64	65	69	71
machine Y	68	72	64	63	75	82	63	63	72	74

(a) Use a t-test to test the manufacturer's claim at the 1% significance level. [7]

(b) State an assumption required for the test in **3(a)** to be valid. [1]

The manufacturer notices that the amount of juice extracted from the oranges in pair 4 were recorded incorrectly. In fact, the amount of juice extracted from machine X was 60 ml and from machine Y was 62 ml.

(c) Explain, with justification, whether the conclusion of the test in **3(a)** remains the same. [2]

1 The manager of a hardware store is interested in whether there is a difference in the amount spent per customer on weekdays ($\$x$) compared to weekends ($\$y$). Random samples of 120 customers on weekdays and 80 customers on weekends are taken and the amount spent by each customer is recorded. The results are summarised as follows.

$$\Sigma x = 10\,470 \qquad \Sigma(x-\overline{x})^2 = 12\,283 \qquad \Sigma y = 6560 \qquad \Sigma(y-\overline{y})^2 = 13\,520$$

Test at the 1% significance level whether there is a difference in the mean amount spent per customer on weekdays compared to weekends. You should not assume that the population variances of the amounts spent on weekdays and weekends are equal. [7]

4 A researcher is interested in whether there is a difference between two schools in students' aptitude for English. She randomly chooses 10 students from school X and 8 students of a similar age from school Y to take a written English test. The scores for the students from school X (x) and school Y (y) are summarised as follows.

$$\Sigma x = 612 \qquad \Sigma x^2 = 40\,104 \qquad \Sigma y = 444 \qquad \Sigma y^2 = 27\,460$$

You should assume that the two distributions are normal and have the same population variance.

(a) Find a 95% confidence interval for the difference in the mean scores for students from school X and students from school Y in the written English test. [6]

(b) Use the confidence interval you found in part **(a)** to explain why there is insufficient evidence at a 5% significance level to suggest that the English scores of students from school *X* and students from school *Y* are different. [1]

2 A farmer is investigating whether using a new fertiliser will increase the yield of tomato plants. The farmer selects 40 tomato plants at random and gives them the new fertiliser. The crop mass, x kg, of each of these 40 plants is recorded. The farmer selects a further 60 tomato plants at random and gives them a standard fertiliser. The crop mass, y kg, of each of these 60 plants is recorded. The results are summarised as follows.

$$\Sigma x = 168 \quad \Sigma x^2 = 720 \quad \Sigma y = 228 \quad \Sigma y^2 = 900$$

Find a 90% confidence interval for the difference in mean crop mass associated with each type of fertiliser. [7]

5 A doctor is investigating the concentration of blood glucose in patients at risk of developing type 2 diabetes, where blood glucose is measured in appropriate units. The doctor claims that a particular intervention reduces the concentration by more than k units on average. A group of 8 at risk patients is selected at random and each patient follows the intervention for six months. The blood glucose concentrations before and after the intervention are given in the following table.

Patient	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*
Before	183	165	172	165	143	176	161	153
After	164	148	164	149	134	153	155	148

(a) Use a t-test at the 5% significance level to find the range of values of k for which the result of the test is to reject the null hypothesis. [7]

(b) State an assumption necessary for the test in part **(a)** to be valid. [1]