1. (a) Explain why it is often useful to take samples as a means of obtaining information. **(2 marks)**

   (b) Briefly define the term **sampling frame**. **(1 mark)**

   (c) Suggest a suitable sampling frame for a sample survey on a proposal to install speed humps on a road. **(1 mark)**

2. An insurance company conducts its business by using a Call Centre. The average number of calls per minute is 3·5. In the first minute after a TV advertisement is shown, the number of calls received is 7.

   (a) Stating your hypotheses carefully, and working at the 5% significance level, test whether the advertisement has had an effect. **(5 marks)**

   (b) Find the number of calls that would be required in the first minute for the null hypothesis to be rejected at the 0·1% significance level. **(3 marks)**

3. On average, 35% of the candidates in a certain subject get an A or B grade in their exam. In a class of 20 students, find the probability that

   (a) less than 5 get A or B grades, **(2 marks)**

   (b) exactly 8 get A or B grades. **(2 marks)**

   Five such classes of 20 students are combined to sit the exam.

   (c) Use a suitable approximation to find the probability that less than a quarter of the total get A or B grades. **(6 marks)**

4. Light bulbs produced in a certain factory have lifetimes, in 100s of hours, whose distribution is modelled by the random variable X with probability density function

$$f(x) = \frac{2x(3-x)}{9}, \qquad 0 \le x \le 3;$$

$$f(x) = 0 \qquad \text{otherwise.}$$

   (a) Sketch $f(x)$. **(2 marks)**

   (b) Write down the mean lifetime of a bulb. **(1 mark)**

   (c) Show that ten times as many bulbs fail before 200 hours as survive beyond 250 hours. **(5 marks)**

   (d) Given that a bulb lasts for 200 hours, find the probability that it will then last for at least another 50 hours. **(2 marks)**

   (e) State, with a reason, whether you consider that the density function f is a realistic model for the lifetimes of light bulbs. **(1 mark)**

5. In a packet of 40 biscuits, the number of currants in each biscuit is as follows

Number of currants, x	0	1	2	3	4	5	6
Number of biscuits	4	9	11	8	4	3	1

(a) Find the mean and variance of the random variable X representing the number of currants per biscuit. **(4 marks)**

(b) State an appropriate model for the distribution of X, giving two reasons for your answer. **(2 marks)**

Another machine produces biscuits with a mean of 1·9 currants per biscuit.

(c) Determine which machine is more likely to produce a biscuit with at least two currants. **(5 marks)**

6. A greengrocer sells apples from a barrel in his shop. He claims that no more than 5% of the apples are of poor quality. When he takes 10 apples out for a customer, 2 of them are bad.

(a) Stating your hypotheses clearly, test his claim at the 1% significance level. **(5 marks)**

(b) State an assumption that has been made about the selection of the apples. **(1 mark)**

(c) When five other customers also buy 10 apples each, the numbers of bad apples they get are 1, 3, 1, 2 and 1 respectively. By combining all six customers' results, and using a suitable approximation, test at the 1% significance level whether the combined results provide evidence that the proportion of bad apples in the barrel is greater than 5%. **(5 marks)**

(d) Comment briefly on your results in parts (a) and (c). **(1 mark)**

7. Some children are asked to mark the centre of a scale 10 cm long. The position they choose is indicated by the variable X, where $0 \leq X \leq 10$. Initially, X is modelled as a random variable with a continuous uniform distribution.

(a) Find the mean and the standard deviation of X. **(3 marks)**

It is suggested that a better model would be the distribution with probability density function

$f(x) = cx,\ 0 \leq x \leq 5,$ $\quad f(x) = c(10 - x),\ 5 < x \leq 10,$ $\quad f(x) = 0$ otherwise.

(b) Write down the mean of X. **(1 mark)**

(c) Find c, and hence find the standard deviation of X in this model. **(7 marks)**

(d) Find $P(4 < X < 6)$. **(3 marks)**

It is then proposed that an even better model for X would be a Normal distribution with the mean and standard deviation found in parts (b) and (c).

(e) Use these results to find $P(4 < X < 6)$ in the third model. **(4 marks)**

(f) Compare your answer with (d). Which model do you think is most appropriate? **(1 mark)**

