SECTION A

1. A statistics test was taken by 30 students, and their marks (as percentages) were as follows.

41, 60, 63, 65, 73, 74, 75, 75, 76, 77, 78, 80, 81, 82, 83,
84, 84, 86, 86, 86, 87, 88, 88, 89, 90, 90, 90, 91, 93, 95.

(i) Display the above data as a stem-and-leaf plot. [4 marks]

(ii) Describe the shape of the distribution of marks. [2 marks]

(iii) Calculate the median value. For these data, would you prefer the median or the mean as a measure of location? Explain your answer. [2 marks]

2. (i) For two events A, B, express $P(A \cup B)$ in terms of $P(A)$, $P(B)$ and $P(A \cap B)$. [2 marks]

(ii) For three mutually exclusive events C, D, E, express $P(C \cup D \cup E)$ in terms of $P(C)$, $P(D)$ and $P(E)$. [2 marks]

(iii) Suppose three events F, G, H are such that G and H are independent, while F and G are mutually exclusive. Express $P(F \cup (G \cap H))$ in terms of $P(F)$, $P(G)$ and $P(H)$. [4 marks]

3. A continuous random variable X has probability density function $f(x)$ given by

$$f(x) = \begin{cases} K\,(9 + 6x - 3x^2) & \text{if } 0 \le x \le 2, \\ 0 & \text{otherwise,} \end{cases}$$

for some constant K.

(i) Find the value of K. [2 marks]

(ii) Calculate $E[X]$. [3 marks]

(iii) Calculate the variance of X. [3 marks]

4. The number of phone calls X received per ten-minute period at a call centre follows a Poisson distribution with mean 7.

(i) Write down an expression for $P(X = x)$, $x = 0, 1, 2, \ldots$ [2 mark]

(ii) Calculate $P(X = 10)$. [1 mark]

(iii) Calculate $P(X \leq 2)$. [2 marks]

(iv) Calculate $P(X \leq 1 \mid X \leq 2)$. [3 marks]

5. A random variable X is Normally distributed with mean 20 and variance 16. Calculate

(i) $P(X > 20)$ [1 mark]

(ii) $P(14 < X < 22)$ [3 marks]

(iii) The value a such that $P(X > a) = 0.2266$. [4 marks]

SECTION B

6. (a) A company has three factories producing a particular component, in Aberdeen, Bristol and Cambridge. 40% of the components are produced in Aberdeen; 40% in Bristol; and 20% in Cambridge. Of those components produced in Aberdeen, 5% are defective; of those produced in Bristol, 4% are defective; of those produced in Cambridge, 6% are defective.

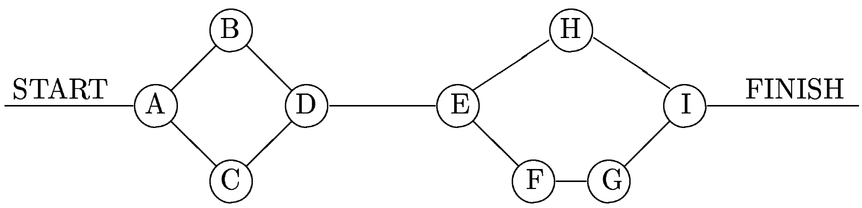
(i) A component is chosen at random from those produced by the company. Calculate the probability that this component is defective. [2 marks]

(ii) Given that a component, chosen at random, is found to be defective, calculate the probability that it was produced in the Bristol factory. [4 marks]

(iii) Given that a component, chosen at random, is found **not** to be defective, calculate the probability that it was produced in the Bristol factory. [4 marks]

(b) The diagram below shows a network of unreliable components. For each component X, the probability that X is working is denoted by $P(X)$, and whether X is working or not is independent of whether or not any other component is working. Calculate the network reliability (that is, the probability that a signal will pass successfully from START to FINISH) given that

$$P(A) = 0.9, \quad P(B) = 0.9, \quad P(C) = 0.8, \quad P(D) = 0.8,$$
$$P(E) = 0.8, \quad P(F) = 0.9, \quad P(G) = 0.9, \quad P(H) = 0.8, \quad P(I) = 0.8.$$



[10 marks]

7. (a) In sample of 80 families, each with 4 children, the number of male children in each family was recorded, with the following results.

Number of male children	0	1	2	3	4
Number of families	3	10	31	24	12

Denoting by p the probability that a child is born male, test at the 5% level the hypothesis that these data are drawn from a binomial distribution with $p = 0.5$.

[7 marks]

Rather than assuming that $p = 0.5$, we could use the data to estimate the value of p and then test whether a binomial distribution with p estimated from the data provides an adequate fit to the data.

Use the given data to estimate the value of p. [1 mark]

For a binomial distribution with p estimated from the data, the goodness-of-fit test statistic is calculated to be $X^2 = 2.01$. Test at the 5% level whether this X^2 value provides evidence against the hypothesis that the data are drawn from a binomial distribution with p estimated from the data. [2 marks]

(b) In a clinical trial, 76 patients were treated with either a vaccine or a placebo, and each patient's response recorded as 'Small', 'Moderate' or 'Large'. The results were as follows.

	Response		
	Small	Moderate	Large
Vaccine	6	20	12
Placebo	25	8	5

Test at the 5% significance level the hypothesis that a patient's response is independent of the treatment received.

[10 marks]

8. (a) Machine components are mass-produced at a factory. Components are acceptable provided they are between 45.95 mm and 46.05 mm long. In practice, it is found that 6.3% of those produced are over-sized and 2.5% are under-sized.

   (i) Assuming that the lengths of components are Normally distributed with mean μ and standard deviation σ, find the values of μ and σ. [8 marks]

   (ii) If a sample of 5 components is taken, calculate the probability that at least 4 are of acceptable length. [4 marks]

(b) In a study of low-birthweight infants, the Psychomotor Development Index (PDI) was recorded for each of 56 low-birthweight infants. The sample mean and standard deviation were calculated to be $\bar{x} = 104.125$ and $s = 12.584$, respectively. Assuming that PDI values are Normally distributed, calculate a 95% confidence interval for the mean.

The mean PDI value for the entire infant population is 100. On the basis of the confidence interval which you have calculated, comment on whether you believe the mean PDI value for low-birthweight infants differs from 100.

[8 marks]

9. (a) A company producing a particular electrical component tests the lifetimes of a sample of 200 components, recording the time (in days of continuous use) until failure. The table below summarises the resulting data.

Lifetime (days)	Frequency
0–2	41
2–4	24
4–6	24
4–8	18
8–10	13
10–15	37
15–20	27
20–30	11
30–40	5

Draw a histogram of these data. [8 marks]

Comment, with reasoning, on whether the histogram suggests the data are Normally distributed. [2 marks]

(b) A continuous random variable X has probability density function

$$f(x) = \begin{cases} 0 & \text{if } x < 0, \\ \lambda e^{-\lambda x} & \text{if } 0 \leq x < \infty, \end{cases}$$

where λ is some constant with $\lambda > 0$.

(i) Find $P(X < 10)$ in terms of λ. [2 marks]

(ii) Suppose now it is thought that the random variable X with density function given above may be used to model the data on lifetimes of electrical components given in part (a).
Use the data given in part (a) to estimate the probability that a randomly chosen component has a lifetime of less than 10 days. [1 mark]

(iii) By equating your answers for (i) and (ii) above, estimate the value of λ. [3 marks]

(iv) Using the λ value computed in (iii), estimate the median lifetime of the electrical components — that is, find the value m such that $P(X < m) = 0.5$. [4 marks]

10. (a) For a random variable X which is Poisson distributed with parameter λ, show that the expectation $E[X]$ is given by $E[X] = \lambda$.

Show also that the variance of X is equal to λ. [10 marks]

(b) The Maths Department photocopier breaks down, on average, 3 times per week. Assuming that the number of breakdowns in a week follows a Poisson distribution, calculate the probability that the number of breakdowns in one particular week is 4 or more. [3 marks]

Assuming that the number of breakdowns during a 10 week period follows a Poisson distribution, write down the value of the mean of this distribution, and by using the Normal approximation to the Poisson distribution estimate the probability that the number of breakdowns in a particular 10 week period is 40 or more. [7 marks]