SECTION A

1. For a sample of 20 graduating MBA students, the highest starting salary offer received by each student (in thousands of dollars) was recorded, the data being as follows.

   61, 51, 53, 42, 55, 42, 39, 72, 48, 47, 49, 62, 40, 40, 55, 56, 47, 55, 62, 43

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

   (ii) Calculate the median and the interquartile range of the above data. [4 marks]

2. A discrete random variable X taking values $0, 1, \ldots, 6$ has probability mass function $p(x)$ given by

x	0	1	2	3	4	5	6
$p(x)$	$3k$	$2k$	k	0	$2k$	$4k$	$6k$

   for some constant k.

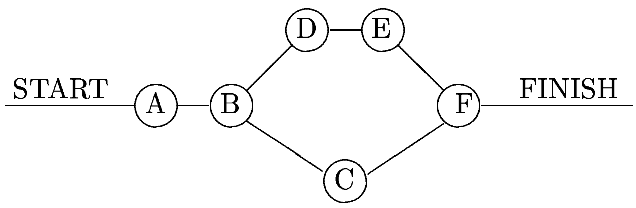
   (i) Find the value of k. [1 mark]

   (ii) Calculate $P(X > 1)$. [2 marks]

   (ii) Calculate $E[X]$. [2 marks]

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

3. The diagram below shows a network of unreliable components. For each component X, the probability that X is working is $P(X) = 0.8$, 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).



[8 marks]

4. A company produces fuses on a production line, and it is known from past experience that on average 2% of fuses are defective. A random sample of 10 fuses is taken from the production line.

   (i) Denoting by X the number of defective fuses in the sample, write down an expression for $P(X = x)$, $x = 0, 1, 2, \ldots, 10$. [2 mark]

   (ii) Calculate the expected number of defective fuses in the sample. [1 mark]

   (iii) Calculate the probability that there are no defective fuses in the sample.. [2 marks]

   (iv) Calculate the probability that there are more than 3 defective fuses in the sample. [3 marks]

5. A random variable X is normally distributed with mean 9 and variance 4. Calculate

   (i) $P(X > 7)$ [2 marks]

   (ii) $P(7 < X < 12)$ [3 marks]

   (iii) The value a such that $P(X < a) = 0.3085$. [3 marks]

SECTION B

6. (a) For two events A, B with $P(B) \neq 0$, write down the definition of the conditional probability $P(A \mid B)$. [2 marks]

   (b) In a survey of male business managers, 55% of those surveyed had cheated at sports. Of those who had cheated at sports, 30% had also lied in business. Of those who had not cheated at sports, only 20% had lied in business.

   (i) For a business manager chosen at random from those surveyed, calculate the probability that he had lied in business. [2 marks]

   (ii) Given that a business manager, chosen at random from those surveyed, had lied in business, calculate the probability that he had cheated at sports. [4 marks]

   (c) Suppose A, B, C are three events such that A and B are independent, while B and C are mutually exclusive. Express the following probabilities in terms of $P(A)$, $P(B)$ and $P(C)$.

   (i) $P(A \cap B)$ [1 mark]

   (ii) $P(A \cup B)$ [2 marks]

   (iii) $P(B \cup C)$ [2 marks]

   (iv) $P(B \mid A)$ [1 mark]

   (v) $P(B \mid (C \cap A))$ [2 marks]

   (vi) $P(C \cup (A \cap B))$ [4 marks]

7. (a) In a study of the effects of alcohol consumption on patients who had suffered a non-fatal heart attack, patients were classified according to whether or not they had Congestive Heart Failure (CHF, a serious ongoing heart problem following on from the heart attack), and also according to how much alcohol they usually drank each week, classified as 'none', 'low' or 'high'. The results were as follows.

	Alcohol consumption		
	None	Low	High
CHF	146	106	29
No CHF	750	590	292

Test at the 5% significance level the hypothesis that whether or not a patient suffers CHF is independent of their level of alcohol consumption. [10 marks]

(b) A can company believes that the number of breakdowns per 8-hour shift on its machine-operated assembly line follows a Poisson distribution with mean 1.1. Recording the numbers of breakdowns per shift for 70 shifts, the following data were obtained.

Number of breakdowns	0	1	2	3	4	≥ 5
Number of shifts	16	22	18	9	5	0

Test at the 5% level the hypothesis that these data are indeed drawn from a Poisson distribution with mean 1.2.

Use the data to estimate the mean number of breakdowns per shift. Test at the 5% level the hypothesis that the data are drawn from a Poisson distribution with the mean determined from the data. [10 marks]

8. (a) In a study of the effect of a particular drug on response times, 100 rats were each injected with a fixed dose of the drug. Each rat was then subjected to a stimulus and its response time (in seconds) recorded. The sample mean and standard deviation were calculated to be $\bar{x} = 1.16$ seconds and $s = 0.16$ seconds, respectively. Assuming that the response time values are normally distributed, calculate a 95% confidence interval for the mean.

The mean response time for rats who have not received the drug is known to be 1.2 seconds. On the basis of the confidence interval which you have calculated, comment on whether you believe the mean response time for rats who have received the drug differs from 1.2 seconds. [8 marks]

(b) In the context of statistical hypothesis testing, explain briefly, using the example of part (a) above to illustrate your answer, what is meant by each of the following terms.

Null hypothesis; Alternative hypothesis; Type I error; Type II error; Significance level; Power. [6 marks]

(c) For the example of part (a) above, denoting by μ the mean response time for rats who have receieved the drug, carry out a test at the 5% level of the hypothesis

$$H_0 : \mu = 1.2$$

$$\text{versus } H_1 : \mu \neq 1.2.$$

[6 marks]

9. (a) At city centre parking meters, the length of stay is limited to at most 2 hours. The length of time T (in hours) for which a randomly chosen car will park at such a meter is thought to be a continuous random variable with probability density function $f(t)$ given by

$$f(t) \quad = \quad \begin{cases} K\,(t^3 - 4t^2 + 5t) & \text{if } 0 \le t \le 2, \\ 0 & \text{otherwise,} \end{cases} \qquad (*)$$

for some constant K.

(i) Find the value of K.

(ii) Calculate $E[T]$.

(iii) Calculate the variance of T. [8 marks]

(b) For a random sample of 10 cars parked at meters as described in part (a) above, assuming that cars behave independently of one another, calculate the probability that none of the 10 cars spends more than 1 hour parked at the meter. [4 marks]

(c) The cost of parking at a meter as described in part (a) above depends upon the length of stay as follows.

Length of stay T (hours)	0–0.5	0.5–1.0	1.0–1.5	1.5–2.0
Cost (£)	0.4	0.8	1.2	1.6

Denoting by X the cost (in £) of parking for a randomly selected car with length of stay T having probability density function $f(t)$ given by $(*)$ above, calculate the probability mass function of X (that is, find $P(X = x)$ for $x = 0.4,\ 0.8,\ 1.2,\ 1.6$).

Calculate the expected cost of parking $E[X]$. [8 marks]

10. (a) For a random variable X which is binomially distributed with parameters n, p, show that the expectation $E[X]$ is given by $E[X] = np$.

Write down (without proof) an expression in terms of n and p for the variance of X.

[8 marks]

(b) In the USA, 22% of all births take place by Caesarian section. A researcher takes a random sample of 50 births. Assuming that the number of births in the sample which take place by Caesarian section follows a binomial distribution, calculate the probability that the number of births X which take place by Caesarian section lies in the range $11 \leq X \leq 14$. [4 marks]

Use the normal approximation to the binomial distribution to estimate the probability that the number of births X which take place by Caesarian section lies in the range $11 \leq X \leq 14$. [6 marks]

Comment on the accuracy of your approximation. [2 marks]