Answer **all** questions.

1 A council claims that 80 per cent of households are generally satisfied with the services it provides.

A random sample of 250 households shows that 209 are generally satisfied with the council's provision of services.

(a) Construct an approximate 95% confidence interval for the proportion of households that are generally satisfied with the council's provision of services. *(6 marks)*

(b) Hence comment on the council's claim. *(2 marks)*

2 The table below shows the heart rates, x beats per minute, and the systolic blood pressures, y milligrams of mercury, of a random sample of 10 patients undergoing kidney dialysis.

Patient	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
x	83	86	88	92	94	98	101	111	115	121
y	157	172	161	154	171	169	179	180	192	182

(a) Calculate the value of the product moment correlation coefficient for these data. *(3 marks)*

(b) Assuming that these data come from a bivariate normal distribution, investigate, at the 1% level of significance, the claim that, for patients undergoing kidney dialysis, there is a positive correlation between heart rate and systolic blood pressure. *(4 marks)*

3 Each enquiry received by a business support unit is dealt with by Ewan, Fay or Gaby. The probabilities of them dealing with an enquiry are 0.2, 0.3 and 0.5 respectively.

Of enquiries dealt with by Ewan, 60% are answered immediately, 25% are answered later the same day and the remainder are answered at a later date.

Of enquiries dealt with by Fay, 75% are answered immediately, 15% are answered later the same day and the remainder are answered at a later date.

Of enquiries dealt with by Gaby, 90% are answered immediately and the remainder are answered at a later date.

 (a) Determine the probability that an enquiry:

   (i)   is dealt with by Gaby and answered immediately;   *(1 mark)*

   (ii) is answered immediately;   *(3 marks)*

   (iii) is dealt with by Gaby, given that it is answered immediately.   *(3 marks)*

 (b) Determine the probability that an enquiry is dealt with by Ewan, given that it is answered later the same day.   *(4 marks)*

4 The table below shows the probability distribution for the number of students, R, attending classes for a particular mathematics module.

r	6	7	8
$P(R=r)$	0.1	0.6	0.3

 (a) Find values for $E(R)$ and $Var(R)$.   *(4 marks)*

 (b) The number of students, S, attending classes for a different mathematics module is such that

$$E(S) = 10.9, \quad Var(S) = 1.69 \quad \text{and} \quad \rho_{RS} = \tfrac{2}{3}$$

   Find values for the mean and variance of:

   (i) $T = R + S$;   *(4 marks)*

   (ii) $D = S - R$.   *(2 marks)*

Turn over for the next question

5 The number of letters per week received at home by Rosa may be modelled by a Poisson distribution with parameter 12.25.

   (a) Using a normal approximation, estimate the probability that, during a 4-week period, Rosa receives at home at least 42 letters but at most 54 letters. *(5 marks)*

   (b) Rosa also receives letters at work. During a 16-week period, she receives at work a total of 248 letters.

       (i) Assuming that the number of letters received at work by Rosa may also be modelled by a Poisson distribution, calculate a 98% confidence interval for the average number of letters per week received at work by Rosa. *(5 marks)*

       (ii) Hence comment on Rosa's belief that she receives, on average, fewer letters at home than at work. *(2 marks)*

6 The random variable X has a Poisson distribution with parameter λ.

   (a) Prove that $E(X) = \lambda$. *(3 marks)*

   (b) By first proving that $E(X(X-1)) = \lambda^2$, or otherwise, prove that $Var(X) = \lambda$. *(5 marks)*

7 A shop sells cooked chickens in two sizes: medium and large.

   The weights, X grams, of medium chickens may be assumed to be normally distributed with mean μ_X and standard deviation 45.

   The weights, Y grams, of large chickens may be assumed to be normally distributed with mean μ_Y and standard deviation 65.

   A random sample of 20 medium chickens had a mean weight, $\bar{x}$ grams, of 936.

   A random sample of 10 large chickens had the following weights in grams:

   1165   1202   1077   1144   1195   1275   1136   1215   1233   1288

   (a) Calculate the mean weight, $\bar{y}$ grams, of this sample of large chickens. *(1 mark)*

   (b) Hence investigate, at the 1% level of significance, the claim that the mean weight of large chickens exceeds that of medium chickens by more than 200 grams. *(8 marks)*

   (c) (i) Deduce that, for your test in part (b), the critical value of $(\bar{y} - \bar{x})$ is 253.24, correct to two decimal places. *(2 marks)*

       (ii) Hence determine the power of your test in part (b), given that $\mu_Y - \mu_X = 275$. *(5 marks)*

       (iii) Interpret, in the context of this question, the value that you obtained in part (c)(ii). *(3 marks)*

END OF QUESTIONS.

Practice 2

Leave blank

1. Some biologists were studying a large group of wading birds. A random sample of 36 were measured and the wing length, x mm of each wading bird was recorded. The results are summarised as follows

$$\sum x = 6046 \qquad \sum x^2 = 1\,016\,338$$

(a) Calculate unbiased estimates of the mean and the variance of the wing lengths of these birds.

(3)

Given that the standard deviation of the wing lengths of this particular type of bird is actually 5.1 mm,

(b) find a 99% confidence interval for the mean wing length of the birds from this group.

(5)

2

Question 1 continued

(Total 8 marks)

Q1

Turn over

Leave
blank

2. Students in a mixed sixth form college are classified as taking courses in either Arts, Science or Humanities. A random sample of students from the college gave the following results

		Course		
		Arts	Science	Humanities
Gender	Boy	30	50	35
	Girl	40	20	42

Showing your working clearly, test, at the 1% level of significance, whether or not there is an association between gender and the type of course taken. State your hypotheses clearly.

(11)

4

Question 2 continued

Leave blank

Turn over

Question 2 continued

6

Leave
blank

Question 2 continued

Q2

(Total 11 marks)

7

Turn over

Leave blank

3. The product moment correlation coefficient is denoted by r and Spearman's rank correlation coefficient is denoted by r_s.

(a) Sketch separate scatter diagrams, with five points on each diagram, to show

(i) $r = 1$,

(ii) $r_s = -1$ but $r > -1$.

(3)

Two judges rank seven collie dogs in a competition. The collie dogs are labelled A to G and the rankings are as follows

Rank	1	2	3	4	5	6	7
Judge 1	A	C	D	B	E	F	G
Judge 2	A	B	D	C	E	G	F

(b) (i) Calculate Spearman's rank correlation coefficient for these data.

(6)

(ii) Stating your hypotheses clearly, test, at the 5% level of significance, whether or not the judges are generally in agreement.

(5)

8

Question 3 continued

Leave
blank

9

Turn over

Leave
blank

Question 3 continued

Leave blank

Question 3 continued

(Total 14 marks)

Q3

Leave blank

4. The weights of adult men are normally distributed with a mean of 84 kg and a standard deviation of 11 kg.

   (a) Find the probability that the total weight of 4 randomly chosen adult men is less than 350 kg.

   (5)

   The weights of adult women are normally distributed with a mean of 62 kg and a standard deviation of 10 kg.

   (b) Find the probability that the weight of a randomly chosen adult man is less than one and a half times the weight of a randomly chosen adult woman.

   (6)

12

Leave
blank

Question 4 continued

Leave
blank

Question 4 continued

14

Question 4 continued

Leave blank

Q4

(Total 11 marks)

Leave blank

5. A researcher is hired by a cleaning company to survey the opinions of employees on a proposed pension scheme. The company employs 55 managers and 495 cleaners.

To collect data the researcher decides to give a questionnaire to the first 50 cleaners to leave at the end of the day.

(a) Give 2 reasons why this method is likely to produce biased results.

(2)

(b) Explain briefly how the researcher could select a sample of 50 employees using

   (i) a systematic sample,

   (ii) a stratified sample.

(6)

Using the random number tables in the formulae book, and starting with the top left hand corner (8) and working across, 50 random numbers between 1 and 550 inclusive were selected. The first two suitable numbers are 384 and 100.

(c) Find the next two suitable numbers.

(2)

16

Leave
blank

Question 5 continued

Q5

(Total 10 marks)

Leave
blank

6. Ten cuttings were taken from each of 100 randomly selected garden plants. The numbers of cuttings that did not grow were recorded.

The results are as follows

No. of cuttings which did not grow	0	1	2	3	4	5	6	7	8, 9 or 10
Frequency	11	21	30	20	12	3	2	1	0

(a) Show that the probability of a randomly selected cutting, from this sample, not growing is 0.223

(2)

A gardener believes that a binomial distribution might provide a good model for the number of cuttings, out of 10, that do not grow.

He uses a binomial distribution, with the probability 0.2 of a cutting not growing. The calculated expected frequencies are as follows

No. of cuttings which did not grow	0	1	2	3	4	5 or more
Expected frequency	r	26.84	s	20.13	8.81	t

(b) Find the values of r, s and t.

(4)

(c) State clearly the hypotheses required to test whether or not this binomial distribution is a suitable model for these data.

(2)

The test statistic for the test is 4.17 and the number of degrees of freedom used is 4.

(d) Explain fully why there are 4 degrees of freedom.

(2)

(e) Stating clearly the critical value used, carry out the test using a 5% level of significance.

(3)

Question 6 continued

19

Question 6 continued

http://www.mppe.org.uk

Leave
blank

Question 6 continued

Q6

(Total 13 marks)

Leave blank

7. A sociologist is studying how much junk food teenagers eat. A random sample of 100 female teenagers and an independent random sample of 200 male teenagers were asked to estimate what their weekly expenditure on junk food was. The results are summarised below.

	n	mean	s.d.
Female teenagers	100	£5.48	£3.62
Male teenagers	200	£6.86	£4.51

(a) Using a 5% significance level, test whether or not there is a difference in the mean amounts spent on junk food by male teenagers and female teenagers. State your hypotheses clearly.

(7)

(b) Explain briefly the importance of the central limit theorem in this problem.

(1)

22

Question 7 continued

23

Turn over

Question 7 continued

(Total 8 marks)

TOTAL FOR PAPER: 75 MARKS

END

Q7

24

Answer **all** questions.

---

1 As part of an investigation into the starting salaries of graduates in a European country, the following information was collected.

		Starting salary (€)	
	Sample size	**Sample mean**	**Sample standard deviation**
Science graduates	175	19 268	7321
Arts graduates	225	17 896	8205

  (a) Stating a necessary assumption about the samples, construct a 98% confidence interval for the difference between the mean starting salary of science graduates and that of arts graduates. *(6 marks)*

  (b) What can be concluded from your confidence interval? *(2 marks)*

2 A hill-top monument can be visited by one of three routes: road, funicular railway or cable car. The percentages of visitors using these routes are 25, 35 and 40 respectively.

The age distribution, in percentages, of visitors using **each** route is shown in the table. For example, 15 per cent of visitors using the road were under 18.

		Percentage of visitors using		
		Road	**Funicular railway**	**Cable car**
	Under 18	15	25	10
Age (years)	**18 to 64**	80	60	55
	Over 64	5	15	35

Calculate the probability that a randomly selected visitor:

  (a) who used the road is aged 18 or over; *(1 mark)*

  (b) is aged between 18 and 64; *(3 marks)*

  (c) used the funicular railway and is aged over 64; *(2 marks)*

  (d) used the funicular railway, given that the visitor is aged over 64. *(5 marks)*

3 Kutz and Styler are two unisex hair salons. An analysis of a random sample of 150 customers at Kutz shows that 28 per cent are male. An analysis of an independent random sample of 250 customers at Styler shows that 34 per cent are male.

(a) Test, at the 5% level of significance, the hypothesis that there is no difference between the proportion of male customers at Kutz and that at Styler. *(9 marks)*

(b) State, with a reason, the probability of making a Type I error in the test in part (a) if, in fact, the actual difference between the two proportions is 0.05. *(2 marks)*

4 A machine is used to fill 5-litre plastic containers with vinegar. The volume, in litres, of vinegar in a container filled by the machine may be assumed to be normally distributed with mean μ and standard deviation 0.08.

A quality control inspector requires a 99% confidence interval for μ to be constructed such that it has a width of at most 0.05 litres.

Calculate, to the nearest 5, the sample size necessary in order to achieve the inspector's requirement. *(6 marks)*

5 The duration, X minutes, of a timetabled 1-hour lesson may be assumed to be normally distributed with mean 54 and standard deviation 2.

The duration, Y minutes, of a timetabled $1\frac{1}{2}$-hour lesson may be assumed to be normally distributed with mean 83 and standard deviation 3.

Assuming the durations of lessons to be independent, determine the probability that the total duration of a random sample of three 1-hour lessons is less than the total duration of a random sample of two $1\frac{1}{2}$-hour lessons. *(7 marks)*

Turn over for the next question

6 (a) The random variable X has a binomial distribution with parameters n and p.

 (i) Prove that $\mathrm{E}(X) = np$. *(4 marks)*

 (ii) Given that $\mathrm{E}(X^2) - \mathrm{E}(X) = n(n-1)p^2$, show that $\mathrm{Var}(X) = np(1-p)$.

 (3 marks)

 (iii) Given that X is found to have a mean of 3 and a variance of 2.97, find values for n and p. *(3 marks)*

 (iv) Hence use a distributional approximation to estimate $\mathrm{P}(X > 2)$. *(3 marks)*

 (b) Dressher is a nationwide chain of stores selling women's clothes. It claims that the probability that a customer who buys clothes from its stores uses a Dressher store card is 0.45.

 Assuming this claim to be correct, use a distributional approximation to estimate the probability that, in a random sample of 500 customers who buy clothes from Dressher stores, at least half of them use a Dressher store card. *(7 marks)*

7 In a town, the total number, R, of houses sold during a week by estate agents may be modelled by a Poisson distribution with a mean of 13.

A new housing development is completed in the town. During the first week in which houses on this development are offered for sale by the developer, the estate agents sell a total of 10 houses.

 (a) Using the 10% level of significance, investigate whether the offer for sale of houses by the developer has resulted in a reduction in the mean value of R. *(6 marks)*

 (b) Determine, for your test in part (a), the critical region for R. *(2 marks)*

 (c) Assuming that the offer for sale of houses on the new housing development has reduced the mean value of R to 6.5, determine, for a test at the 10% level of significance, the probability of a Type II error. *(4 marks)*

END OF QUESTIONS

Leave blank

1. A telephone directory contains 50 000 names. A researcher wishes to select a systematic sample of 100 names from the directory.

   (a) Explain in detail how the researcher should obtain such a sample.

   (2)

   (b) Give one advantage and one disadvantage of

      (i) quota sampling,

      (ii) systematic sampling.

      (4)

Leave
blank

Question 1 continued

(Total 6 marks)

Leave blank

2. The heights of a random sample of 10 imported orchids are measured. The mean height of the sample is found to be 20.1 cm. The heights of the orchids are normally distributed.

Given that the population standard deviation is 0.5 cm,

(a) estimate limits between which 95% of the heights of the orchids lie,

(3)

(b) find a 98% confidence interval for the mean height of the orchids.

(4)

A grower claims that the mean height of this type of orchid is 19.5 cm.

(c) Comment on the grower's claim. Give a reason for your answer.

(2)

4

Leave
blank

Question 2 continued

(Total 9 marks)

Q2

5

Turn over

3. A doctor is interested in the relationship between a person's Body Mass Index (BMI) and their level of fitness. She believes that a lower BMI leads to a greater level of fitness. She randomly selects 10 female 18 year-olds and calculates each individual's BMI. The females then run a race and the doctor records their finishing positions. The results are shown in the table.

Individual	A	B	C	D	E	F	G	H	I	J
BMI	17.4	21.4	18.9	24.4	19.4	20.1	22.6	18.4	25.8	28.1
Finishing position	3	5	1	9	6	4	10	2	7	8

(a) Calculate Spearman's rank correlation coefficient for these data.

(5)

(b) Stating your hypotheses clearly and using a one tailed test with a 5% level of significance, interpret your rank correlation coefficient.

(5)

(c) Give a reason to support the use of the rank correlation coefficient rather than the product moment correlation coefficient with these data.

(1)

Leave
blank

Question 3 continued

Q3

(Total 11 marks)

7

Turn over

Leave blank

4. A sample of size 8 is to be taken from a population that is normally distributed with mean 55 and standard deviation 3. Find the probability that the sample mean will be greater than 57.

(5)

Leave blank

Question 4 continued

Q4

(Total 5 marks)

9

Turn over

5. The number of goals scored by a football team is recorded for 100 games. The results are summarised in Table 1 below.

Number of goals	Frequency
0	40
1	33
2	14
3	8
4	5

Table 1

(a) Calculate the mean number of goals scored per game.

(2)

The manager claimed that the number of goals scored per match follows a Poisson distribution. He used the answer in part (a) to calculate the expected frequencies given in Table 2.

Number of goals	Expected Frequency
0	34.994
1	r
2	s
3	6.752
$\geqslant 4$	2.221

Table 2

(b) Find the value of r and the value of s giving your answers to 3 decimal places.

(3)

(c) Stating your hypotheses clearly, use a 5% level of significance to test the manager's claim.

(7)

Leave
blank

Question 5 continued

11

Turn over

Leave blank

Question 5 continued

Leave
blank

Question 5 continued

Q5

(Total 12 marks)

13

Turn over

Leave blank

6. The lengths of a random sample of 120 limpets taken from the upper shore of a beach had a mean of 4.97 cm and a standard deviation of 0.42 cm. The lengths of a second random sample of 150 limpets taken from the lower shore of the same beach had a mean of 5.05 cm and a standard deviation of 0.67 cm.

    (a) Test, using a 5% level of significance, whether or not the mean length of limpets from the upper shore is less than the mean length of limpets from the lower shore. State your hypotheses clearly.

(8)

    (b) State two assumptions you made in carrying out the test in part (a).

(2)

14

Leave
blank

Question 6 continued

Q6

(Total 10 marks)

15

Turn over

Leave blank

7. A company produces climbing ropes. The lengths of the climbing ropes are normally distributed. A random sample of 5 ropes is taken and the length, in metres, of each rope is measured. The results are given below.

    120.3   120.1   120.4   120.2   119.9

    (a) Calculate unbiased estimates for the mean and the variance of the lengths of the climbing ropes produced by the company.

    (5)

    The lengths of climbing rope are known to have a standard deviation of 0.2 m. The company wants to make sure that there is a probability of at least 0.90 that the estimate of the population mean, based on a random sample size of n, lies within 0.05 m of its true value.

    (b) Find the minimum sample size required.

    (6)

16

Leave
blank

Question 7 continued

Q7

(Total 11 marks)

17

Turn over

Leave
blank

8. The random variable A is defined as

$$A = 4X - 3Y$$

where $X \sim N(30, 3^2)$, $Y \sim N(20, 2^2)$ and X and Y are independent.

Find

(a) $E(A)$,

(2)

(b) $Var(A)$.

(3)

The random variables Y_1, Y_2, Y_3 and Y_4 are independent and each has the same distribution as Y. The random variable B is defined as

$$B = \sum_{i=1}^{4} Y_i$$

(c) Find $P(B > A)$.

(6)

Leave
blank

Question 8 continued

Turn over

Leave blank

Question 8 continued

Q8

(Total 11 marks)

TOTAL FOR PAPER: 75 MARKS

END

20

Answer **all** questions.

1   The best performances of a random sample of 20 junior athletes in the long jump, x metres, and in the high jump, y metres, were recorded. The following statistics were calculated from the results.

$$S_{xx} = 7.0036 \quad S_{yy} = 0.8464 \quad S_{xy} = 1.3781$$

(a)   Calculate the value of the product moment correlation coefficient between x and y.

(2 marks)

(b)   Assuming that these data come from a bivariate normal distribution, investigate, at the 1% level of significance, the claim that for junior athletes there is a positive correlation between x and y.   *(4 marks)*

(c)   Interpret your conclusion in the context of this question.   *(1 mark)*

2   A survey of a random sample of 200 passengers on UK internal flights revealed that 132 of them were on business trips.

(a)   Construct an approximate 98% confidence interval for the proportion of passengers on UK internal flights that are on business trips.   *(6 marks)*

(b)   Hence comment on the claim that more than 60 per cent of passengers on UK internal flights are on business trips.   *(2 marks)*

3   Pitted black olives in brine are sold in jars labelled "340 grams net weight". Two machines, A and B, independently fill these jars with olives before the brine is added.

The weight, X grams, of olives delivered by machine A may be modelled by a normal distribution with mean μ_X and standard deviation 4.5.

The weight, Y grams, of olives delivered by machine B may be modelled by a normal distribution with mean μ_Y and standard deviation 5.7.

The mean weight of olives from a random sample of 10 jars filled by machine A is found to be 157 grams, whereas that from a random sample of 15 jars filled by machine B is found to be 162 grams.

Test, at the 1% level of significance, the hypothesis that $\mu_X = \mu_Y$.   *(6 marks)*

4 A manufacturer produces three models of washing machine: basic, standard and deluxe. An analysis of warranty records shows that 25% of faults are on basic machines, 60% are on standard machines and 15% are on deluxe machines.

For basic machines, 30% of faults reported during the warranty period are electrical, 50% are mechanical and 20% are water-related.

For standard machines, 40% of faults reported during the warranty period are electrical, 45% are mechanical and 15% are water-related.

For deluxe machines, 55% of faults reported during the warranty period are electrical, 35% are mechanical and 10% are water-related.

(a) Draw a tree diagram to represent the above information. *(3 marks)*

(b) Hence, or otherwise, determine the probability that a fault reported during the warranty period:

   (i) is electrical; *(2 marks)*

   (ii) is on a deluxe machine, given that it is electrical. *(2 marks)*

(c) A random sample of 10 electrical faults reported during the warranty period is selected. Calculate the probability that exactly 4 of them are on deluxe machines. *(3 marks)*

5 The daily number of emergency calls received from district A may be modelled by a Poisson distribution with a mean of λ_A.

The daily number of emergency calls received from district B may be modelled by a Poisson distribution with a mean of λ_B.

During a period of 184 days, the number of emergency calls received from district A was 3312, whilst the number received from district B was 2760.

(a) Construct an approximate 95% confidence interval for $\lambda_A - \lambda_B$. *(6 marks)*

(b) State one assumption that is necessary in order to construct the confidence interval in part (a). *(1 mark)*

6 An aircraft, based at airport A, flies regularly to and from airport B.

The aircraft's flying time, X minutes, from A to B has a mean of 128 and a variance of 50.

The aircraft's flying time, Y minutes, on the return flight from B to A is such that

$$\text{E}(Y) = 112, \quad \text{Var}(Y) = 50 \quad \text{and} \quad \rho_{XY} = -0.4$$

(a) Given that $F = X + Y$:

    (i) find the mean of F;

    (ii) show that the variance of F is 60.   *(4 marks)*

(b) At airport B, the stopover time, S minutes, is independent of F and has a mean of 75 and a variance of 36.

Find values for the mean and the variance of:

    (i) $T = F + S$;   *(2 marks)*

    (ii) $M = F - 3S$.   *(3 marks)*

(c) Hence, assuming that T and M are normally distributed, determine the probability that, on a particular round trip of the aircraft from A to B and back to A:

    (i) the time from leaving A to returning to A exceeds 300 minutes;   *(3 marks)*

    (ii) the stopover time is greater than one third of the total flying time.   *(6 marks)*

7 (a) The random variable X has a Poisson distribution with $E(X) = \lambda$.

    (i) Prove, from first principles, that $E(X(X-1)) = \lambda^2$.   *(4 marks)*

    (ii) Hence deduce that $\text{Var}(X) = \lambda$.   *(2 marks)*

  (b) The independent Poisson random variables X_1 and X_2 are such that $E(X_1) = 5$ and $E(X_2) = 2$.

    The random variables D and F are defined by

$$D = X_1 - X_2 \quad \text{and} \quad F = 2X_1 + 10$$

    (i) Determine the mean and the variance of D.   *(2 marks)*

    (ii) Determine the mean and the variance of F.   *(3 marks)*

    (iii) For **each** of the variables D and F, give a reason why the distribution is **not** Poisson.   *(2 marks)*

  (c) The daily number of black printer cartridges sold by a shop may be modelled by a Poisson distribution with a mean of 5.

    Independently, the daily number of colour printer cartridges sold by the same shop may be modelled by a Poisson distribution with a mean of 2.

    Use a distributional approximation to estimate the probability that the total number of black and colour printer cartridges sold by the shop during a 4-week period (24 days) exceeds 175.   *(6 marks)*

<div align="center">

END OF QUESTIONS

</div>

Answer **all** questions.

1 An analysis of a random sample of 150 urban dwellings for sale showed that 102 are semi-detached.

An analysis of an independent random sample of 80 rural dwellings for sale showed that 36 are semi-detached.

(a) Construct an approximate 99% confidence interval for the difference between the proportion of urban dwellings for sale that are semi-detached and the proportion of rural dwellings for sale that are semi-detached. *(6 marks)*

(b) Hence comment on the claim that there is no difference between these two proportions. *(2 marks)*

2 A hotel chain has hotels in three types of location: city, coastal and country. The percentages of the chain's reservations for each of these locations are 30, 55 and 15 respectively.

Each of the chain's hotels offers three types of reservation: Bed & Breakfast, Half Board and Full Board.

The percentages of these types of reservation for **each** of the three types of location are shown in the table.

		Type of location		
		City	**Coastal**	**Country**
Type of reservation	**Bed & Breakfast**	80	10	30
	Half Board	15	65	50
	Full Board	5	25	20

For example, 80 per cent of reservations for hotels in city locations are for Bed & Breakfast.

(a) For a reservation selected at random:

(i) show that the probability that it is for Bed & Breakfast is 0.34; *(2 marks)*

(ii) calculate the probability that it is for Half Board in a hotel in a coastal location; *(2 marks)*

(iii) calculate the probability that it is for a hotel in a coastal location, given that it is for Half Board. *(4 marks)*

(b) A random sample of 3 reservations for Half Board is selected.

Calculate the probability that these 3 reservations are for hotels in different types of location. *(5 marks)*

3 The proportion, p, of an island's population with blood type A Rh$^+$ is believed to be approximately 0.35.

A medical organisation, requiring a more accurate estimate, specifies that a 98% confidence interval for p should have a width of at most 0.1.

Calculate, to the nearest 10, an estimate of the minimum sample size necessary in order to achieve the organisation's requirement. *(6 marks)*

4 Holly, a horticultural researcher, believes that the mean height of stems on Tahiti daffodils exceeds that on Jetfire daffodils by more than $15\,\text{cm}$.

She measures the heights, x centimetres, of stems on a random sample of 65 Tahiti daffodils and finds that their mean, $\bar{x}$, is 40.7 and that their standard deviation, s_x, is 3.4.

She also measures the heights, y centimetres, of stems on a random sample of 75 Jetfire daffodils and finds that their mean, $\bar{y}$, is 24.4 and that their standard deviation, s_y, is 2.8.

Investigate, at the 1% level of significance, Holly's belief. *(8 marks)*

5 The random variable X has a binomial distribution with parameters n and p.

(a) Given that

$$\text{E}(X) = np \quad \text{and} \quad \text{E}(X(X-1)) = n(n-1)p^2$$

find an expression for $\text{Var}(X)$. *(3 marks)*

(b) Given that X has a mean of 36 and a standard deviation of 4.8:

(i) find values for n and p; *(3 marks)*

(ii) use a distributional approximation to estimate $\text{P}(30 < X < 40)$. *(4 marks)*

Turn over ▶

6 The table shows the probability distribution for the number of weekday (Monday to Friday) morning newspapers, X, purchased by the Reed household per week.

x	0	1	2	3	4	5
P($X=x$)	0.16	0.15	0.25	0.25	0.15	0.04

(a) Find values for $E(X)$ and $Var(X)$. *(3 marks)*

(b) The number of weekday (Monday to Friday) evening newspapers, Y, purchased by the same household per week is such that

$$E(Y) = 2.0\,, \quad Var(Y) = 1.5 \quad \text{and} \quad Cov(X, Y) = -0.43$$

Find values for the mean and variance of:

(i) $S = X + Y$;

(ii) $D = X - Y$. *(5 marks)*

(c) The total cost per week, *L*, of the Reed household's weekday morning and evening newspapers may be assumed to be normally distributed with a mean of £2.31 and a standard deviation of £0.89.

The total cost per week, *M*, of the household's weekend (Saturday and Sunday) newspapers may be assumed to be independent of *L* and normally distributed with a mean of £2.04 and a standard deviation of £0.43.

Determine the probability that the total cost per week of the Reed household's newspapers is more than £5. *(5 marks)*

7 The daily number of customers visiting a small arts and crafts shop may be modelled by a Poisson distribution with a mean of 24.

  (a) Using a distributional approximation, estimate the probability that there was a total of at most 150 customers visiting the shop during a given 6-day period. *(5 marks)*

  (b) The shop offers a picture framing service. The daily number of requests, Y, for this service may be assumed to have a Poisson distribution.

  Prior to the shop advertising this service in the local free newspaper, the mean value of Y was 2. Following the advertisement, the shop received a total of 17 requests for the service during a period of 5 days.

   (i) Using a Poisson distribution, carry out a test, at the 10% level of significance, to investigate the claim that the advertisement increased the mean daily number of requests for the shop's picture framing service. *(5 marks)*

   (ii) Determine the critical value of Y for your test in part (b)(i). *(3 marks)*

   (iii) Hence, assuming that the advertisement increased the mean value of Y to 3, determine the power of your test in part (b)(i). *(4 marks)*

END OF QUESTIONS