A Level

Mathematics

Session: 2010 June
Type: Question paper
Code: 3890-7890; 3892-7892
Units: 4732, 4733, 4734, 4735

ADVANCED SUBSIDIARY GCE

MATHEMATICS 4732

Probability & Statistics 1

QUESTION PAPER

Candidates answer on the Printed Answer Book

OCR Supplied Materials:
- Printed Answer Book 4732
- List of Formulae (MF1)

Other Materials Required:
- Scientific or graphical calculator

Friday 18 June 2010
Afternoon

Duration: 1 hour 30 minutes

INSTRUCTIONS TO CANDIDATES

These instructions are the same on the Printed Answer Book and the Question Paper.

- Write your name clearly in capital letters, your Centre Number and Candidate Number in the spaces provided on the Printed Answer Book.
- **The questions are on the inserted Question Paper.**
- **Write your answer to each question in the space provided in the Printed Answer Book.** Additional paper may be used if necessary but you must clearly show your Candidate Number, Centre Number and question number(s).
- Use black ink. Pencil may be used for graphs and diagrams only.
- Read each question carefully and make sure that you know what you have to do before starting your answer.
- Answer **all** the questions.
- Do **not** write in the bar codes.
- You are permitted to use a graphical calculator in this paper.
- Give non-exact numerical answers correct to 3 significant figures unless a different degree of accuracy is specified in the question or is clearly appropriate.

INFORMATION FOR CANDIDATES

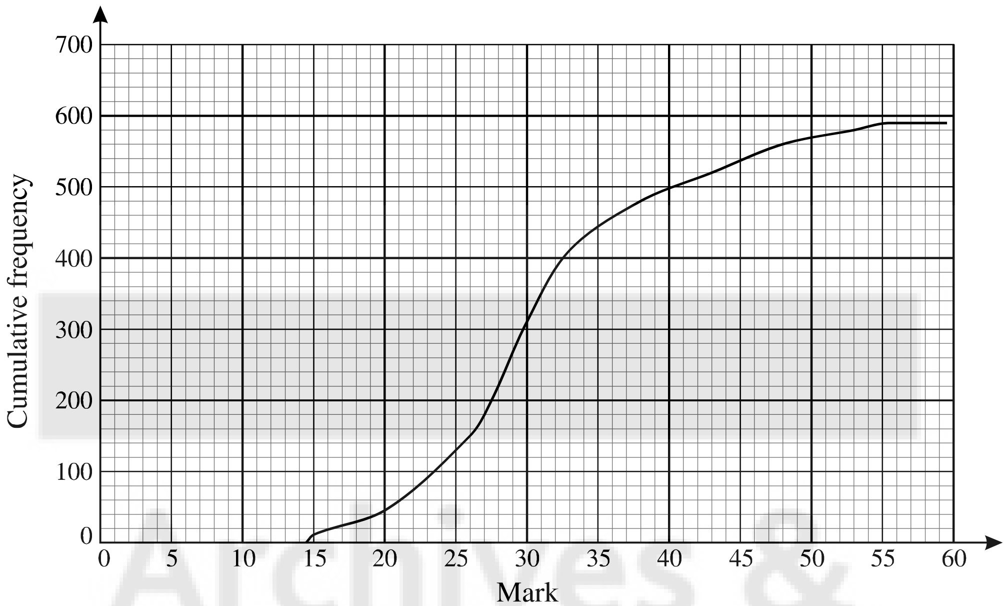
This information is the same on the Printed Answer Book and the Question Paper.

- The number of marks is given in brackets **[ ]** at the end of each question or part question on the Question Paper.
- **You are reminded of the need for clear presentation in your answers.**
- The total number of marks for this paper is **72**.
- The Printed Answer Book consists of **12** pages. The Question Paper consists of **4** pages. Any blank pages are indicated.

INSTRUCTION TO EXAMS OFFICER / INVIGILATOR

- Do not send this Question Paper for marking; it should be retained in the centre or destroyed.

1 The marks of some students in a French examination were summarised in a grouped frequency distribution and a cumulative frequency diagram was drawn, as shown below.



(i) Estimate how many students took the examination. **[1]**

(ii) How can you tell that no student scored more than 55 marks? **[1]**

(iii) Find the greatest possible range of the marks. **[1]**

(iv) The minimum mark for Grade C was 27. The number of students who gained exactly Grade C was the same as the number of students who gained a grade lower than C. Estimate the maximum mark for Grade C. **[3]**

(v) In a German examination the marks of the same students had an interquartile range of 16 marks. What does this result indicate about the performance of the students in the German examination as compared with the French examination? **[3]**

2 Three skaters, A, B and C, are placed in rank order by four judges. Judge P ranks skater A in 1st place, skater B in 2nd place and skater C in 3rd place.

(i) Without carrying out any calculation, state the value of Spearman's rank correlation coefficient for the following ranks. Give a reason for your answer. **[1]**

Skater	A	B	C
Judge P	1	2	3
Judge Q	3	2	1

(ii) Calculate the value of Spearman's rank correlation coefficient for the following ranks. **[3]**

Skater	A	B	C
Judge P	1	2	3
Judge R	3	1	2

(iii) Judge S ranks the skaters at random. Find the probability that the value of Spearman's rank correlation coefficient between the ranks of judge P and judge S is 1. **[3]**

3 **(i)** Some values, (x, y), of a bivariate distribution are plotted on a scatter diagram and a regression line is to be drawn. Explain how to decide whether the regression line of y on x or the regression line of x on y is appropriate. **[2]**

(ii) In an experiment the temperature, x °C, of a rod was gradually increased from 0 °C, and the extension, y mm, was measured nine times at 50 °C intervals. The results are summarised below.

$$n = 9 \qquad \Sigma x = 1800 \qquad \Sigma y = 14.4 \qquad \Sigma x^2 = 510\,000 \qquad \Sigma y^2 = 32.6416 \qquad \Sigma xy = 4080$$

(a) Show that the gradient of the regression line of y on x is 0.008 and find the equation of this line. **[4]**

(b) Use your equation to estimate the temperature when the extension is 2.5 mm. **[1]**

(c) Use your equation to estimate the extension for a temperature of −50 °C. **[1]**

(d) Comment on the meaning and the reliability of your estimate in part **(c)**. **[2]**

4 **(i)** The random variable W has the distribution $\mathrm{B}\left(10, \frac{1}{3}\right)$. Find

(a) $\mathrm{P}(W \leqslant 2)$, **[1]**

(b) $\mathrm{P}(W = 2)$. **[2]**

(ii) The random variable X has the distribution $\mathrm{B}(15, 0.22)$.

(a) Find $\mathrm{P}(X = 4)$. **[2]**

(b) Find $\mathrm{E}(X)$ and $\mathrm{Var}(X)$. **[3]**

5 Each of four cards has a number printed on it as shown.

$$\boxed{1} \qquad \boxed{2} \qquad \boxed{3} \qquad \boxed{3}$$

Two of the cards are chosen at random, without replacement. The random variable X denotes the sum of the numbers on these two cards.

(i) Show that $\mathrm{P}(X = 6) = \frac{1}{6}$ and $\mathrm{P}(X = 4) = \frac{1}{3}$. **[3]**

(ii) Write down all the possible values of X and find the probability distribution of X. **[4]**

(iii) Find $\mathrm{E}(X)$ and $\mathrm{Var}(X)$. **[5]**

6 There are 10 numbers in a list. The first 9 numbers have mean 6 and variance 2. The 10th number is 3. Find the mean and variance of all 10 numbers. **[6]**

[Questions 7 and 8 are printed overleaf.]

7 The menu below shows all the dishes available at a certain restaurant.

Rice dishes	**Main dishes**	**Vegetable dishes**
Boiled rice	Chicken	Mushrooms
Fried rice	Beef	Cauliflower
Pilau rice	Lamb	Spinach
Keema rice	Mixed grill	Lentils
	Prawn	Potatoes
	Vegetarian	

A group of friends decide that they will share a total of 2 different rice dishes, 3 different main dishes and 4 different vegetable dishes from this menu. Given these restrictions,

(i) find the number of possible combinations of dishes that they can choose to share, **[3]**

(ii) assuming that all choices are equally likely, find the probability that they choose boiled rice. **[2]**

The friends decide to add a further restriction as follows. If they choose boiled rice, they will not choose potatoes.

(iii) Find the number of possible combinations of dishes that they can now choose. **[3]**

8 The proportion of people who watch *West Street* on television is 30%. A market researcher interviews people at random in order to contact viewers of *West Street*. Each day she has to contact a certain number of viewers of *West Street*.

(i) Near the end of one day she finds that she needs to contact just one more viewer of *West Street*. Find the probability that the number of further interviews required is

(a) 4, **[3]**

(b) less than 4. **[3]**

(ii) Near the end of another day she finds that she needs to contact just two more viewers of *West Street*. Find the probability that the number of further interviews required is

(a) 5, **[4]**

(b) more than 5. **[2]**

OCR RECOGNISING ACHIEVEMENT

ADVANCED GCE
MATHEMATICS
Probability & Statistics 2

4733

Candidates answer on the Answer Booklet

OCR Supplied Materials:
- 8 page Answer Booklet
- List of Formulae (MF1)

Other Materials Required:
- Scientific or graphical calculator

Tuesday 22 June 2010
Afternoon

Duration: 1 hour 30 minutes

INSTRUCTIONS TO CANDIDATES

- Write your name clearly in capital letters, your Centre Number and Candidate Number in the spaces provided on the Answer Booklet.
- Use black ink. Pencil may be used for graphs and diagrams only.
- Read each question carefully and make sure that you know what you have to do before starting your answer.
- Answer **all** the questions.
- Do **not** write in the bar codes.
- Give non-exact numerical answers correct to 3 significant figures unless a different degree of accuracy is specified in the question or is clearly appropriate.
- You are permitted to use a graphical calculator in this paper.

INFORMATION FOR CANDIDATES

- The number of marks is given in brackets **[ ]** at the end of each question or part question.
- **You are reminded of the need for clear presentation in your answers.**
- The total number of marks for this paper is **72**.
- This document consists of **4** pages. Any blank pages are indicated.

1 **(i)** The number of inhabitants of a village who are selected for jury service in the course of a 10-year period is a random variable with the distribution Po(4.2).

(a) Find the probability that in the course of a 10-year period, at least 7 inhabitants are selected for jury service. **[2]**

(b) Find the probability that in 1 year, exactly 2 inhabitants are selected for jury service. **[3]**

(ii) Explain why the number of inhabitants of the village who contract influenza in 1 year can probably not be well modelled by a Poisson distribution. **[2]**

2 A university has a large number of students, of whom 35% are studying science subjects. A sample of 10 students is obtained by listing all the students, giving each a serial number and selecting by using random numbers.

(i) Find the probability that fewer than 3 of the sample are studying science subjects. **[3]**

(ii) It is required that, in selecting the sample, the same student is not selected twice. Explain whether this requirement invalidates your calculation in part **(i)**. **[2]**

3 Tennis balls are dropped from a standard height, and the height of bounce, H cm, is measured. H is a random variable with the distribution $N(40, \sigma^2)$. It is given that $P(H < 32) = 0.2$.

(i) Find the value of σ. **[3]**

(ii) 90 tennis balls are selected at random. Use an appropriate approximation to find the probability that more than 19 have $H < 32$. **[6]**

4 The proportion of commuters in a town who travel to work by train is 0.4. Following the opening of a new station car park, a random sample of 16 commuters is obtained, and 11 of these travel to work by train. Test at the 1% significance level whether there is evidence of an increase in the proportion of commuters in this town who travel to work by train. **[7]**

5 The time T seconds needed for a computer to be ready to use, from the moment it is switched on, is a normally distributed random variable with standard deviation 5 seconds. The specification of the computer says that the population mean time should be not more than 30 seconds.

(i) A test is carried out, at the 5% significance level, of whether the specification is being met, using the mean $\bar{t}$ of a random sample of 10 times.

(a) Find the critical region for the test, in terms of $\bar{t}$. **[4]**

(b) Given that the population mean time is in fact 35 seconds, find the probability that the test results in a Type II error. **[3]**

(ii) Because of system degradation and memory load, the population mean time μ seconds increases with the number of months of use, m. A formula for μ in terms of m is $\mu = 20 + 0.6m$. Use this formula to find the value of m for which the probability that the test results in rejection of the null hypothesis is 0.5. **[4]**

6 **(a)** The random variable D has the distribution Po(24). Use a suitable approximation to find $P(D > 30)$. **[5]**

(b) An experiment consists of 200 trials. For each trial, the probability that the result is a success is 0.98, independent of all other trials. The total number of successes is denoted by E.

(i) Explain why the distribution of E cannot be well approximated by a Poisson distribution. **[1]**

(ii) By considering the number of failures, use an appropriate Poisson approximation to find $\mathrm{P}(E \leqslant 194)$. **[4]**

7 A machine is designed to make paper with mean thickness 56.80 micrometres. The thicknesses, x micrometres, of a random sample of 300 sheets are summarised by

$$n = 300, \qquad \Sigma x = 17\,085.0, \qquad \Sigma x^2 = 973\,847.0.$$

Test, at the 10% significance level, whether the machine is producing paper of the designed thickness. **[11]**

8 The continuous random variable X has probability density function given by

$$\mathrm{f}(x) = \begin{cases} kx^{-a} & x \geqslant 1, \\ 0 & \text{otherwise,} \end{cases}$$

where k and a are constants and a is greater than 1.

(i) Show that $k = a - 1$. **[3]**

(ii) Find the variance of X in the case $a = 4$. **[5]**

(iii) It is given that $\mathrm{P}(X < 2) = 0.9$. Find the value of a, correct to 3 significant figures. **[4]**

OCR
RECOGNISING ACHIEVEMENT

ADVANCED GCE

MATHEMATICS **4734**

Probability & Statistics 3

Candidates answer on the Answer Booklet

OCR Supplied Materials:
- 8 page Answer Booklet
- List of Formulae (MF1)

Other Materials Required:
- Scientific or graphical calculator

Thursday 24 June 2010
Morning

Duration: 1 hour 30 minutes

INSTRUCTIONS TO CANDIDATES

- Write your name clearly in capital letters, your Centre Number and Candidate Number in the spaces provided on the Answer Booklet.
- Use black ink. Pencil may be used for graphs and diagrams only.
- Read each question carefully and make sure that you know what you have to do before starting your answer.
- Answer **all** the questions.
- Do **not** write in the bar codes.
- Give non-exact numerical answers correct to 3 significant figures unless a different degree of accuracy is specified in the question or is clearly appropriate.
- You are permitted to use a graphical calculator in this paper.

INFORMATION FOR CANDIDATES

- The number of marks is given in brackets [ ] at the end of each question or part question.
- **You are reminded of the need for clear presentation in your answers.**
- The total number of marks for this paper is **72**.
- This document consists of **4** pages. Any blank pages are indicated.

1 The numbers of minor flaws that occur on reels of copper wire and reels of steel wire have Poisson distributions with means 0.21 per metre and 0.24 per metre respectively. One length of 5 m is cut from each reel.

(i) Calculate the probability that the total number of flaws on these two lengths of wire is at least 2. **[4]**

(ii) State one assumption needed in the calculation. **[1]**

2 A coffee machine used in a bar is claimed by the manager to dispense 170 ml of coffee per cup on average. A customer believes that the average amount of coffee dispensed is less than 170 ml. She measures the amount of coffee in 6 randomly chosen cups. The results, in ml, are as follows.

167 171 164 169 168 166

Assuming a relevant normal distribution, test the manager's claim at the 5% significance level. **[7]**

3 The developers of a shopping mall sponsored a study of the shopping habits of its users. Each of a random sample of 100 users was asked whether their weekend shopping was mainly on Saturday or mainly on Sunday. The results, classified according to whether the user lived in the city or the country, are shown in the table.

	City dweller	Country dweller
Saturday shopper	23	19
Sunday shopper	42	16

(i) Test, at the 10% significance level, whether there is an association between the area in which shoppers live and the day on which they shop at the weekend. **[8]**

(ii) State, with a reason, whether the conclusion of the test would be different at the 3% significance level. **[1]**

4 Part of an ecological study involved measuring the heights of trees in a young forest. In order to obtain an estimate of the mean height of all the trees in the forest, a random sample of 70 trees was selected and their heights measured. These heights, x metres, are summarised by $\Sigma x = 246.6$ and $\Sigma x^2 = 1183.65$. The mean height of all trees in the forest is denoted by μ metres.

(i) Calculate a symmetric 90% confidence interval for μ. **[5]**

(ii) A student was asked to interpret the interval and said,

"If 100 independent 90% confidence intervals were calculated then 90 of them would contain μ."

Explain briefly what is wrong with this statement. **[1]**

(iii) Four independent 90% confidence intervals for μ are obtained. Calculate the probability that at least three of the intervals contain μ. **[2]**

5 A random variable X is believed to have (cumulative) distribution function given by

$$F(x) = \begin{cases} 0 & x < 0, \\ 1 - e^{-x^2} & x \geqslant 0. \end{cases}$$

In order to test this, a random sample of 150 observations of X were taken, and their values are summarised in the following grouped frequency table.

Values	$0 \leqslant x < 0.5$	$0.5 \leqslant x < 1$	$1 \leqslant x < 1.5$	$1.5 \leqslant x < 2$	$x \geqslant 2$
Frequency	41	50	32	23	4

The expected frequencies, correct to 1 decimal place, corresponding to the above distribution, are 33.2, 61.6 and 39.4 respectively for the first 3 cells.

(i) Find the expected frequencies for the last 2 cells. **[4]**

(ii) Carry out a goodness of fit test at the $2\frac{1}{2}\%$ significance level. **[6]**

6 It has been suggested that people who suffer anxiety when they are about to undergo surgery can have their anxiety reduced by listening to relaxation tapes. A study was carried out on 18 experimental subjects who listened to relaxation tapes, and 13 control subjects who listened to neutral tapes. After listening to the tapes, the subjects were given a test which produced an anxiety score, X. Higher scores indicated higher anxiety. The results are summarised in the table.

	Sample size	$\bar{x}$	$\Sigma(x - \bar{x})^2$
Experimental subjects	18	32.16	1923.56
Control subjects	13	38.21	1147.58

(i) Use a two-sample t-test, at the 5% significance level, to test whether anxiety is reduced by listening to relaxation tapes. State two necessary assumptions for the validity of your test. **[10]**

(ii) State why a test using a normal distribution would not have been appropriate. **[1]**

7 The employees of a certain company have masses which are normally distributed. Female employees have a mean of 66.7 kg and standard deviation 9.3 kg, and male employees have a mean of 78.3 kg and standard deviation 8.5 kg. It may be assumed that all employees' masses are independent. On the ground floor 6 women and 9 men enter the empty staff lift for which it is stated that the maximum load is 1150 kg.

(i) Calculate the probability that the maximum load is exceeded. **[6]**

At the first floor all 15 passengers leave and 6 women, 8 men and an unknown employee enter.

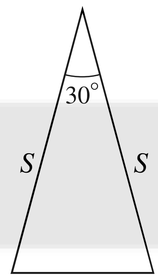
(ii) Assuming that the unknown employee is equally likely to be a woman or a man, calculate the probability that the maximum load is now exceeded. **[6]**

[Question 8 is printed overleaf.]

8 The continuous random variable S has probability density function given by

$$f(s) = \begin{cases} \dfrac{8}{3s^3} & 1 \leqslant s \leqslant 2, \\ 0 & \text{otherwise.} \end{cases}$$

An isosceles triangle has equal sides of length S, and the angle between them is 30° (see diagram).



(i) Find the (cumulative) distribution function of the area X of the triangle, and hence show that the probability density function of X is $\dfrac{1}{3x^2}$ over an interval to be stated. **[7]**

(ii) Find the median value of X. **[3]**

OCR
RECOGNISING ACHIEVEMENT

ADVANCED GCE

MATHEMATICS **4735**

Probability & Statistics 4

Candidates answer on the Answer Booklet

OCR Supplied Materials:
- 8 page Answer Booklet
- List of Formulae (MF1)

Other Materials Required:
- Scientific or graphical calculator

Thursday 24 June 2010
Morning

Duration: 1 hour 30 minutes

INSTRUCTIONS TO CANDIDATES

- Write your name clearly in capital letters, your Centre Number and Candidate Number in the spaces provided on the Answer Booklet.
- Use black ink. Pencil may be used for graphs and diagrams only.
- Read each question carefully and make sure that you know what you have to do before starting your answer.
- Answer **all** the questions.
- Do **not** write in the bar codes.
- Give non-exact numerical answers correct to 3 significant figures unless a different degree of accuracy is specified in the question or is clearly appropriate.
- You are permitted to use a graphical calculator in this paper.

INFORMATION FOR CANDIDATES

- The number of marks is given in brackets **[ ]** at the end of each question or part question.
- **You are reminded of the need for clear presentation in your answers.**
- The total number of marks for this paper is **72**.
- This document consists of **4** pages. Any blank pages are indicated.

1 For the variables A and B, it is given that $\text{Var}(A) = 9$, $\text{Var}(B) = 6$ and $\text{Var}(2A - 3B) = 18$.

(i) Find $\text{Cov}(A, B)$. **[3]**

(ii) State with a reason whether A and B are independent. **[1]**

2 The probability generating function of the discrete random variable X is $\dfrac{e^{4t^2}}{e^4}$. Find

(i) $E(X)$, **[3]**

(ii) $P(X = 2)$. **[3]**

3 X_1 and X_2 are continuous random variables. Random samples of 5 observations of X_1 and 6 observations of X_2 are taken. No two observations are equal. The 11 observations are ranked, lowest first, and the sum of the ranks of the observations of X_1 is denoted by R.

(i) Assuming that all rankings are equally likely, show that $P(R \leqslant 17) = \frac{2}{231}$. **[5]**

The marks of 5 randomly chosen students from School A and 6 randomly chosen students from School B, who took the same examination, achieving different marks, were ranked. The rankings are shown in the table.

Rank	1	2	3	4	5	6	7	8	9	10	11
School	*A*	*A*	*A*	*B*	*A*	*A*	*B*	*B*	*B*	*B*	*B*

(ii) For a Wilcoxon rank-sum test, obtain the exact smallest significance level for which there is evidence of a difference in performance at the two schools. **[2]**

4 The moment generating function of a continuous random variable Y, which has a χ^2 distribution with n degrees of freedom, is $(1 - 2t)^{-\frac{1}{2}n}$, where $0 \leqslant t < \frac{1}{2}$.

(i) Find $E(Y)$ and $\text{Var}(Y)$. **[5]**

For the case $n = 1$, the sum of 60 independent observations of Y is denoted by S.

(ii) Write down the moment generating function of S and hence identify the distribution of S. **[2]**

(iii) Use a normal approximation to estimate $P(S \geqslant 70)$. **[3]**

5 In order to test whether the median salary of employees in a certain industry who had worked for three years was £19 500, the salaries x, in thousands of pounds, of 50 randomly chosen employees were obtained.

(i) The values $|x - 19.5|$ were calculated and ranked. No two values of x were identical and none was equal to 19.5. The sum of the ranks corresponding to positive values of $(x - 19.5)$ was 867. Stating a required assumption, carry out a suitable test at the 5% significance level. **[10]**

(ii) If the assumption you stated in part **(i)** does not hold, what test could have been used? **[1]**

6 Nuts and raisins occur in randomly chosen squares of a particular brand of chocolate. The numbers of nuts and raisins are denoted by N and R respectively and the joint probability distribution of N and R is given by

$$f(n, r) = \begin{cases} c(n+2r) & n = 0, 1, 2 \text{ and } r = 0, 1, 2, \\ 0 & \text{otherwise,} \end{cases}$$

where c is a constant.

(i) Find the value of c. **[3]**

(ii) Find the probability that there is exactly one nut in a randomly chosen square. **[2]**

(iii) Find the probability that the total number of nuts and raisins in a randomly chosen square is more than 2. **[2]**

(iv) For squares in which there are 2 raisins, find the mean number of nuts. **[4]**

(v) Determine whether N and R are independent. **[2]**

7 The continuous random variable X has probability density function given by

$$f(x) = \begin{cases} \dfrac{x}{2\theta^2} & 0 \leqslant x \leqslant 2\theta, \\ 0 & \text{otherwise,} \end{cases}$$

where θ is an unknown positive constant.

(i) Find $E(X^n)$, where $n \neq -2$, and hence write down the value of $E(X)$. **[3]**

(ii) Find

(a) $\text{Var}(X)$,

(b) $\text{Var}(X^2)$.

[4]

(iii) Find $E(X_1 + X_2 + X_3)$ and $E(X_1^2 + X_2^2 + X_3^2)$, where X_1, X_2 and X_3 are independent observations of X. Hence construct unbiased estimators, T_1 and T_2, of θ and $\text{Var}(X)$ respectively, which are based on X_1, X_2 and X_3. **[6]**

(iv) Find $\text{Var}(T_2)$. **[2]**

8 For the events L and M, $P(L \mid M) = 0.2$, $P(M \mid L) = 0.4$ and $P(M) = 0.6$.

(i) Find $P(L)$ and $P(L' \cup M')$. **[3]**

(ii) Given that, for the event N, $P\big(N \mid (L \cap M)\big) = 0.3$, find $P(L' \cup M' \cup N')$. **[3]**