

UNIVERSITY OF CAMBRIDGE INTERNATIONAL EXAMINATIONS
General Certificate of Education Ordinary Level

STATISTICS

4040/01

Paper 1

October/November 2005

2 hours 15 minutes

Additional Materials: Answer Booklet/Paper
Graph paper (2 sheets)
Mathematical tables
Pair of compasses
Protractor

READ THESE INSTRUCTIONS FIRST

If you have been given an Answer Booklet, follow the instructions on the front cover of the Booklet.
Write your Centre number, candidate number and name on all the work you hand in.
Write in dark blue or black pen on both sides of the paper.
You may use a soft pencil for any diagrams or graphs.
Do not use staples, paper clips, highlighters, glue or correction fluid.

Answer **all** questions in Section A and not more than **four** questions from Section B.
Write your answers on the separate Answer Booklet/Paper provided.
All working must be clearly shown.
At the end of the examination, fasten all your work securely together.

The number of marks is given in brackets [] at the end of each question or part question.
The use of an electronic calculator is expected in this paper.

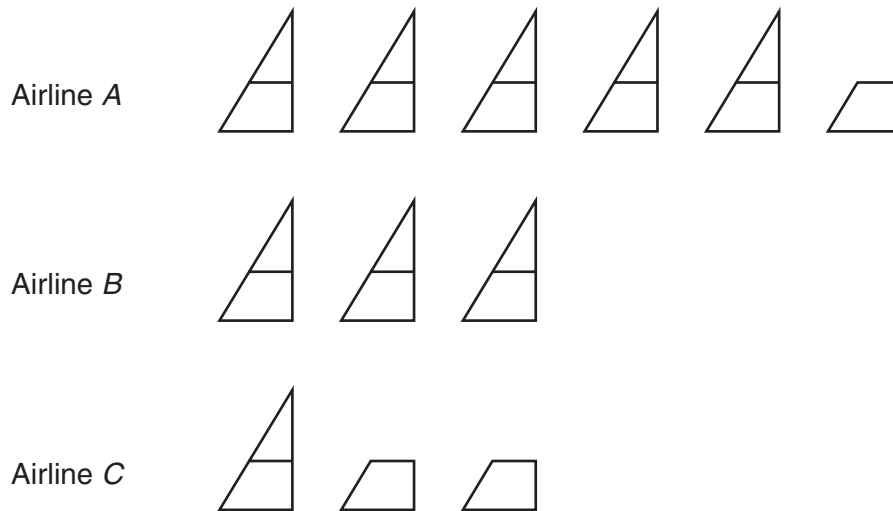
This document consists of **12** printed pages.



Section A [36 marks]

Answer **all** of the questions 1 to 6.


- 1 The pictogram below illustrates the numbers of aircraft operated by three airlines in 2003.



The symbol  represents 40 aircraft.

- (i) Calculate the number of aircraft operated in 2003 by Airline B. [1]

The total number of aircraft operated by the three airlines is 450.
One aircraft is chosen at random from the three airlines.

- (ii) Calculate, as a fraction in its lowest terms, the probability that the aircraft is operated by Airline B. [1]
- (iii) Calculate how many aircraft the symbol  represents. [2]

Two aircraft are chosen at random from the 3 airlines.

- (iv) Calculate the probability that one aircraft is operated by Airline B and one aircraft is operated by Airline C. [3]

2 In a small village there are 240 adults whose ages are classified below.

<i>Age</i>	<i>Men</i>	<i>Women</i>
18 – 35	35	22
36 – 60	45	63
61 or over	26	49

A committee of 20 is to be selected by stratified random sampling to represent the 240 adults.

- (i) Show that the number of men required for the committee from the 18 – 35 category is 3. [1]
- (ii) Construct a table showing the number required for the committee from each category. [2]
- (iii) One person is selected at random from the committee. Calculate the probability that the person selected is a woman from the 36 – 60 category. [1]
- (iv) One person is selected at random from the adults in the village. Calculate the probability that the person selected is a woman from the 36 – 60 category. [1]

3 The marks scored in a test by a group of 10 students were

20 3 8 9 6 6 8 10 6 4.

Three different measures of average can be found from the data above.

- (i) State the name of the statistical measure which is the most common mark.
Give the numerical value of this statistical measure. [2]
- (ii) State the name of the statistical measure which is the mark such that 5 of the students score less than it and 5 of the students score more than it.
Give the numerical value of this statistical measure. [2]
- (iii) State the name of the statistical measure which is calculated from the total marks of all 10 students.
Give the numerical value of this statistical measure. [2]

- 4 The table below shows the percentages in each of four age groups of male full-time students in a certain college in 2000.

<i>Age group</i>	<i>Percentage</i>
18 or under	15.2
19 – 20	35.8
21 – 24	37.8
25 or over	11.2

The data are to be illustrated by a pie chart.

- (i) Calculate, showing your method, the sector angles of the pie chart. Give your answers to the nearest degree. [2]
- (ii) Draw the pie chart using a circle of radius 5 cm. Label each sector with the appropriate age group. [2]

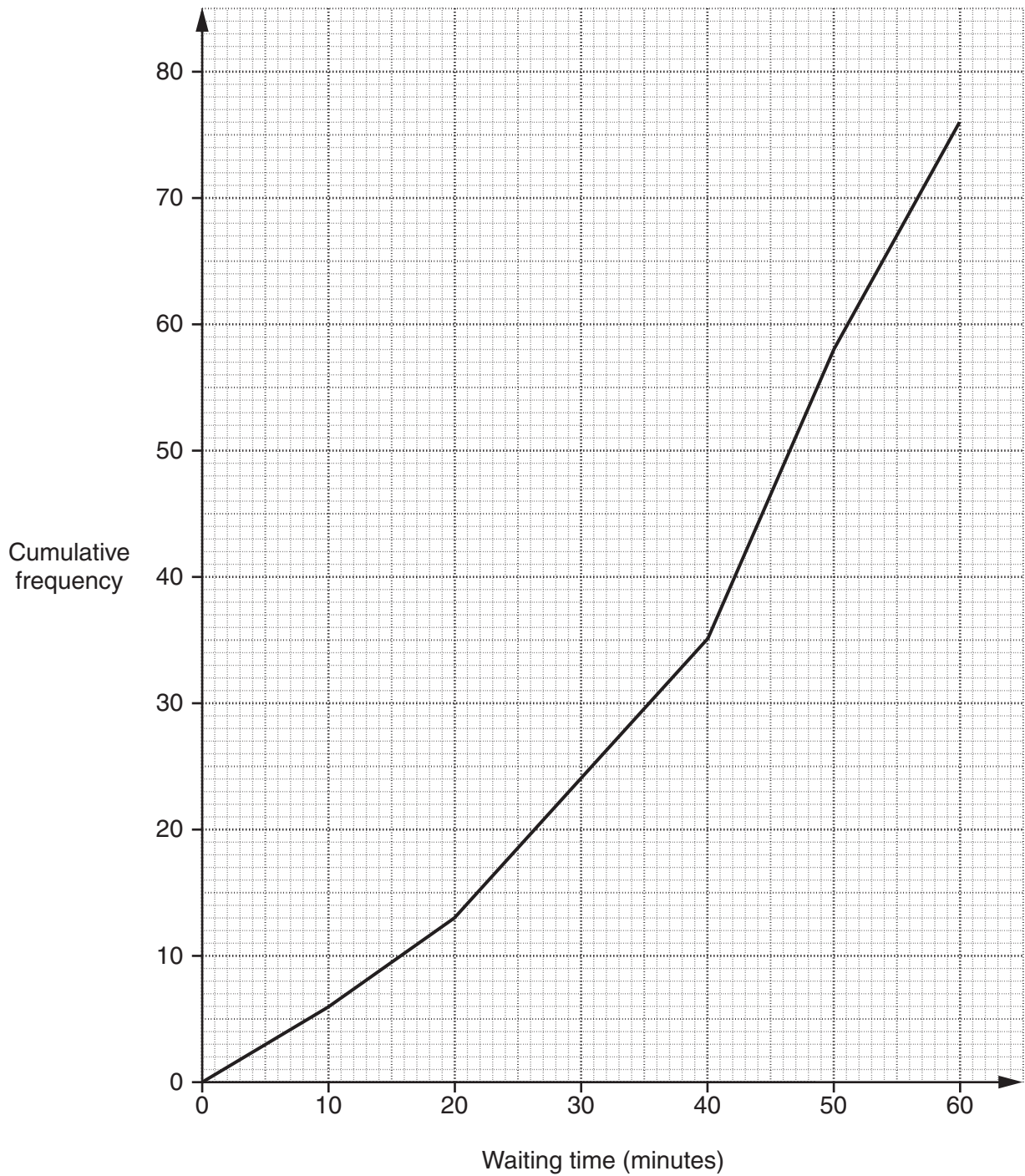
The table below shows the numbers of female full-time students in the same age groups at the same college for 1980 and also for 2000.

<i>Age group</i>	1980	2000
18 or under	42	63
19 – 20	90	158
21 – 24	54	91
25 or over	32	48

A pie chart of radius 6 cm was used to represent the numbers of female full-time students by age in 1980.

- (iii) Calculate, to 1 decimal place, the radius of a comparable pie chart to represent the numbers of female full-time students by age in 2000. [2]

- 5 The cumulative frequency polygon below shows the waiting times, in minutes, of a sample of 76 patients at a doctor's surgery in a certain town.



Use the cumulative frequency polygon to estimate

- (i) the median waiting time, to the nearest minute, [1]
- (ii) the interquartile range of the waiting times, to the nearest minute, [3]
- (iii) the percentage, to 1 decimal place, of patients waiting more than 20 minutes to see the doctor. [3]

- 6 A survey was taken of the main method of travelling to work of 1200 people working for a company.

The table below summarizes the data for this survey.

<i>Type of travel</i>	<i>Male</i>	<i>Female</i>	<i>Total</i>
Car	180	570	750
Bus or train	45	405	450
Totals	225	975	1200

Use the table to calculate the probability that a person chosen at random

- (i) travels to work by car and is a female, [1]
- (ii) travels by car or is a female who does not travel by car, [3]
- (iii) travels by bus or train, given that the person is a male. [1]

Section B [64 marks]

Answer not more than **four** of the questions 7 to 11.

Each question in this section carries 16 marks.

- 7 (a) The table below relates to area A in 1998.

<i>Age group</i>	<i>Population</i>	<i>Number of deaths</i>	<i>Death rate per 1000</i>	<i>Standard Population</i>
0 – 21	3000	45	<i>P</i>	4000
22 – 49	2500	<i>Q</i>	8	3500
50 and over	<i>R</i>	40	20	2500

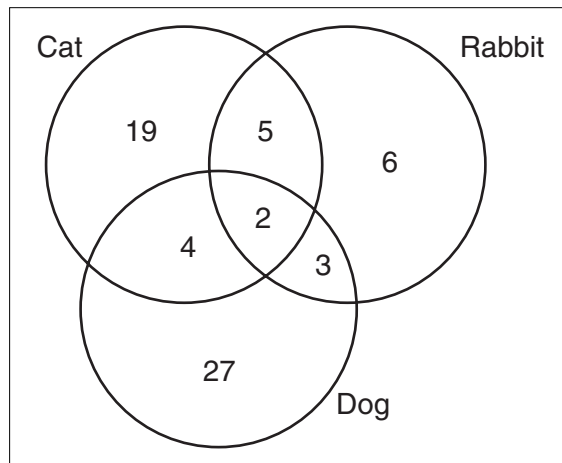
- (i) Calculate each of the numbers represented by the letters *P*, *Q* and *R*. [3]
- (ii) Calculate the crude death rate of area A. [2]
- (iii) Calculate the standardized death rate of area A. [3]

- (b) A survey of 80 households was undertaken to find out whether the household had a dog, a cat or a rabbit as a pet.

None of the households had more than one dog or more than one cat or more than one rabbit.

Some households did not have any pets.

The diagram below summarises **most** of the results of the survey.



For example, 4 households had exactly 1 cat and 1 dog but no rabbit
6 households had only a rabbit.

- (i) Copy and complete the frequency table for the number of pets per household.

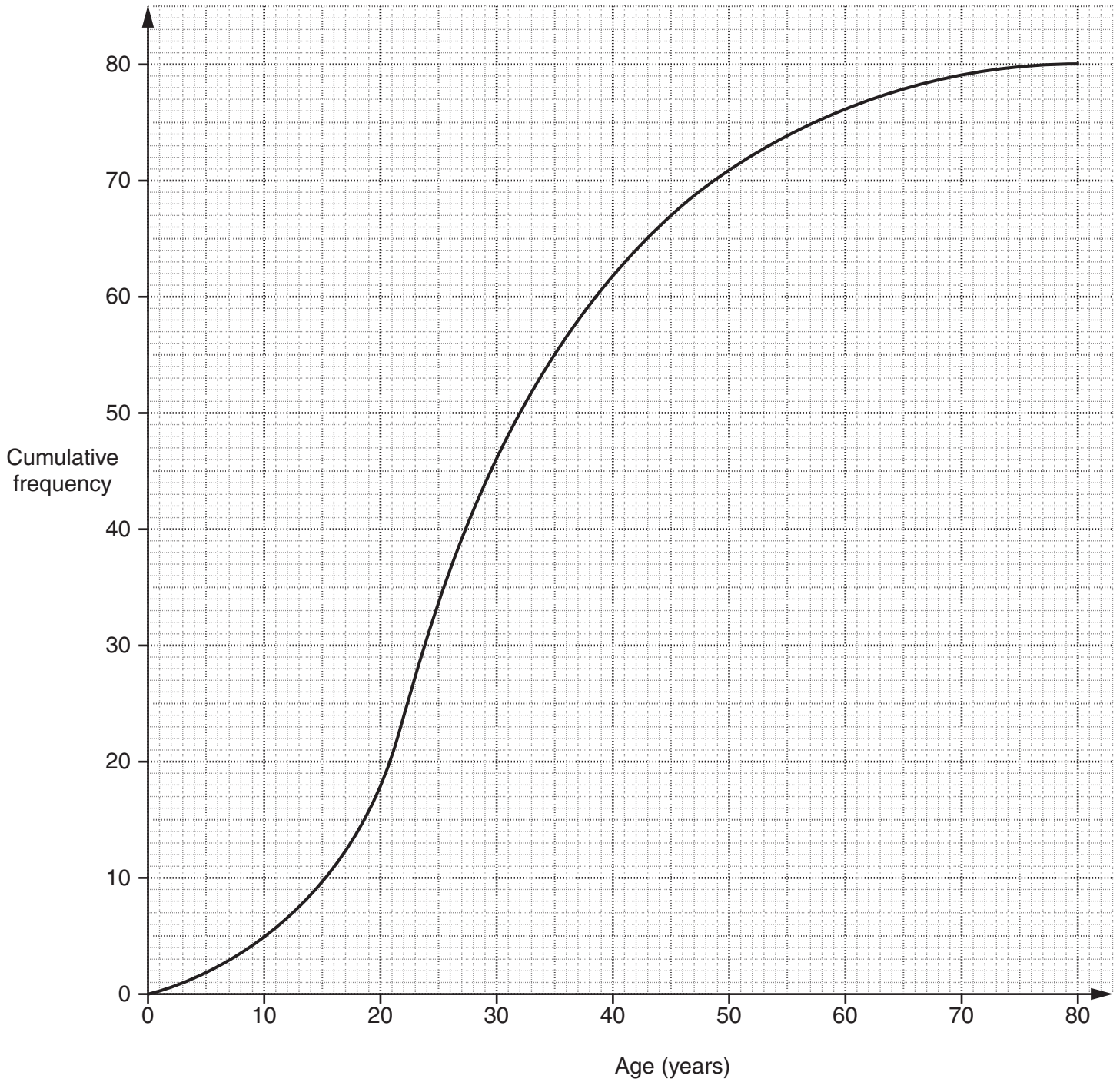
<i>Number of pets</i>	<i>Frequency</i>
0	
1	
2	
3	

[3]

Calculate the probability that a randomly selected household had

- (ii) one pet only, [1]
- (iii) no rabbit, [3]
- (iv) only a dog and a rabbit. [1]

- 8 A survey was taken of the age distribution of 80 people swimming at a pool. The cumulative frequency graph below shows the age distribution of the 80 people.



- (i) Use the cumulative frequency graph to obtain an estimate of

- | | |
|--------------------------|-----|
| (a) the median, | [1] |
| (b) the lower quartile, | [1] |
| (c) the upper quartile, | [1] |
| (d) the 10th percentile, | [1] |
| (e) the 90th percentile. | [1] |

(ii) Use the cumulative frequency graph to estimate the number of people at the pool aged 55 years or over. [3]

(iii) Using the cumulative frequency graph, copy and complete the frequency distribution table below.

<i>Age (y years)</i>	<i>Number of people</i>
$0 \leq y < 15$	
$15 \leq y < 25$	
$25 \leq y < 50$	
$50 \leq y < 80$	

[4]

A person is selected at random from the people at the swimming pool.

(iv) Calculate the probability that this person is aged between 15 years and 25 years. [1]

Exactly 40% of the people were aged Y years or below.

(v) Use the cumulative frequency graph to estimate the value of Y . [2]

A similar survey was taken of the age distribution of people at a garden centre. The survey showed that the median age of the people at the garden centre was 46 years and the interquartile range was 28 years.

(vi) Compare the age distribution of the people at the garden centre with that of the people at the swimming pool. [1]

- 9 The following table shows the annual salaries of the 70 employees of a small manufacturing company.

<i>Salary (\$x)</i>	<i>Number of employees</i>
$10\,000 \leq x < 15\,000$	20
$15\,000 \leq x < 17\,000$	15
$17\,000 \leq x < 20\,000$	15
$20\,000 \leq x < 22\,000$	16
$22\,000 \leq x < 23\,000$	4

- (i) State the median salary of the employees. [1]
- (ii) Using 2 cm to represent \$2000 on the horizontal axis, starting at \$10 000, and a column of height 10 cm to represent the $17\,000 \leq x < 20\,000$ salary group, draw on graph paper a histogram to illustrate the distribution of the salaries of the employees. [5]
- (iii) Use your histogram to identify the modal class for this distribution. [1]
- (iv) Write down the mid-values of the five groups of salaries. [2]
- (v) Calculate, to the nearest \$50, an estimate of the mean salary of the employees. [2]
- (vi) Calculate, to the nearest \$50, an estimate of the standard deviation of the salaries of the employees. [3]
- (vii) If you were the owner of the manufacturing company, which statistical measure of the average salary would you use? Give a reason for your answer. [2]

- 10 The prices in cents of eight items of food purchased from the same shop on 1st August 2000 and 1st August 2001 are given below.

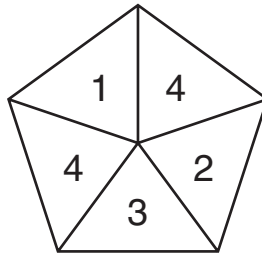
<i>Item</i>	<i>Price 1st August 2000 (x)</i>	<i>Price 1st August 2001 (y)</i>
<i>A</i>	104	108
<i>B</i>	60	66
<i>C</i>	128	144
<i>D</i>	84	88
<i>E</i>	140	148
<i>F</i>	36	42
<i>G</i>	52	54
<i>H</i>	28	30

- (i) Using 2 cm to represent 20 cents on each axis, plot on graph paper a scatter diagram of the price on 1st August 2001 (y) against the price on 1st August 2000 (x). [3]
- (ii) By calculating and plotting (\bar{x}, \bar{y}) and the two semi-averages, draw the line of best fit. [6]
- (iii) Using your graph, state the item which had a greater proportionate price increase than the other seven items between 1st August 2000 and 1st August 2001. [1]

The prices of two other items, K and L , increased similarly to the items in the table.

- (iv) Use your graph to estimate the price of item K on 1st August 2001 if its price on 1st August 2000 was 72 cents. [1]
- (v) Use your graph to estimate the price of item L on 1st August 2000 if its price on 1st August 2001 was 96 cents. [1]
- (vi) Obtain the equation of your line of best fit in the form $y = mx + c$. [3]
- (vii) Give an interpretation for the value of m in your equation. [1]

- 11 (a) The diagram below shows a spinner. The probability that the spinner lands on any of its five outside edges is the same.



The spinner is used three times. Calculate the probability of getting

- (i) a 4 on the first spin only, [3]
- (ii) a 4 on exactly two of the spins. [3]
- (b) Bag *A* contains 3 red discs and 4 blue discs. Bag *B* contains 2 red discs and 5 blue discs.

Hanif takes a disc at random from Bag *A* and puts it into Bag *B*.
Ruth **then** takes a disc at random from Bag *B* and puts it into Bag *A*.

Calculate the probability that Bag *A* now contains 3 red discs. [5]

- (c) In a certain basketball competition all games are either won or lost.

If a team wins a game then the probability that the team wins its next game is $\frac{3}{4}$.

If a team loses a game then the probability of the team losing the next game is $\frac{5}{8}$.

Assuming that a team won their last game, calculate the probability that the team wins exactly one game from its next three games. [5]