1. (a) Explain briefly what is meant by a **discrete random variable**. **(1 mark)**

A family has 3 cats and 4 dogs. Two of the family's animals are to be chosen at random. The random variable X represents the number of dogs chosen.

(b) Copy and complete the table to show the probability distribution of X:

x	0	1	2
$\mathrm{P}(X = x)$			

(4 marks)

(c) Calculate (i) $\mathrm{E}(X)$, (ii) $\mathrm{Var}(X)$, (iii) $\mathrm{Var}(2X)$. **(5 marks)**

2. The discrete random variable X can take any value in the set $\{1, 2, 3, 4, 5, 6, 7, 8\}$.

Arthur, Beatrice and Chris each carry out trials to investigate the distribution of X.

Arthur finds that $\mathrm{P}(X = 1) = 0{\cdot}125$ and that $\mathrm{E}(X) = 4{\cdot}5$.

Beatrice finds that $\mathrm{P}(X = 2) = \mathrm{P}(X = 3) = \mathrm{P}(X = 4) = p$.

Chris finds that the values of X greater than 4 are all equally likely, with each having probability q.

(a) Calculate the values of p and q. **(7 marks)**

(b) Give the name for the distribution of X. **(1 mark)**

(c) Calculate the standard deviation of X. **(3 marks)**

3. The marks obtained by ten students in a Geography test and a History test were as follows:

Student	A	B	C	D	E	F	G	H	I	J
Geography (x)	34	57	49	21	84	53	10	77	61	85
History (y)	40	49	55	40		71	39	47	65	73

(a) Given that $\sum y = 547$, calculate the mark obtained by student E in History. **(1 mark)**

Given further that $\sum x^2 = 34\,087$, $\sum y^2 = 31\,575$ and $\sum xy = 31\,342$, calculate

(b) the product moment correlation coefficient between x and y, **(4 marks)**

(c) an equation of the regression line of y on x, **(4 marks)**

(d) an estimate of the History mark of student K, who scored 70 in Geography. **(2 marks)**

(e) State, with a reason, whether you would expect your answer to part (d) to be reliable. **(2 marks)**

4. The random variable X is normally distributed with mean μ and variance σ^2.

   (a) If $2\mu = 3\sigma$, find $P(X < 2\mu)$. **(5 marks)**

   (b) If, instead, $P(X < 3\mu) = 0{\cdot}86$,

      (i) find μ in terms of σ, **(4 marks)**

      (ii) calculate $P(X > 0)$. **(4 marks)**

5. The stem-and-leaf diagram shows the values taken by two variables A and B.

A		B
8, 7, 4, 1, 0	1	1, 1, 2, 5, 6, 8, 9
9, 8, 7, 6, 6, 5, 2	2	0, 3, 4, 6, 7, 7, 9
9, 7, 6, 4, 2, 1, 0	3	1, 4, 5, 5, 8
8, 6, 3, 2, 2	4	0, 2, 6, 6, 9, 9
6, 4, 0	5	2, 3, 5, 7
5, 3, 1	6	0, 1

Key : 3 | 1 | 2 means $A = 13, B = 12$

   (a) For each set of data, calculate estimates of the median and the quartiles. **(6 marks)**

   (b) Calculate the 42nd percentile for A. **(2 marks)**

   (c) On graph paper, indicating your scale clearly, construct box and whisker plots for both sets of data. **(4 marks)**

   (d) Describe the skewness of the distribution of A and of B. **(2 marks)**

6. The values of the two variables A and B given in the table in Question 5 are written on cards and placed in two separate packs, which are labelled A and B. One card is selected from Pack A. Let A_i represent the event that the first digit on this card is i.

   (a) Write down the value of $P(A_2)$. **(1 mark)**

   The card taken from Pack A is now transferred into Pack B, and one card is picked at random from Pack B. Let B_i represent the event that the first digit on this card is i.

   (b) Show that $P(A_1 \cap B_1) = \frac{1}{24}$. **(3 marks)**

   (c) Show that $P(A_6 \mid B_5) = \frac{4}{41}$. **(5 marks)**

   (d) Find the value of $P(A_1 \cup B_3)$. **(5 marks)**