1. Give one example of a quantity which could be modelled as a discrete random variable. [1]
Name two discrete probability distributions. [2]

2. Spearman's coefficient of rank correlation between two variables $x$ and $y$ is 0·7697 correct to 4 significant figures. If, with the usual notation, $\sum d^2 = 38$, find how many pairs of values $(x, y)$ were used to calculate the coefficient. [4]

3. The students in a large Sixth Form can choose to do exactly one of Community Service, Games or Private Study on Wednesday afternoons. The probabilities that a randomly chosen student does Games and Private Study are $\frac{3}{8}$ and $\frac{1}{5}$ respectively, and it may be assumed that the number of students is large enough for these probabilities to remain constant.

   (i) Find the probability that a randomly chosen student does Community Service. [2]

   (ii) If two students are chosen at random, find the probability that they both do the same activity. [2]

   Two-fifths of the students are girls, and a quarter of these girls do Private Study.

   (iii) Find the probability that a randomly chosen student who does Private Study is a boy. [4]

4. A regular tetrahedron has its faces numbered 1, 2, 3 and 4. It is weighted so that when it is thrown, the probability of each face being in contact with the table is inversely proportional to the number on that face. This number is represented by the random variable $X$.

   (i) Show that $P(X = 1) = \frac{12}{25}$ and find the probabilities of the other values of $X$. [4]

   (ii) Calculate the mean and the variance of $X$. [5]

5. Two variables $x$ and $y$ are such that, for a sample of ten pairs of values,
$$\sum x = 104 \cdot 5, \quad \sum y = 113 \cdot 6, \quad \sum x^2 = 1954 \cdot 1, \quad \sum y^2 = 2100 \cdot 6.$$
The regression line of $x$ on $y$ has gradient 0·8. Find

   (i) $\sum xy$, [3]

   (ii) the equation of the regression line of $y$ on $x$, [4]

   (iii) the product moment correlation coefficient between $y$ and $x$. [3]

   (iv) Describe the kind of correlation indicated by your answer to part (iii). [1]

6. The following table gives the weights, in grams, of 60 items delivered to a company in a day.

| Weight (g) | 0 - 10 | 10 - 20 | 20 - 30 | 30 - 40 | 40 - 50 | 50 - 60 | 60 - 80 |
|---|---|---|---|---|---|---|---|
| No. of items | 2 | 11 | 18 | 12 | 9 | 7 | 1 |

(i) On graph paper, draw a cumulative frequency graph of this data. [4]

(ii) Use your graph to calculate estimated values of (a) the median weight,

(b) the interquartile range, (c) the thirty-third percentile. [4]

The lightest item weighed 3 g and the heaviest weighed 79 g. The mean weight was 32·0 g and the standard deviation of the weights was 14·9 g.

(iii) State, with a reason, whether you would choose to summarise the data by using the mean and standard deviation or the median and interquartile range. [2]

On another day, items were delivered whose weights ranged from 14 g to 58 g; the median was 32 g and the interquartile range was 26 g.

(iv) Briefly compare the two sets of data using this information. [2]

7. On a production line, bags are filled with cement and weighed as they emerge. It is found that 15% of the bags are underweight. In a random sample consisting of $n$ bags, the variance of the number of underweight bags is found to be 2·295.

(i) Show that $n = 18$. [2]

(ii) Use cumulative binomial probability tables to find the probability that, in a further random sample of 18 bags, the number that are underweight is

(a) less than 3, (b) at least 5. [5]

Ten samples of 18 bags are tested. Find the probability that

(iii) all these batches contain less than 5 underweight bags, [3]

(iv) the fourth batch tested is the first to contain less than 5 underweight bags. [3]

# STATISTICS 1 (C) TEST PAPER 10 : ANSWERS AND MARK SCHEME

| | | | |
|---|---|---|---|
| 1. | (i) e.g. score on a die | B1 | |
| | (ii) e.g. geometric and binomial distributions | B1 B1 | 3 |
| 2. | $1 - 6(38)/n(n-1) = 0{\cdot}7697$ $\quad 228/n(n^2-1) = 0{\cdot}2303$ | M1 A1 | |
| | $n^3 - n = 990$ $\quad$ By inspection, $n = 10$ | M1 A1 | 4 |
| 3. | (i) $\text{P(C.S.)} = 1 - \frac{3}{8} - \frac{1}{5} = \frac{17}{40}$ | M1 A1 | |
| | (ii) $\text{P(both do the same)} = \frac{3^2}{8} + \frac{1^2}{5} + \frac{17^2}{40} = 0{\cdot}361$ | M1 A1 | |
| | (iii) Let P(Boy does P.S.) $= p$ $\quad$ By tree diagram or otherwise, | | |
| | $\frac{3}{5}p + \frac{1}{10} = \frac{1}{5}$ $\quad p = \frac{1}{6}$ | M1 A1 | |
| | So $\text{P(Boy} \mid \text{P.S.)} = (\frac{3}{5} \times \frac{1}{6}) \div \frac{1}{5} = \frac{1}{2}$ | M1 A1 | 8 |
| 4. | (i) $\text{P}(X = x) = \frac{k}{x}$ $\quad \frac{k}{1} + \frac{k}{2} + \frac{k}{3} + \frac{k}{4} = 1$ $\quad 25k = 12$ $\quad k = \frac{12}{25}$ | M1 M1 | |
| | so $\text{P}(1) = \frac{12}{25}$ $\quad \text{P}(2) = \frac{6}{25}$ $\quad \text{P}(3) = \frac{4}{25}$ $\quad \text{P}(4) = \frac{3}{25}$ | A1 A1 | |
| | (ii) $\text{E}(X) = \frac{4 \times 12}{25} = 1{\cdot}92$ | M1 A1 | |
| | $\text{E}(X^2) = \frac{12 + 24 + 36 + 48}{25} = 4{\cdot}8$ $\quad \text{Var}(X) = 4.8 - 1.92^2 = 1{\cdot}11$ | M1 A1 A1 | 9 |
| 5. | (i) $S_{yy} = 810{\cdot}104$ $\quad S_{xy} = 0{\cdot}8 \times 810{\cdot}104 = 648{\cdot}0832$ | B1 M1 | |
| | $\sum xy = S_{xy} + (\sum x \sum y)/10 = 1835{\cdot}2$ | A1 | |
| | (ii) $S_{xx} = 862{\cdot}075$ $\quad y - 11{\cdot}36 = \frac{648{\cdot}0832}{862{\cdot}075}(x - 10{\cdot}45)$ | M1 A1 A1 | |
| | $y = 0{\cdot}752x + 3{\cdot}50$ | A1 | |
| | (iii) $r = \sqrt{(0{\cdot}776 \times 0{\cdot}8)} = 0{\cdot}776$ $\quad$ (iv) Moderate positive corr. | M1 A1 A1; B1 | 11 |
| 6. | (i) Cumulative freq.'s 2, 13, 31, 43, 52, 59, 60 $\quad$ Graph | B1 B3 | |
| | (ii) (a) Median = 30th value ≈ 29·5 | B1 | |
| | (b) $Q_1 \approx 21$, $Q_3 \approx 42$, so IQR ≈ 21 | M1 A1 | |
| | (c) Approx. 23·5 | B1 | |
| | (iii) Median and IQR, as they are not affected by the outlier 79 | B1 B1 | |
| | (iv) Second set is slightly higher overall, with wider spread | B1 B1 | 12 |
| 7. | (i) $B(n, 0{\cdot}15)$ : $\text{Var} = n(0{\cdot}15)(0{\cdot}85) = 2{\cdot}295$ $\quad n = 18$ | M1 A1 | |
| | (ii) $X \sim B(18, 0{\cdot}15)$ $\quad$ (a) $\text{P}(X < 3) = \text{P}(X \leq 2) = 0{\cdot}480$ | M1 A1 M1 | |
| | (b) $\text{P}(X \geq 5) = 1 - \text{P}(X \leq 4) = 1 - 0{\cdot}8794 = 0{\cdot}121$ | M1 A1 | |
| | (iii) $X \sim B(10, 0{\cdot}8794)$ : $\text{P}(X = 10) = 0{\cdot}8794^{10} = 0{\cdot}277$ | B1 M1 A1 | |
| | (iv) $0{\cdot}1206^3 \times 0{\cdot}8794 = 0{\cdot}00154$ | M1 A1 A1 | 13 |