1.  A histogram is to be drawn to represent the following grouped continuous data:

| Group | 0 - 10 | 10 - 20 | 20 - 25 | 25 - 30 | 30 - 50 | 50 - 100 |
|-----------|--------|---------|---------|---------|---------|----------|
| Frequency | $2x$ | $3x$ | $5x$ | $6x$ | $2x$ | $x$ |

The '10 - 20' bar has height 6 cm and width 4 cm.  Calculate

(a)  the height of the '20 - 25' bar,                              **(4 marks)**

(b)  the total area under the histogram.                          **(3 marks)**

2.  The events $A$ and $B$ are independent.  Given that $P(A) = 0\cdot4$ and $P(A \cap B) = 0\cdot12$, find

(a)  $P(B)$,                                                       **(2 marks)**

(b)  $P(A \cup B)$,                                               **(2 marks)**

(c)  $P(A' \cap B)$,                                              **(2 marks)**

(d)  $P(A \mid B')$.                                              **(1 mark)**

3.  The random variable $X$ has the discrete uniform distribution over the set of consecutive integers $\{-7, -6, \ldots, 10\}$.

Calculate  (a)  the expectation and variance of $X$,             **(5 marks)**

(b)  $P(X > 7)$,                                                  **(2 marks)**

(c)  the value of $n$ for which $P(-n \le X \le n) = \dfrac{7}{18}$.   **(2 marks)**

4.  The marks, $x$ out of 100, scored by 30 candidates in an examination were as follows:

| | | | | | | | | | |
|----|----|----|----|----|----|----|----|----|-----|
| 5  | 19 | 20 | 21 | 23 | 25 | 31 | 37 | 39 | 41  |
| 42 | 44 | 47 | 51 | 56 | 57 | 60 | 61 | 62 | 65  |
| 67 | 70 | 71 | 73 | 75 | 77 | 81 | 82 | 98 | 100 |

Given that $\sum x = 1600$ and $\sum x^2 = 102\,400$,

(a)  find the median, the mean and the standard deviation of these marks.   **(5 marks)**

The marks were scaled to give modified scores, $y$, using the formula $y = \dfrac{4x}{5} + 20$.

(b)  Find the median, the mean and the standard deviation of the modified scores.   **(4 marks)**

---

Statistics 1 Test Paper 1 Version A

5.  The table shows the numbers of cars and vans in a company's fleet having registrations with the prefix letters shown.

| Registration letter | $K$ | $L$ | $M$ | $N$ | $P$ | $R$ | $S$ | $T$ | $V$ |
|---|---|---|---|---|---|---|---|---|---|
| Number of cars ($x$) | 6 | 7 | 9 | 11 | 15 | 14 | 12 | 10 | 7 |
| Number of vans ($y$) | 8 | 10 | 14 | 13 | 13 | 15 | 14 | 9 | 8 |

(a) Plot a scatter graph of this data, with the number of cars on the horizontal axis and the number of vans on the vertical axis. **(5 marks)**

(b) If there were 4 $J$-registered cars, estimate the number of $J$-registered vans. **(2 marks)**

Given that $\sum x^2 = 1001$, $\sum y^2 = 1264$ and $\sum xy = 1106$,

(c) calculate the product-moment correlation coefficient between $x$ and $y$. Give a brief interpretation of your answer. **(5 marks)**

6.  The distributions of two independent discrete random variables $X$ and $Y$ are given in the tables:

| $x$ | 0 | 1 | 2 |
|---|---|---|---|
| $P(X=x)$ | $\dfrac{3}{5}$ | $\dfrac{3}{10}$ | $\dfrac{1}{10}$ |

| $y$ | 0 | 1 |
|---|---|---|
| $P(Y=y)$ | $\dfrac{5}{8}$ | $\dfrac{3}{8}$ |

The random variable $Z$ is defined to be the sum of one observation from $X$ and one from $Y$.

(a) Tabulate the probability distribution for $Z$. **(5 marks)**

(b) Calculate $E(Z)$. **(3 marks)**

(c) Calculate (i) $E(Z^2)$, (ii) $Var(Z)$. **(4 marks)**

(d) Calculate $Var(3Z - 4)$. **(2 marks)**

7.  The times taken by a large number of people to read a certain book can be modelled by a normal distribution with mean 5·2 hours. It is found that 62·5% of the people took more than 4·5 hours to read the book.

(a) Show that the standard deviation of the times is approximately 2·2 hours. **(5 marks)**

(b) Calculate the percentage of the people who took between 4 and 7 hours to read the book. **(4 marks)**

(c) Calculate the probability that two of the people chosen at random both took less than 5 hours to read the book, stating any assumption that you make. **(4 marks)**

(d) If a number of extra people were taken into account, all of whom took exactly 5·2 hours to read the book, state with reasons what would happen to

(i) the mean, (ii) the variance

and explain briefly why the distribution would no longer be normal. **(4 marks)**

---