STATISTICS (C) UNIT 1 TEST PAPER 5

1. Four pairs of positive bivariate data (x, y) are such that the product moment correlation coefficient between them is r and Spearman's coefficient of rank correlation between them is ρ.
   Sketch scatter diagrams to illustrate each of the following possible situations :
   (i) $r = -1, \rho = -1$, [2]
   (ii) $0 < r < 1, \rho = 1$. [2]

2. A child's toy consists of a board with a large number of equal-sized holes and a set of 25 different coloured pegs. In every set made, some of the pegs are too big for the holes. The number of over-sized pegs in a randomly-chosen set is denoted by X. Given that $P(X \leq 4) = 0{\cdot}902$, use tables of cumulative binomial probabilities to find
   (i) the probability that a randomly-chosen peg in a set is too big, [2]
   (ii) $P(X > 6)$. [2]
   Also find the mean and the variance of X. [3]

3. The discrete random variable X has probability function
   $P(X=x) = cx^2 \quad x = -3, -2, -1, 1, 2, 3$
   (i) Show that $c = {}^1/_{28}$. [3]
   (ii) Calculate (a) $E(X)$, (b) $E(X^2)$, (c) $Var(X)$. [5]

4. Aldith and Bernard play a game. To decide who starts, they use a spinner in the form of a regular pentagon, whose five sectors are numbered 1, 2, 3, 4 and 5

   The first person to score an even number starts the game. Aldith spins first. What assumptions must be made if a geometric distribution with parameter 0·4 is to be used to model the probability that an even number is scored on any particular spin? [2]
   Find the probability that
   (i) Aldith starts the game with her second spin, [2]
   (ii) Bernard starts the game with either his first or second spin, [3]
   (iii) at least four spins are needed before the game is started. [2]

5. The heights, h m, of eight children were measured, giving the following values of h:
   1·20, 1·12, 1·43, 0·98, 1·31, 1·26, 1·02, 1·41.
   (i) Find the mean height of the children. [2]

   (ii) Calculate the standard deviation of the heights. [2]

PMT

The children were also weighed. It was found that their masses, w kg, were such that

$$\Sigma w = 324, \qquad \Sigma w^2 = 13\,532, \qquad \Sigma wh = 403.$$

(iii) Calculate the product-moment correlation coefficient between w and h. [4]

(iv) Comment briefly on the value you have obtained. [1]

6. At a driving test centre, the probabilities of candidates in different age groups passing their theory and practical tests are given by the following table:

	Under 20 years	20 - 25 years	Over 25 years
Practical	0·4	0·5	0·6
Theory	0·7	0·8	x

On a certain morning there are three candidates : Calum, aged 18; Debbie, aged 22; Ed, aged 25.
Use the given figures to find the probability that

(i) all three candidates pass the practical test, [1]

(ii) exactly two of the three pass the practical test, [2]

(iii) Calum passes one component of the test but fails the other. [2]

State any assumptions you have made, and discuss their validity. [2]

If the probability that exactly two of these three candidates pass the theory test is 0·425,

(iv) find the value of x. [4]

7. The heights of 110 applicants to the fire service ranged from 155 cm to 190 cm.

The distribution of these heights is shown in the grouped frequency table :

Height (cm)	155 - 160	160 - 165	165 - 170	170 - 175	175 - 180	180 - 185	185 - 190
Frequency	7	12	18	21	25	17	10

(i) On graph paper, construct a cumulative frequency graph to illustrate this data. [4]

(ii) Use your graph to estimate the median height and the interquartile range. [3]

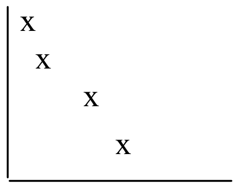
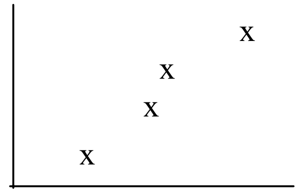
(iii) On graph paper, draw a box-and-whisker plot for this data. Show your scale. [3]

The box-and-whisker plot for another set of applicants is shown below :

150 155 160 165 170 175 180 185 190 195 200 Height (cm)

(iv) Use the box plots to compare the two sets of data briefly. [2]

STATISTICS 1 (C) TEST PAPER 5 : ANSWERS AND MARK SCHEME

1. (i)  (ii)  B2 B2 4

2. (i) $P(X \le 4) = 0{\cdot}902$ in $B(25, p)$, so from tables $p = 0{\cdot}1$ M1 A1
   (ii) $P(X > 6) = 1 - P(X \le 6) = 1 - 0{\cdot}9905 = 0{\cdot}0095$ M1 A1
   $E(X) = 2{\cdot}5$ $\quad$ $Var(X) = 2{\cdot}5 \times 0{\cdot}9 = 2{\cdot}25$ B1 M1 A1 7

3. (i) $c(9 + 4 + 1 + 0 + 1 + 4 + 9) = 1$ $\quad$ $c = {}^{1}/_{28}$ M1 A1 A1
   (ii) (a) $E(X) = 0$ (b) $E(X^2) = (81 + 16 + 1 + 1 + 16 + 81)/28 = 7$ B1 M1 A1
   (c) $Var(X) = 7 - 0 = 7$ M1 A1 8

4. All scores are equally likely; successive trials are independent B1 B1
   (i) $\frac{3}{5}\times\frac{3}{5}\times\frac{2}{5} = \frac{18}{125} = 0.144$ (ii) $\frac{6}{25} + \frac{54}{625} = \frac{204}{625} = 0{\cdot}326$ M1 A1; M1 A1 A1
   (iii) $\left(\frac{3}{5}\right)^4 = \frac{81}{625} = 0{\cdot}130$ M1 A1 9

5. (i) $\Sigma h = 9{\cdot}73$ $\quad$ $9{\cdot}73 \div 8 = 1{\cdot}22$ m M1 A1
   (ii) $\Sigma h^2 = 12{\cdot}0319$ $\quad$ Var. $= 12{\cdot}0319 \div 8 - 1{\cdot}21625^2 = 0{\cdot}0247$ M1 A1
   (iii) $S_{hh} = 0{\cdot}1978$, $S_{ww} = 410$, $S_{hw} = 8{\cdot}935$ $\quad$ $r = 0{\cdot}992$ M1 A1 A1 A1
   (iv) Shows strong positive correlation B1 9

6. (i) $0{\cdot}4 \times 0{\cdot}5 \times 0{\cdot}6 = 0{\cdot}12$ B1
   (ii) $0{\cdot}4 \times 0{\cdot}5 \times 0{\cdot}4 + 0{\cdot}6 \times 0{\cdot}5 \times 0{\cdot}6 + 0{\cdot}4 \times 0{\cdot}5 \times 0{\cdot}6 = 0{\cdot}38$ M1 A1
   (iii) $0{\cdot}4 \times 0{\cdot}3 + 0{\cdot}6 \times 0{\cdot}7 = 0{\cdot}54$ M1 A1
   Assumed : different candidates' results independent : probably so B1
   and results in theory and practical independent : probably not so B1
   (iv) $0{\cdot}3 \times 0{\cdot}8 \times x + 0{\cdot}7 \times 0{\cdot}2 \times x + 0{\cdot}7 \times 0{\cdot}8 \times (1 - x) = 0{\cdot}425$ M1 A1
   $0{\cdot}56 - 0{\cdot}18x = 0{\cdot}425$ $\quad$ $x = 0{\cdot}75$ M1 A1 11

7. (i) Cum. freqs. 7, 19, 37, 58, 83, 100, 110 B1
   Graph drawn B3
   (ii) Median $\approx 174{\cdot}5$ $\quad$ $Q1 \approx 167{\cdot}5$, $Q3 \approx 179{\cdot}5$, so IQR = 12 B1 M1 A1
   (iii) Box plot drawn, with scale shown B3
   (iv) First set taller on average, with more consistent heights B1 B1 12