STATISTICS (C) UNIT 1 TEST PAPER 3

1. A die is biased so that the probability of scoring a 6 is $^1/_5$.
   (i) Find the probability that the die is thrown at least 10 times without scoring a 6 before the first 6 is scored. [2]
   (ii) Write down an expression for the probability that the first 6 is scored on the nth throw. [2]

2. The mean mark obtained by a class of 20 pupils in a test was μ. The marks, x, were such that $\Sigma (x - \mu)^2 = 7578{\cdot}2$.
   (i) Find the standard deviation of the pupils' marks. [2]
   (ii) Given further that $\mu = 58{\cdot}7$, find the value of Σx^2. [3]

3. The discrete random variable X has the probability function given by the following table:

x	0	1	2	3	4	5	6
P($X = x$)	0·09	0·12	0·22	0·16	p	$2p$	0·2

   (i) Show that $p = 0{\cdot}07$ [2]
   (ii) Find the values of E(X) and Var(X). [5]

4. Of the students who enrol on a university course, 15% drop out within the first year and a further 5% of the original total drop out before the end of the course. One year, 25 students enrol. Find the probability that there are still more than 20 students on the course
   (i) after one year, [2]
   (ii) at the end of the course. [2]

   State an assumption that you have made. Is this assumption likely to be valid? [2]
   10 of the students are mature students. The probability that at least 8 of these will still be on the course after a year is calculated to be 0·3, to 1 decimal place.
   Find the proportion of mature students who drop out during the first year. [4]

5. Sixteen cards have been lost from a pack, which therefore contains only 36 cards. Two cards are drawn at random from the pack. The probability that **both** cards are red is $\frac{1}{3}$.
   (i) Show that r, the number of red cards in the pack, satisfies the equation
   $$r(r - 1) = 420.$$ [3]
   (ii) Hence or otherwise find the value of r. [3]
   (iii) Find the conditional probability that, when two cards are drawn at random from the pack, the first one is red given that at least one is red. [4]

6. Twenty pairs of observations are made of two variables x and y, which are believed to be related. It is found that

$$\Sigma x = 200, \quad \Sigma y = 174, \quad \Sigma x^2 = 6201, \quad \Sigma y^2 = 5102, \quad \Sigma xy = 5200.$$

Find

(i) the product-moment correlation coefficient between x and y, [2]

(ii) the equation of the regression line of y on x. [3]

Given that $p = x + 30$ and $q = y + 50$,

(iii) write down the product-moment correlation coefficient between p and q. [1]

(iv) Find the equation of the regression line of q on p, in the form $q = mp + c$. [3]

(v) Estimate the value of q when $p = 46$, stating any assumptions you make. [3]

7. In a survey of natural habitats, the numbers of trees in sixty equal areas of land were recorded, as follows:

17	12	9	23	40	32	11	5	34	22	31	8
15	45	10	52	14	13	29	43	69	30	15	47
35	6	24	13	19	26	9	31	27	18	6	20
22	18	30	51	49	35	50	25	8	10	26	31
33	29	40	37	38	44	24	34	42	38	11	23

(i) Construct a stem-and-leaf diagram to illustrate this data, using the groupings 5 - 9, 10 - 14, 15 - 19, 20 - 24, etc. [6]

(ii) Estimate the median of the distribution. [2]

The first and third quartiles of the distribution are 14 and 35 respectively.

(iii) On graph paper construct a box-and-whisker plot for the data, showing your scale and clearly indicating any outliers. [4]

STATISTICS 1 (C) TEST PAPER 3 : ANSWERS AND MARK SCHEME

1. (i) $(0.8)^{10} = 0{\cdot}107$ (ii) $(0.8)^{n-1}(0.2)$ M1 A1; M1 A1 4

2. (i) s.d. $= \sqrt{(7578{\cdot}2 \div 20) = 19{\cdot}5}$ M1 A1
   (ii) $\Sigma x^2 \div 20 - 58{\cdot}7^2 = 19{\cdot}466^2$ $\Sigma x^2 = 76\,492$ M1 A1 A1 5

3. (i) $3p + 0{\cdot}79 = 1$ $3p = 0{\cdot}21$ $p = 0{\cdot}07$ M1 A1
   (ii) $E(X) = 3{\cdot}22$ M1 A1
   $\text{Var}(X) = E(X^2) - 3{\cdot}22^2 = 14{\cdot}26 - 3{\cdot}22^2 = 3{\cdot}892$ M1 A1 A1 7

4. (i) X~B(25, 0·85) : $P(X > 20) = 1 - P(X \le 20) = 1 - 0{\cdot}3179 = 0{\cdot}682$ M1 A1
   (ii) X~B(25, 0·8) : $P(X > 20) = 1 - P(X \le 20) = 1 - 0{\cdot}5793 = 0{\cdot}421$ M1 A1
   Assumed each student's leaving is independent of all others B1
   May not be true - some may influence others B1
   B(10, p) : $P(X \ge 8) = 0{\cdot}3$ so $P(X \le 7) = 0{\cdot}7$ M1 A1 A1
   From tables, $p = {}^2/_3$ so ${}^1/_3$ drop out in first year A1 10

5. (i) $\frac{r}{36} \times \frac{r-1}{35} = \frac{1}{3}$ $r(r-1) = 12 \times 35 = 420$ M1 A1 A1
   (ii) $r^2 - r - 420 = 0$ $(r - 21)(r + 20) = 0$ $r = 21$ M1 M1 A1
   (iii) P(at least one red) $= 1 - {}^5/_{12} \times {}^{14}/_{35} = {}^5/_6$ M1 A1
   $P(\text{1st red} \mid \text{at least one red}) = \frac{\frac{7}{12}}{\frac{5}{6}} = \frac{7}{10}$ M1 A1 10

6. (i) $S_{xx} = 4201$, $S_{yy} = 3588{\cdot}2$, $S_{xy} = 3460$ $r = 0{\cdot}891$ M1 A1 A1
   (ii) $y - 8{\cdot}7 = \frac{3460}{4201}(x - 10)$ $y = 0{\cdot}824x + 0{\cdot}464$ M1 M1 A1
   (iii) Same as between x and y : 0·891 B1
   (iv) $q - 50 = 0{\cdot}824(p - 30) + 0{\cdot}464$ $q = 0{\cdot}824p + 25{\cdot}8$ M1 A1 A1
   (v) When $p = 46$, $q \approx 63{\cdot}6$ Assumed these values of p and q are M1 A1
   within or close to the range from which the data was collected B1 12

7. (i) Stem-and-leaf diagram drawn M2 A1 any one group
   Totals in groups : 7, 8, 6, 7, 6, 9, 5, 5, 3, 3, 0, 0, 1 A1 A1 A1
   (ii) Median= 30th - 31st value = 26 M1 A1
   (iii) Box plot drawn Outlier at 69 B3 B1 12