STATISTICS (C) UNIT 1 TEST PAPER 2

1. 15 observations are made of a variable X. Taking 10 as an assumed value for the mean, it is found that $\Sigma (x - 10) = 30$ and $\Sigma(x - 10)^2 = 130$.
   Find the mean and the standard deviation of the 15 values of X. [6]

2. The integers from 1 to 9 inclusive are written on nine cards, which are shuffled together. Three cards are then drawn at random. Find the probability that the numbers on these three cards are
   (i) consecutive numbers, [3]
   (ii) all even numbers. [3]

3. The performance figures for seven cars are summarised in the table :

Car	1	2	3	4	5	6	7
Maximum speed (km h^{-1})	120	200	235	195	150	180	210
Fuel consumption (l / 100 km)	4·33	7·52	9·38	7·95	4·88	5·35	6·67

   Calculate Spearman's coefficient of rank correlation for this data. [4]
   The data for an eighth car was subsequently obtained. When this data was taken into account, the coefficient of rank correlation remained unchanged. Find the new value of Σd^2, where d denotes the difference between the ranks of corresponding pairs of data items. [2]

4. Explain briefly what is meant by a **discrete random variable**. [1]
   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.
   (i) Show that $P(X = 1) = {}^4/_7$ and complete the table to show the probability distribution of X:

x	0	1	2
$P(X = x)$		${}^4/_7$	

   [4]
   (ii) Calculate $E(X)$ and $Var(X)$. [4]

5. A hurdler estimates that she has a 0·95 probability of clearing each hurdle in a race.
   Specify fully the distribution of
   (a) X, the random variable representing the number of hurdles attempted up to and including the first failure, assuming that she continues until a failure occurs; [1]
   (b) Y, the random variable representing the number of hurdles cleared altogether in a race with n hurdles. [1]

Q.5 continued on next page …

5. continued …

Find the probability that, in a race with 20 hurdles,

(i) she clears all 20, [2]

(ii) the 15th hurdle is the first one she fails to clear, [2]

(iii) she clears at least 18. [3]

6. 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

(i) Given that $\Sigma y = 547$, calculate the mark obtained by student *E* in History. [1]

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

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

(iii) an equation of the regression line of y on x,

[3]

(iv) an estimate of the History mark of student *K*, who scored 70 in Geography. [2]

(v) State, with a reason, whether you would expect your answer to part (iv) to be reliable. [2]

7. The stem-and-leaf diagram shows the values taken by a variable *X*.

Key : 1 | 2 means 12

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

(i) Copy and complete the cumulative frequency table :

x	< 19	< 29	< 39	< 49	< 59	< 69
Frequency	7					

[2]

(ii) On graph paper, indicating the scale clearly, construct a cumulative frequency graph for this data. [4]

(iii) Use your graph to find estimates of the median, the quartiles and the 42nd percentile. Make your method clear. [5]

(iv) On graph paper, showing your scale, construct a box-and-whisker plot for the data. [3]

STATISTICS 1 (C) TEST PAPER 2 : ANSWERS AND MARK SCHEME

1. $\Sigma (x - 10) = 30$ Mean = $^{30}/_{15} + 10 = 2+10 = 12$ M1 A1
   $\Sigma (x - 10)^2 = 130$ Var = $^{130}/_{15} - 2^2 = 4.67$ M1 A1
   s.d. = 2·16 M1 A1 6

2. (i) No. of ways of drawing 3 cards = 9C3 = 84 B1
   No. of consecutive triads = 7, ie 123,234,345,456,567,678,789
   so P(consecutive) = $\frac{1}{12}$ M1 A1
   (ii) No. of even triads = 4, so P(all even) = $^{1}/_{21}$ M1 A1 A1 6

3. | Ranks | 1 | 5 | 7 | 4 | 2 | 3 | 6 | |
   |---|---|---|---|---|---|---|---|---|
   | | 1 | 5 | 7 | 6 | 2 | 3 | 4 | B1 |
   | d | 0 | 0 | 0 | 2 | 0 | 0 | 2 | B1 |

   $\Sigma d^2 = 8$ $r_s = 1 - \frac{48}{7\times 48} = \frac{6}{7}$ M1 A1
   Now $1 - 6\Sigma d^2 \div (8 \times 63) = \frac{6}{7}$ $\Sigma d^2 = 12$ M1 A1 6

4. A quantity which can take only certain distinct values with
   fixed probabilities B1
   (i) 2 animals : $7 \times 6 \div 2 = 21$ ways
   1 dog, 1 cat : $4 \times 3 = 12$ ways

x	0	1	2
P($X = x$)	$^{1}/_{7}$	$^{4}/_{7}$	$^{2}/_{7}$

   M1 M1 A1 A1
   (ii) E(X) = $^{8}/_{7}$ E(X^2) = $^{12}/_{7}$ Var (X) = $^{12}/_{7} - {}^{64}/_{49} = {}^{20}/_{49}$ B1 B1 M1 A1 9

5. (a) Geometric : $X \sim$ Geo(0·05) (b) Binomial : $Y \sim$ B(n, 0·95) B1 B1
   (i) $0{\cdot}95^{20} = 0{\cdot}358$ (ii) $0{\cdot}95^{14} \times 0{\cdot}05 = 0{\cdot}0244$ M1 A1; M1 A1
   (iii) P($Y > 18$) = 1 – P($Y < 17$) = 1 – 0·0755 = 0·925 M1 A1 A1 9

6. (i) 547 – 479 = 68 B1
   (ii) $\Sigma x = 531$
   $S_{xx} = 5890{\cdot}8$, $S_{yy} = 1654{\cdot}1$, $S_{xy} = 2296{\cdot}3$ $r = 0{\cdot}736$ M1 A1
   (iii) $y - 54{\cdot}7 = (2296{\cdot}3/5890{\cdot}8)(x - 53{\cdot}1) = 0{\cdot}3898x - 20{\cdot}699$ M1 A1
   $y = 0{\cdot}390x + 34{\cdot}0$ A1
   (iv) When $x = 70$, $y \approx 61{\cdot}3$ M1 A1
   (v) Nor very reliable, as value of r shows only moderate correlation B1 B1 10

7. (i) 7, 14, 19, 25, 29, 31 (ii) Graph drawn B2; B4
   (iii) Appropriate lines drawn Median≈ 34, $Q_1 \approx 20$, $Q_3 \approx 49$ M1 A1 A1 A1
   0·42 x 31 = 13·0, so 13th value, ≈27 A1
   (iv) Box plot drawn B3 14