STATISTICS (C) UNIT 1 **TEST PAPER 9**

1. Briefly describe what is meant by

   (i) a statistical model, [2]

   (ii) a refinement of a model. [2]

2. Every week a social club holds a raffle in which each ticket has a probability of 0·05 that it will win a prize. Marjory decides to buy one ticket every week until she wins a prize, after which she will stop taking part. Find the probability that

   (i) she wins a prize in the third week, [3]

   (ii) she is still taking part in the fifth week. [2]

3. The random variable X has the binomial distribution B(n, p).

   Given that E(X) = k Var(X), find an expression for p in terms of k. [3]

   If n = 8 and k = 4, find

   (i) P(X = 5), (ii) P(X > 6). [5]

4. The length of time, in minutes, that visitors queued for a tourist attraction is given by the following table, where, for example, '20 - ' means from 20 up to but not including 30 minutes.

Queuing time (mins)	0 -	10 -	15 -	20 -	30 -	40 - 60
Number of visitors	15	24	x	13	10	y

   A histogram is drawn to represent this data. The total area under the histogram is 36 cm^2. The '10 - ' bar has width 1 cm and height 9·6 cm. The '15 - ' bar is ten times as high as the '40 - 60' bar.

   (i) Find the values of x and y. [6]

   (ii) On graph paper, construct the histogram accurately. [4]

5. The discrete random variable X takes only the values 4, 5, 6, 7, 8 and 9. The probabilities of these values are given in the table:

x	4	5	6	7	8	9
P(X = x)	p	0·1	q	q	0·3	0·2

   It is known that E(X) = 6·7. Find

   (i) the values of p and q, [6]

   (ii) P(X < 7), [2]

   (iii) Var(X). [3]

6. Among the families with two children in a large city, the probability that the elder child is a boy is $\frac{5}{12}$ and the probability that the younger child is a boy is $\frac{9}{16}$. The probability that the younger child is a girl, given that the elder child is a girl, is $\frac{1}{4}$.

One of the families is chosen at random. Using a tree diagram, or otherwise,

(i) show that the probability that both children are boys is $\frac{1}{8}$. [4]

(ii) Find the probability that one child is a boy and the other is a girl. [3]

If three of the families are chosen at random,

(iii) find the probability that exactly two of the families have two boys. [3]

(iv) State an assumption that you have made in answering part (iii). [1]

7. The marks out of 75 obtained by a group of ten students in their first and second Statistics modules were as follows:

Student	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*
Module 1 (x)	54	33	42	71	60	27	39	46	59	64
Module 2 (y)	50	22	44	58	42	19	35	46	55	60

(i) Find $\sum x$ and $\sum y$. [2]

Given that $\sum x^2 = 26\,353$ and $\sum xy = 22\,991$,

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

(iii) Estimate the Module 2 result of a student whose mark in Module 1 was (a) 65, (b) 5. Explain why one of these estimates is less reliable than the other. [3]

The equation of the regression line of x on y is $x = 0{\cdot}921y + 9{\cdot}81$.

(iv) Deduce the product moment correlation coefficient between x and y, and briefly interpret its value. [3]

STATISTICS 1 (C) TEST PAPER 9 : ANSWERS AND MARK SCHEME

1. (i) A mathematical representation which uses probabilities to describe and predict the behaviour of a real-life situation B2
   (ii) An improved mathematical formulation of the problem which aims to represent the reality more closely. B2 4

2. (i) P(wins in third week) = 0·95 x 0·95 x 0·05 = 0·045 M1 A1 A1
   (ii) $0{\cdot}95^4 = 0{\cdot}815$ M1 A1 5

3. $np = knp(1-p)$ $k(1-p) = 1$ $p = 1 - \frac{1}{k}$ M1 A1 A1
   (i) $p = 0{\cdot}75$ $P(X \le 5) - P(X \le 4) = 0{\cdot}3215 - 0{\cdot}1138 = 0{\cdot}208$ B1 M1 A1
   (ii) $P(X > 6) = 1 - P(X \le 6) = 1 - 0{\cdot}6329 = 0{\cdot}367$ M1 A1 8

4. (i) 10 - 15 has area 9·6 cm^2, so 2.5 visitors : 1 cm^2 B1 M1
   and 36 cm^2 = 90 visitors $62 + x + y = 90$ $x + y = 28$ M1 A1
   Also $\frac{x}{2{\cdot}5} = 10 \times \frac{y}{10}$ so $x = 2{\cdot}5y$ Hence $x = 20$, $y = 8$ M1 A1
   (ii) Freq. densities 1·5, 4·8, 4, 1·3, 1, 0·4 Histogram drawn B1 B3 10

5. (i) $4p + 13q + 4{\cdot}7 = 6{\cdot}7$ $4p + 13q = 2$ M1 A1
   $p + 2q + 0{\cdot}6 = 1$ $p + 2q = 0{\cdot}4$ B1
   Solve: $p = 0{\cdot}24$, $q = 0{\cdot}08$ M1 A1 A1
   (ii) $P(X < 7) = p + q + 0{\cdot}1 = 0{\cdot}42$ M1 A1
   (iii) $E(X^2) = 48{\cdot}54$ $Var(X) = 48{\cdot}54 - 6{\cdot}7^2 = 3{\cdot}65$ M1 A1 A1 11

6. (i) Let P(younger child is a boy, given elder is a boy) = p
   $\frac{5}{12}p + \frac{7}{12} \times \frac{3}{4} = \frac{9}{16}$ $\frac{5}{12}p = \frac{1}{8}$ $p = \frac{3}{10}$ M1 A1
   $P(B, B) = \frac{5}{12} \times \frac{3}{10} = \frac{1}{8}$ M1 A1
   (ii) $P(B, G \text{ or } G, B) = \frac{5}{12} \times \frac{7}{10} + \frac{7}{12} \times \frac{3}{4} = \frac{35}{48}$ or 0·729 M1 A1 A1
   (iii) $\frac{1}{8} \times \frac{1}{8} \times \frac{7}{8} \times 3 = \frac{21}{512}$ or 0·0410 M1 A1 A1
   (iv) Assumed independence B1 11

7. (i) $\sum x = 495$, $\sum y = 431$ B1 B1
   (ii) $S_{xx} = 1850{\cdot}5$, $S_{xy} = 1656{\cdot}5$ $y - 43{\cdot}1 = 0{\cdot}895(x - 49{\cdot}5)$ M1 A1
   $y = 0{\cdot}895x - 1{\cdot}21$ A1
   (iii) (a) 57, (b) 3 B1 B1
   (b) is less reliable as it is well outside the range of given values B1
   (iv) $r = \sqrt{(0{\cdot}895 \times 0{\cdot}921)} = 0{\cdot}908$ Quite good positive correlation M1 A1 B1 11