

STATISTICS (C) UNIT 1 TEST PAPER 1

1. The random variable X has the geometric distribution $\text{Geo}(0.75)$.
 Calculate (i) the expectation of X , [1]
 (ii) $P(X \leq 3)$, [3]
2. 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
 (i) the height of the '20 - 25' bar, [3]
 (ii) the total area under the histogram. [2]
 In the case when $x = 10$, *sketch* a cumulative frequency graph of this data. [2]
3. The 12 letters of the word INDEPENDENCE are written on cards, which are shuffled and then laid out on a table.
 Find the number of distinguishable ways in which
 (i) the 12 cards can be arranged in a straight line, [2]
 (ii) a pair of cards can be chosen from the 12. [2]
 The 12 cards are shuffled again and two cards are picked out at random. Find the probability that the letters on the two cards are the same. [3]
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 |
- (i) Write down the median mark. [1]
 (ii) Given that $\sum x = 1600$ and $\sum x^2 = 102\,400$, calculate the mean and the standard deviation of the 30 marks. [4]
 (iii) One candidate is chosen at random. Find the probability that this candidate's mark differs from the mean by more than the standard deviation. [2]
 The marks of another group of 30 candidates had mean 49 and standard deviation 17.5.
 (iv) Write down two brief comparisons between the two groups. [2]
5. A die is made in the form of a square-based pyramid. When it is thrown onto a table, the probability that it lands 'vertex up' (i.e. with the square face in contact with the table) is 0.55.
 The die is thrown 30 times. Find the probability that it lands vertex up
 (i) exactly 17 times, (ii) between 15 and 20 times (inclusive),
 (iii) more than 10 times.
 [6]
 Three identical dice of this type are thrown together 10 times. Find the probability that on all of these 10 occasions at least one of the dice lands vertex up.
 [4]
6. The distributions of two independent discrete random variables X and Y are given in the tables:
- | | | | | | | |
|----------|-----|-----|-----|----------|-------|-------|
| x | 0 | 1 | 2 | y | 0 | 1 |
| $P(X=x)$ | 0.6 | 0.3 | 0.1 | $P(Y=y)$ | 0.625 | 0.375 |
- The random variable Z is defined to be the sum of one observation from X and one from Y .
 (i) Tabulate the probability distribution for Z . [4]

- (ii) Calculate $E(Z)$. [2]
- (iii) Calculate (a) $E(Z^2)$, (b) $\text{Var}(Z)$. [4]
7. The table shows the numbers of cars and vans in a company's fleet having registrations with the prefix letters shown.
- | Registration letter | <i>K</i> | <i>L</i> | <i>M</i> | <i>N</i> | <i>P</i> | <i>R</i> | <i>S</i> | <i>T</i> | <i>V</i> |
|-----------------------------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| Number of cars (<i>x</i>) | 6 | 7 | 9 | 11 | 15 | 14 | 12 | 10 | 7 |
| Number of vans (<i>y</i>) | 8 | 10 | 14 | 13 | 13 | 15 | 14 | 9 | 8 |
- (i) 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. [3]
- (ii) Obtain the coordinates of a point which must lie on both the regression line of *y* on *x* and that of *x* on *y*. [2]
- (iii) If there were 4 *J*-registered cars, estimate the number of *J*-registered vans. [1]
- Given that $\Sigma x^2 = 1001$, $\Sigma y^2 = 1264$ and $\Sigma xy = 1106$,
- (iv) calculate the product-moment correlation coefficient between *x* and *y*. Give a brief interpretation of your answer. [3]
- (v) Find, in the form $y = ax + b$, an equation of the regression line of *y* on *x*. Use this line to obtain an improved answer to part (iii). [4]

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

1. (i) $E(X) = 1 \div p = 4/3$ B1
 (ii) $P(X \leq 3) = \frac{3}{4} + \frac{1}{4} \times \frac{3}{4} + \frac{1}{4} \times \frac{3}{4} = 63/64 = 0.984$ M1 A1 A1 4
 ie get it or not get it then get it or not get it twice then get it
2. (i) $3x$ corresponds to 24 cm^2 , so $5x$ corresponds to 40 cm^2 M1 A1
 Width = 2 cm, so height = 20 cm A1
 (ii) Area = $19x = 19 \times 8 = 152 \text{ cm}^2$ M1 A1
 Graph sketched B2 7
3. (i) $12! \div (2! \times 3! \times 4!) = 1\,663\,200$ M1 A1
 (ii) 2 the same : 3 ways 2 different : $6C2 = 15$ ways Total : 18 M1 A1
 Either 2E's or 2N's or 2D's
 ${}^4/_{12} \times {}^3/_{11} + {}^2/_{12} \times {}^1/_{11} + {}^3/_{12} \times {}^2/_{11} = {}^{10}/_{66} = 0.152$ B1M1 A1 7
4. (i) Median = 56.5 (ii) Mean = $1600 \div 30 = 53$ B1; B1
 Var = $102400/30 - ({}^{1600}/_{30})^2 = 568.89$, so s.d. = 23.9 M1 A1 A1
 (iii) Range is 29.5 to 77.2 10 outside this, so prob. = $\frac{1}{3}$ M1 A1
 (iv) Second group was less good overall but more consistent B1 B1 9
5. (i) $X \sim B(30, 0.55)$: $P(X \leq 17) - P(X \leq 16) = 0.6408 - 0.4975 = 0.143$ M1 A1
 (ii) $P(X \leq 20) - P(X \leq 14) = 0.9306 - 0.2309 = 0.700$ M1 A1
 (iii) $P(X > 10) = 1 - P(X \leq 10) = 1 - 0.0138 = 0.986$ M1 A1
 $P(\text{at least one in three vertex-up}) = 1 - 0.45^3 = 0.9089$ M1 A1
 Required probability = $0.9089^{10} = 0.385$ M1 A1 10

6. (i)
$$P(Z=z) \begin{matrix} z & 0 & 1 & 2 & 3 \\ & \frac{3}{8} & \frac{33}{80} & \frac{7}{40} & \frac{3}{80} \end{matrix}$$
 M1 A1 A1 A1
- (ii) $E(Z) = \frac{7}{8}$ ie $\sum(z \times P(Z=z))$ M1 A1
- (iii) (a) $E(Z^2) = \frac{29}{20}$ (b) $\text{Var}(Z) = \frac{29}{20} - \frac{49}{64} = \frac{219}{320} = 0.684$ M1 A1 M1 A1 10
7. (i) Scatter graph showing moderate positive correlation B3
- (ii) Means of x and y : (10.1, 11.6) M1 A1
- (iii) 7 or 8 B1
- (iv) $\sum x = 91$, $\sum y = 104$ B1
- $S_{xx} = 80.89$, $S_{yy} = 62.22$, $S_{xy} = 54.44$ $r = 0.767$ M1 A1
- which confirms the moderate positive correlation B1
- (v) $y - 11.6 = 0.673(x - 10.1)$ $y = 0.673x + 4.79$ M1 A1 A1
- When $x = 4$, $y \approx 7.5$, confirming 7 or 8 B1 13