

A LEVEL SECTION TESTS

unit S1

mathematical models in statistics

answers

S1/1

1. (a) The set of all items about which information is to be collected **B1**
(b) A mathematical technique chosen to represent a statistical problem **B1**
(c) An improvement which makes the model a closer fit to reality **B1**
(d) The set of all possible outcomes of an experiment **B1**
4
2. (a) Quantitative : values can be put in numerical order, e.g. lengths **B1**
Qualitative : cannot be put in numerical order, e.g. colours **B1**
(b) Discrete : can take only certain values, e.g. shoe size **B1**
Continuous : can take **any** value in a given range, e.g. temperature **B1**
4
3. (a) Nearest hour : discrete Nearest second and 0·1 second : continuous **B1 B1**
(b) Continuous graphs can be drawn; continuous distributions used as models **B1 B1**
4
4. (i)
(a) Throw the dice a large number of times and record the scores **B2**
(b) Use a tally and hence produce a frequency table for the discrete data **B2**
(c) Data is discrete Bar chart, pie chart **B2**
(d) Mean and standard deviation **B2**
(e) Model : a discrete random variable with probability distribution given by the relative frequencies of the scores **B2**
(f) The number of times a particular score will arise in a large number of further trials can be predicted **B2**
(g) The predictions can be compared with the actual outcomes **B2**
14
- (ii)
(a) Collect data by questioning a large sample of the residents **B2**
(b) Use a tally and hence produce a grouped frequency table **B2**
(c) Data is continuous Histogram, cumulative frequency graph **B2**
(d) Mean and standard deviation **or** median and interquartile range **B2**
(e) Model : a normal distribution with mean and variance determined from the sample **B2**
(f) The proportion of the population with heights within a given range can be estimated **B2**
(g) These proportions can be verified by examining a further sample **B2**
14

A LEVEL SECTION TESTS

unit S1

representation and summary of data

answers

S1/2b

1. (a) Mean = 5, standard deviation = $\sqrt{(79.57 - 25)} = 7.39$ **B1, M1 A1**
 (b) (i) Mean = $5 + 7 = 12$, standard deviation = 7.39 **M1 A1 A1**
 (ii) Mean = $5 \times 5 = 25$, standard deviation = $7.39 \times 5 = 36.9$ **M1 A1 A1**
- 9**
2. (a) Median = $(270 + 290) \div 2 = 280$ **B1**
 (b) y values : -70, -50, -40, -10, 10, 20, 30, 30 **B1**
 $\bar{y} = -80 \div 8 = -10$ $\sum y^2 = 11400$, so $\sigma_y = \sqrt{1325} = 36.4$ **A1 M1 A1**
 (c) Hence $\mu = -10 + 280 = 270$, $\sigma = 36.4$ **M1 A1 A1**
 (d) $\text{Var}(X) = 1325 = \sum (x - \mu)^2 / 8$, so $\sum (x - \mu)^2 = 10600$ **M1 A1**
- 10**
3. (a) 10, 15, 20, 25, 30, 35, 40, 45 **B1**
 (b)
- | x | 7.5 | 12.5 | 17.5 | 22.5 | 27.5 | 32.5 | 37.5 | 42.5 | 47.5 |
|-------|-----|------|------|------|------|------|------|------|------|
| y | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 |
| Freq. | 5 | 14 | 16 | 24 | 35 | 5 | 3 | 2 | 6 |
- $\bar{y} = (-20 - 42 - 32 - 24 + 5 + 6 + 6 + 24) \div 110 = -0.7$ **M1 A1**
 $\sum y^2 = 11(16) + 16(9) + 19(4) + 29(1) = 425$ **M1 A1**
 $\sigma_y = \sqrt{(425/110 - 0.7^2)} = 1.84$ **A1**
 $\bar{x} = 5 \times -0.7 + 27.5 = 24$ $\sigma_x = 5 \times 1.837 = 9.18$ **M1 A1 A1**
 (c) Exact distribution of values within groups is not known **B1**
 (d) Median and quartiles better, as median is not affected by the few high values but quartiles show the extent of the spread **B1**
- 14**
4. (a) (i) Median $\approx 79.5 + (4.5 \div 20) \times 20 = 84$ **M1 A1**
 (ii) $Q_1 \approx 39.5 + (11.5 \div 14) \times 20 \approx 56$ **M1 A1**
 $Q_3 \approx 99.5 + (6.5 \div 20) \times 20 \approx 106$ **M1 A1**
 (iii) IQR ≈ 50 **A1**
 (iv) 17% of 90 = 15.3, so $39.5 + (4.3 \div 14) \times 20 = 45.6$ **M1 A1**
 (b) Box plot drawn Negative skew **B2 B1**
 (c) Outliers outside 56 – 75 to 106 + 75, so no outliers **M1 A1**
 (d) Box plot drawn
 Second set has similar spread but its values tend to be slightly lower **B2**
- 17**

A LEVEL SECTION TESTS

unit S1

probability

answers

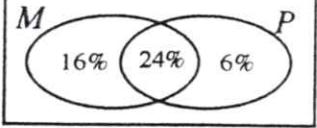
S1/3

1. (a) (i) $P(F) + P(G) = P(F \cup G)$, so $P(G) = \frac{5}{8} - \frac{1}{6} = \frac{11}{24}$ **M1 A1**
 (ii) $P(E \cap G) = P(E) + P(G) - P(E \cup G) = \frac{1}{3} + \frac{11}{24} - \frac{2}{3} = \frac{1}{8}$ **M1 A1**
 (b) $P(E \cap G) \neq P(E) \times P(G)$, so E and G are not independent **M1 A1**
 (c) (i) $P(G|E) = \frac{P(G \cap E)}{P(E)} = \frac{1}{8} \div \frac{1}{3} = \frac{3}{8}$ (ii) $\frac{P(G \cap E')}{P(E')} = \frac{1}{3} \div \frac{2}{3} = \frac{1}{2}$ **M1 A1 M1 A1**

10

2. (a) $(1, 3), (2, 0), (2, 2), (3, 3)$: probability = $\frac{4}{9}$ **M1 A1 A1**
 (b) $(1, 0), (1, 2), (1, 3), (2, 0), (3, 0)$: probability = $\frac{5}{9}$ **M1 A1 A1**
 (c) By tree diagram or otherwise : $\frac{4}{9} \times \frac{1}{4} + \frac{5}{9} \times \frac{3}{10} = \frac{5}{18}$ **M1 A1**
 (b) $1 - \frac{3}{10} = \frac{7}{10}$ **M1 A1**
 (c) $P(\text{red and not ace}) \div P(\text{not ace}) = \frac{5}{9} \times \frac{7}{10} \div \frac{13}{18} = \frac{7}{13}$ **M1 A1 M1 A1**

8

4. (a) 
 (b) (i) $P(P \cap M) = 6\% = 0.06$ **M1 A1**
 (ii) $P((P \cup M)^c) = 54\% = 0.54$ **M1 A1**
 (iii) $P(P'|M) = 0.54 \div 0.6 = 0.9$ **M1 A1**

8

5. (a) $1 - (0.3 + 0.45) = 0.25$ **M1 A1**
 (b) $1 - (0.3^2 + 0.45^2 + 0.25^2) = 0.645$ **M1 A1**
 (c) $(0.3 \times 0.3 \times 0.7) \times 3 = 0.189$ **M1 M1 A1**
 (d) $0.45^4 = 0.0410$ **M1 A1**
 (e) $\frac{0.3 \times 0.7}{0.25} = 0.84$ **M1 A1 A1**

12

6. (a) $P(\text{odd}) = \frac{2}{3}$ $P(\text{odd 12 times}) = \left(\frac{2}{3}\right)^{12} = 0.00771$ **B1 M1 A1**
 (b) $\left(\frac{2}{3}\right)^6 \times \left(\frac{1}{3}\right)^6 \times 2 = 0.000241$ **M1 A1 A1**

6

A LEVEL SECTION TESTS

unit S1

correlation and regression

answers

S1/4

- | | | | | |
|----|---|--|---------------------------------|-----------------|
| 1. | (a) Moderate negative correlation
(c) Only moderate accuracy – correlation is not strong | (b) To predict y , given x | B2 B2 | |
| | | | B2 | |
| | | | 6 | |
| 2. | $\sum x = 10$, $\sum y = 40$, $\sum x^2 = 18.64$, $\sum y^2 = 275$, $\sum xy = 70.42$ | | B1 B1 B1 B1 | |
| | (a) (i) $S_{xx} = 1.973$, $S_{yy} = 8.333$, $S_{xy} = 3.753$ | $y = 1.90x + 3.50$ | M1 A1 A1 | |
| | (ii) $x = 0.450y - 1.34$, so $y = 2.22x + 2.97$ | | M1 A1 A1 | |
| | (b) $r = \sqrt{(1.90 \times 0.450)} = 0.926$ | | M1 A1 A1 | |
| | (c) (i) When $x = 2$, $y = 7.30$ | (ii) When $y = 9.6$, $x = 2.99$ | B1 B1 | |
| | | r is close to 1, so expect quite good accuracy | B1 | |
| | (d) For $y < 2.96$, $x < 0$, so result is meaningless in this range | | B1 B1 | |
| | (e) $x = 1.67$, $y = 6.67$ | Mean values of x and y | B1 B1 B1 | |
| | | | 21 | |
| 3. | (a) Mean fits equation of regression line, so $\bar{q} = 2.5(13) + 15 = 47.5$ | | M1 A1 A1 | |
| | (b) $47.5 = 13m - 37$ | $m = 6.5$ | M1 A1 A1 | |
| | (c) $r = \sqrt{(2.5 / 6.5)} = 0.62$ | Moderate positive correlation | M1 A1 B2 | |
| | (d) 0.62 or their answer from (c) | | B2 | |
| | | | 12 | |
| 4. | (a) $S_{xx} = 615.785$ | $S_{xy}/S_{xx} = 2.4$, so $S_{xy} = 1477.884$ | M1 A1 A1 | |
| | | $\sum xy - (93.3)(90.6)/6 = 1477.884$ | $\sum xy = 2886.7 \approx 2890$ | M1 A1 A1 |
| | (b) $S_{yy} = 3884.44$, so $x = 0.381y + 9.80$ | | M1 A1 A1 | |
| | (c) $r = \sqrt{(2.4 \times 0.3805)} = 0.956$ | | M1 A1 | |
| | | | 11 | |

A LEVEL SECTION TESTS

unit S1

discrete random variables

answers

S1/5

1. (a) (i) 0·15 (ii) $0·3 + 0·25 = 0·55$ **B1 B1**
 (b) Mean = $E(X) = 0·3 + 0·5 + 0·45 + 0·4 = 1·65$ **M1 A1**
 $\text{Var}(X) = (0·3 + 1 + 1·35 + 1·6) - 1·65^2 = 1·5275$ **M1 A1**
 (c) F(x) values: 0·2, 0·5, 0·75, 0·9, 1 **M1 A1 A1**
- 9
2. (a) (i) $E(X) = \frac{1}{3} + \frac{1}{3} + \frac{3}{2} = \frac{13}{6} = 2\frac{1}{6}$ (ii) $E(X^2) = \frac{1}{3} + \frac{2}{3} + \frac{9}{2} = \frac{11}{2} = 5\frac{1}{2}$ **M1 A1 M1 A1**
 (iii) $\text{Var}(X) = 5\frac{1}{2} - 2\frac{1}{6}^2 = 0·806$ **M1 A1**
 (b) $2p = 1 - \frac{1}{2} - \frac{1}{4}$ $p = \frac{1}{8}$ **M1 A1**
 (c)
- | | | | | | | |
|------------|---------------|---------------|----------------|---------------|----------------|---------------|
| z | 2 | 3 | 4 | 5 | 6 | 7 |
| $P(Z = z)$ | $\frac{1}{6}$ | $\frac{1}{8}$ | $\frac{5}{16}$ | $\frac{1}{6}$ | $\frac{5}{48}$ | $\frac{1}{8}$ |
- B1
M1 A1 A1
- 12
3. (a) 4μ (b) $16\sigma^2$ (c) $2\mu - 3$ (d) $4\sigma^2$ **B1 B1 B1 B1**
- 4
4. (a) $E(X) = (15 + 1) \div 2 = 8$ $\text{Var}(X) = (225 - 1) \div 12 = 18·7$ **B1; M1 A1**
 (b) $P(X = 4, 5, \dots, 11) = 8 \div 20 = 0·4$ **M1 A1**
 (c) (i) $(n^2 - 1)/12 = 10$, so $n = 11$ (ii) $E(X) = 6$ **M1 A1 A1**
 (iii) $\text{Var}(3X + 2) = 9 \text{ Var}(X) = 90$ **B1**
- 9
5. (a) $k + 2k + 3k + 4k + 5k + 6k = 1$ $21k = 1$ $k = \frac{1}{21}$ **M1 A1**
 (b) $E(X) = 1 \times \frac{1}{21} + 2 \times \frac{2}{21} + 3 \times \frac{3}{21} + 4 \times \frac{4}{21} + 5 \times \frac{5}{21} + 6 \times \frac{6}{21} = \frac{91}{21} = 4\frac{1}{3}$ **M1 A1**
 $E(X^2) = \frac{441}{21} = 21$ $\text{Var}(X) = 21 - 4\frac{1}{3}^2 = 2·22$ **M1 A1 A1**
- 7
6. (a) $q + r = 0·2$ $0·3 + 15q + 17·5 + 120(0·2 - q) + 10 = 39·2$ **B1 M1 A1**
 $51·8 - 105q = 39·2$ $q = 0·12, r = 0·08$ **M1 A1 A1**
 (b) $E(X^2) = 3722·5$ $\text{Var}(X) = 3722·5 - 39·2^2 = 2185·86$ $\sigma = 46·8$ **M1 A1 A1**
- 9

A LEVEL SECTION TESTS

unit S1

the normal distribution

answers

S1/6

- | | | | |
|--|---|-----------------------|-----------------|
| 1. (a) Graph sketched : bell-shaped curve symmetrical about $z = 0$ | B1 B1 | | |
| $z = 1$ marked at sensible position for standard deviation | B1 | | |
| (b) (i) $P(Z \leq 0.5) = 0.692$ (ii) $P(Z > 1.5) = 1 - 0.9332 = 0.0668$ | B1 B1 | | |
| (iii) $P(-1 < Z < 1) = 2(0.3413) = 0.683$ | B1 | | |
| (iv) $E(2Z + 3) = 2E(Z) + 3 = 3$ (v) $\text{Var}(1 - 3Z) = 9\text{Var}(Z) = 9$ | B1 B1 | | |
| | 8 | | |
| 2. (a) Graph sketched : bell-shaped curve symmetrical about $x = 0$ | B1 B1 | | |
| $x = 5$ marked at sensible position for 1 standard deviation from mean | B1 | | |
| (b) $Z = (X - 3)/2$, so $X = 2Z + 3$ | M1 A1 | | |
| (c) (i) $P(Z > -0.5) = 0.692$ (ii) $P(Z < -1.5) = 0.0668$ | B1 M1 A1 | | |
| (iii) $P(1 < X < 5) = P(-1 < Z < 1) = 0.682$ | M1 A1 | | |
| (iv) $E(2X + 3) = 2E(X) + 3 = 10$ (v) $3\sigma_x = 6$ | B1 B1 | | |
| | 12 | | |
| 3. (a) $P(X < 4.2) = P(Z < (4.2 - 4.5)/2.5) = P(Z < -0.12) = 0.452$ | M1 A1 | | |
| (b) $P(X > 5) = P(Z > (5 - 4.5)/2.5) = P(Z > 0.2) = 0.421$ | M1 A1 | | |
| (c) $P(2.5 < X < 5.5) = P(-0.8 < Z < 0.4) = 0.6554 - 0.2119 = 0.444$ | M1 A1 A1 | | |
| | 7 | | |
| 4. (a) (i) $P(4 < X < 7) = P(-1/1.5 < Z < 2/1.5) = P(-0.67 < Z < 1.33)$
$= 0.90827 - 0.2514 = 0.657$ | M1 A1 | | |
| (ii) If $P(-z < Z < z) = 0.85$ then $P(Z < z) = 0.5 + 0.425 = 0.925$, so
$z = 1.44$ Range is $5 - 1.44 \times 1.5$ to $5 + 1.44 \times 1.5$,
i.e. 2.84 hrs to 7.16 hrs | A1 M1 | | |
| (b) Unlikely to be a good model for values far from the mean: no times will be close to zero, but there may be a lot of high values | A1 | | |
| | 9 | | |
| 5. $(2 - \mu)/\sigma = -1.15$, $(3 - \mu)/\sigma = 0.61$ | B1 B1 | | |
| $2 - \mu = -1.15\sigma$ | $3 - \mu = 0.61\sigma$ | Add: $1.76\sigma = 1$ | M1 M1 A1 |
| $\sigma = 0.568$ | $\mu = 2.65$ | | A1 A1 |
| | | | 7 |
| 6. (a) $\sigma = 30$ | If $P(Z > z) = 0.791$ then $z = -0.81$, so | B1 M1 A1 | |
| $(325 - \mu)/30 = -0.81$ | $\mu = 349.3$ | M1 A1 | |
| (b) $P(X < 325) = 0.209$ | Expected no. = $40 \times 0.209 \approx 8$ | M1 A1 | |
| | | 7 | |