Leave blank

1. A clothes shop manager records the weekly sales figures, £ s, and the average weekly temperature, t °C, for 6 weeks during the summer. The sales figures were coded so that

$$w = \frac{s}{1000}$$

The data are summarised as follows

$$S_{ww} = 50 \quad \sum wt = 784 \quad \sum t^2 = 2435 \quad \sum t = 119 \quad \sum w = 42$$

(a) Find S_{wt} and S_{tt}

(3)

(b) Write down the value of S_{ss} and the value of S_{st}

(2)

(c) Find the product moment correlation coefficient between s and t.

(2)

The manager of the clothes shop believes that a linear regression model may be appropriate to describe these data.

(d) State, giving a reason, whether or not your value of the correlation coefficient supports the manager's belief.

(1)

(e) Find the equation of the regression line of w on t, giving your answer in the form $w = a + bt$

(3)

(f) Hence find the equation of the regression line of s on t, giving your answer in the form $s = c + dt$, where c and d are correct to 3 significant figures.

(2)

(g) Using your equation in part (f), interpret the effect of a 1°C increase in average weekly temperature on weekly sales during the summer.

(1)

(a) $S_{wt} = \Sigma wt - \frac{\Sigma w \Sigma t}{n}$ $\quad$ $S_{tt} = \Sigma t^2 - \frac{(\Sigma t)^2}{n}$

$= 784 - \frac{42 \times 119}{6}$ $\quad$ $= 2435 - \frac{119^2}{6}$

$= -49$ $\quad$ $= 74.83333$

(b) $S_{ss} = 1000^2 S_{ww}$ $\quad$ $S_{st} = 1000 S_{wt}$

$= 50 \times 10^6$ $\quad$ $= -49 \times 10^3$

(c) $r = \frac{S_{st}}{\sqrt{S_{ss} S_{tt}}}$

$= \frac{-49 \times 10^3}{\sqrt{50 \times 10^6 \times 74.8333}}$

$= -0.8011$

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Leave blank

Question 1 continued

(d) It does because the PMCC is fairly close to −1.

(e) $b = \frac{S_{wt}}{S_{tt}}$

$= \frac{-49}{74.8333}$

$= -0.654788$

$a = \bar{w} - b\bar{t}$

$= \frac{42}{6} + 0.654788 \times \frac{119}{6}$

$= 19.9866$

$\therefore\ w = 19.9866 - 0.654788t$

(f) $s = 19986.6 - 654.788t$

$= 20 \times 10^3 - 655t$

(g) A 1°C increase should see a £655 decrease in weekly sales.

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

2. An estate agent is studying the cost of office space in London. He takes a random sample of 90 offices and calculates the cost, £x per square foot. His results are given in the table below.

Cost (£x)	Frequency (f)	cf	Midpoint (£y)	fd
$20 \leqslant x < 40$	12	12	30	0.6
$40 \leqslant x < 45$	13	25	42.5	2.6
$45 \leqslant x < 50$	25	50	47.5	5
$50 \leqslant x < 60$	32	82	55	3.2
$60 \leqslant x < 80$	8	90	70	0.4

(You may use $\sum \mathrm{f} y^2 = 226\,687.5$)

A histogram is drawn for these data and the bar representing $50 \leqslant x < 60$ is 2 cm wide and 8 cm high.

(a) Calculate the width and height of the bar representing $20 \leqslant x < 40$ **(3)**

(b) Use linear interpolation to estimate the median cost. **(2)**

(c) Estimate the mean cost of office space for these data. **(2)**

(d) Estimate the standard deviation for these data. **(2)**

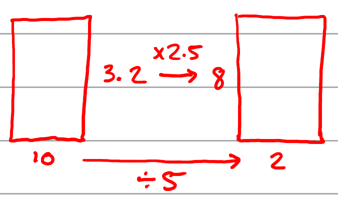
(e) Describe, giving a reason, the skewness. **(1)**

Rika suggests that the cost of office space in London can be modelled by a normal distribution with mean £50 and standard deviation £10

(f) With reference to your answer to part (e), comment on Rika's suggestion. **(1)**

(g) Use Rika's model to estimate the 80th percentile of the cost of office space in London. **(3)**

(a)

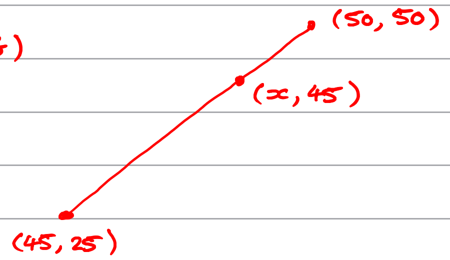


∴ For 20–40 class,
width $= \frac{20}{5} = 4$ cm
height $= 0.6 \times 2.5 = 1.5$ cm

Leave blank

Question 2 continued

(b)



$$\frac{50-25}{50-45} = \frac{45-25}{x-45}$$

$$5 = \frac{20}{x-45}$$

$$x - 45 = 4$$

$$x = £49 \ m^{-2}$$

(c) $\bar{x} = \frac{\Sigma x}{n}$

$= \frac{\Sigma fx}{\Sigma f}$

$= \frac{12 \times 30 + 42.5 \times 13 + 25 \times 47.5 + 55 \times 32 + 8 \times 70}{90}$

$= £49.1 \ m^{-2}$

(d) $\sigma^2 = \frac{\Sigma fx^2}{\Sigma f} - \left(\frac{\Sigma fx}{\Sigma f}\right)^2$

$= \frac{226687.5}{90} - 49.1^2$

$= 106.85$

$\sigma = £10.337 \ m^{-2}$

(e) The median and mean are virtually the same, so the data is approximately symmetrical.

(f) Rika's suggestion is a good one b/c normal distributions are symmetrical.

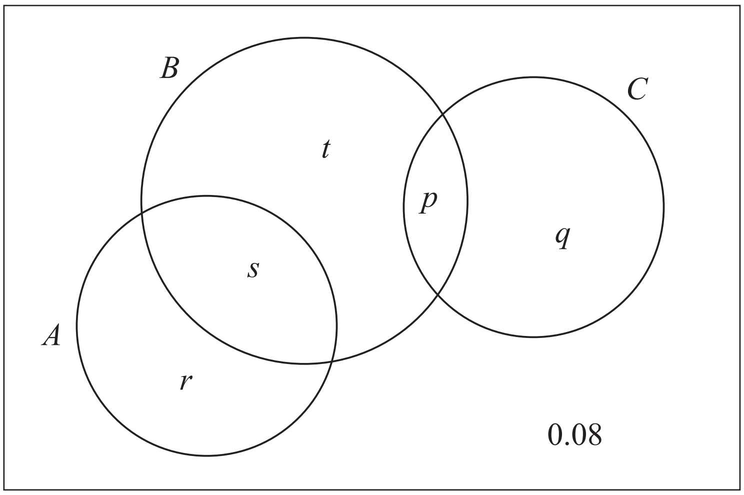
(g) $X \sim N(50, 10^2)$

$P(X < x) = 0.8$

$x = £58.4162$

Leave blank

3. The Venn diagram shows three events A, B and C, where p, q, r, s and t are probabilities.



P(A) = 0.5, P(B) = 0.6 and P(C) = 0.25 and the events B and C are independent.

(a) Find the value of p and the value of q. **(2)**

(b) Find the value of r. **(2)**

(c) Hence write down the value of s and the value of t. **(2)**

(d) State, giving a reason, whether or not the events A and B are independent. **(2)**

(e) Find P($B \mid A \cup C$). **(3)**

(a) $p = P(A)\,P(B)$
$= 0.6 \times 0.25$
$= 0.15$

$q = P(C) - p$
$= 0.25 - 0.15$
$= 0.1$

(b) $P(B \cup C) = P(B) + P(C) - p$
$= 0.6 + 0.25 - 0.15$
$= 0.7$
$r = 1 - 0.7 - 0.08$
$= 0.22$

(c) $s = P(A) - r$
$= 0.5 - 0.22$
$= 0.28$

$t = P(B) - p - s$
$= 0.6 - 0.15 - 0.28$
$= 0.17$

Leave blank

Question 3 continued

(d) $P(A)\,P(B) = 0.5 \times 0.6$

$= 0.3$

But $s = 0.28 \Rightarrow A$ and B are not independent.

(e) $P(B \mid A \cup C) = \dfrac{0.15 + 0.28}{0.5 + 0.25}$

$= 0.5733$

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Leave blank

4. The discrete random variable X has probability distribution

x	-1	0	1	2
$\mathrm{P}(X = x)$	a	b	b	c

The cumulative distribution function of X is given by

x	-1	0	1	2
$\mathrm{F}(x)$	$\frac{1}{3}$	d	$\frac{5}{6}$	e

(a) Find the values of a, b, c, d and e.

(5)

(b) Write down the value of $\mathrm{P}(X^2 = 1)$.

(1)

(a) $a = \frac{1}{3}$

$e = 1$

$c = \frac{1}{6}$

$b = \frac{1}{4}$

$d = \frac{7}{12}$

(b) $P(X^2 = 1) = a + b$

$= \frac{7}{12}$

Leave blank

5. Yuto works in the quality control department of a large company. The time, T minutes, it takes Yuto to analyse a sample is normally distributed with mean 18 minutes and standard deviation 5 minutes.

(a) Find the probability that Yuto takes longer than 20 minutes to analyse the next sample.
(3)

The company has a large store of samples analysed by Yuto with the time taken for each analysis recorded. Serena is investigating the samples that took Yuto longer than 15 minutes to analyse.

She selects, at random, one of the samples that took Yuto longer than 15 minutes to analyse.

(b) Find the probability that this sample took Yuto more than 20 minutes to analyse.
(4)

Serena can identify, in advance, the samples that Yuto can analyse in under 15 minutes and in future she will assign these to someone else.

(c) Estimate the median time taken by Yuto to analyse samples in future.
(5)

(a) $T \sim N(18, 5^2)$

$P(T > 20) = 0.344\,578$

(b) $P(T > 15) = 0.725\,747$

$P(T > 20 \mid T > 15) = \dfrac{P(T > 20)}{P(T > 15)}$

$= \dfrac{0.344\,578}{0.725\,747}$

$= 0.474\,791$

(c) $\frac{1}{2} \times 0.725\,747 = 0.362\,873$

$P(T > t) = 0.362\,873$

$t = 19.753\,94$ min

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Leave blank

6. The score, X, for a biased spinner is given by the probability distribution

x	0	3	6
$P(X = x)$	$\frac{1}{12}$	$\frac{2}{3}$	$\frac{1}{4}$

Find

(a) $E(X)$
(2)

(b) $Var(X)$
(3)

A biased coin has one face labelled 2 and the other face labelled 5
The score, Y, when the coin is spun has

$$P(Y = 5) = p \quad \text{and} \quad E(Y) = 3$$

(c) Form a linear equation in p and show that $p = \frac{1}{3}$
(3)

(d) Write down the probability distribution of Y.
(1)

Sam plays a game with the spinner and the coin.
Each is spun once and Sam calculates his score, S, as follows

if $X = 0$ then $S = Y^2$
if $X \neq 0$ then $S = XY$

(e) Show that $P(S = 30) = \frac{1}{12}$
(2)

(f) Find the probability distribution of S.
(3)

(g) Find $E(S)$.
(2)

Charlotte also plays the game with the spinner and the coin.
Each is spun once and Charlotte ignores the score on the coin and just uses X^2 as her score.
Sam and Charlotte each play the game a large number of times.

(h) State, giving a reason, which of Sam and Charlotte should achieve the higher total score.
(2)

Leave blank

Question 6 continued

(a) $E(X) = 0 \times \frac{1}{12} + 3 \times \frac{2}{3} + 6 \times \frac{1}{4}$

$= 3.5$

(b) $Var(X) = E(X^2) - (E(X))^2$

$= 0^2 \times \frac{1}{12} + 3^2 \times \frac{2}{3} + 6^2 \times \frac{1}{4} - 3.5^2$

$= 2.75$

(c) $3 = 2(1-p) + 5p$

$= 2 - 2p + 5p$

$1 = 3p$

$p = \frac{1}{3}$, QED.

(d)

y	2	5
$P(Y=y)$	$\frac{2}{3}$	$\frac{1}{3}$

(e) $P(S = 30) = P(X=6)\,P(Y=5)$

$= \frac{1}{4} \times \frac{1}{3}$

$= \frac{1}{12}$, QED.

(f)

		x: 0	3	6
y	2	4	6	12
	5	25	15	30

s	4	6	12	15	25	30
$P(S=s)$	$\frac{1}{12} \times \frac{2}{3}$	$\frac{2}{3} \times \frac{2}{3}$	$\frac{1}{4} \times \frac{2}{3}$	$\frac{2}{3} \times \frac{1}{3}$	$\frac{1}{12} \times \frac{1}{3}$	$\frac{1}{4} \times \frac{1}{3}$
	$\frac{1}{18}$	$\frac{4}{9}$	$\frac{1}{6}$	$\frac{2}{9}$	$\frac{1}{36}$	$\frac{1}{12}$

(g) $E(S) = 4 \times \frac{1}{18} + 6 \times \frac{4}{9} + 12 \times \frac{1}{6} + 15 \times \frac{2}{9} + 25 \times \frac{1}{36} + 30 \times \frac{1}{12}$

$= \frac{137}{12}$

(h) $E(X^2) = 15$ (from part (b))

$15 > \frac{137}{12} \Rightarrow$ Charlotte will have the higher score in the long run.