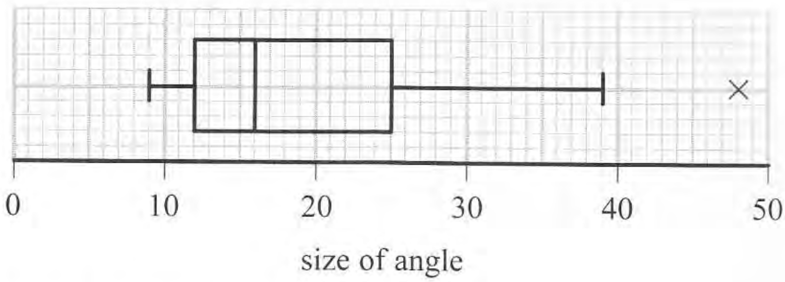
S1S S1

1. Each of 60 students was asked to draw a $20°$ angle without using a protractor. The size of each angle drawn was measured. The results are summarised in the box plot below.



(a) Find the range for these data.

(1)

(b) Find the interquartile range for these data.

(1)

The students were then asked to draw a $70°$ angle.
The results are summarised in the table below.

Angle, *a*, (degrees)	**Number of students**	
$55 \leqslant a < 60$	6	6
$60 \leqslant a < 65$	15	21
$65 \leqslant a < 70$	13	34
$70 \leqslant a < 75$	11	45
$75 \leqslant a < 80$	8	53
$80 \leqslant a < 85$	7	60

(c) Use linear interpolation to estimate the size of the median angle drawn. Give your answer to 1 decimal place.

(2)

(d) Show that the lower quartile is $63°$

(2)

For these data, the upper quartile is $75°$, the minimum is $55°$ and the maximum is $84°$

An outlier is an observation that falls either
more than $1.5 \times$ (interquartile range) above the upper quartile or
more than $1.5 \times$ (interquartile range) below the lower quartile.

(e) (i) Show that there are no outliers for these data.

(ii) Draw a box plot for these data on the grid on page 3.

(5)

(f) State which angle the students were more accurate at drawing. Give reasons for your answer.

(3)

a) $48 - 9 = 39$ b) $IQR = 25 - 12 = 13$

c) $\frac{1}{2}n = 30$

65	Q_2	70
21	30	34

$\frac{Q_2 - 65}{5} = \frac{9}{13}$ $Q_2 = 68.5$

d) $\frac{1}{4}n = 15$

60	Q_1	65
6	15	21

$\frac{Q_1 - 60}{5} = \frac{9}{15}$

$Q_1 = \frac{9}{15} \times 5 + 60 = 63$

e) lower limit $= Q_1 - 1.5\,IQR = 63 - 1.5(75 - 63) = 45$

$\therefore$ no lower outlier

upper limit $= Q_3 + 1.5\,IQR = 75 + 1.5(75 - 63) = 93$

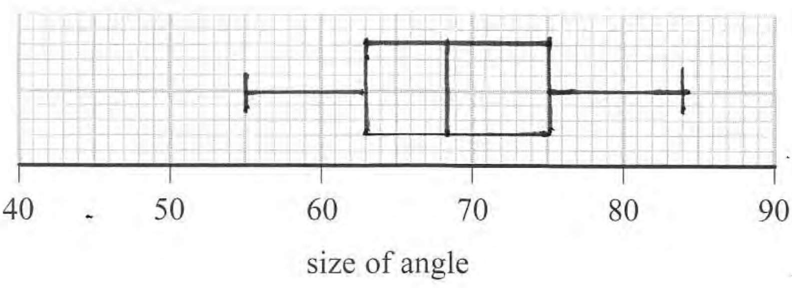
$\therefore$ no upper outlier

f) median for 70° was 68.5, so 1.5° out
median for 20° was 16°, so 4° out
$\therefore$ On average 70° was closer to actual.

IQR 70° was 12°
IQR 20° was 13°
$\therefore$ 70° guesses were more consistent

20° contained an outlier, 70° did not.

$\therefore$ Students were more acurate at drawing 70°.



2. An estate agent recorded the price per square metre, p £/m^2, for 7 two-bedroom houses.

He then coded the data using the coding $q = \dfrac{p - a}{b}$, where a and b are positive constants.

His results are shown in the table below.

p	1840	1848	1830	1824	1819	1834	1850
q	4.0	4.8	3.0	2.4	1.9	3.4	5.0

(a) Find the value of a and the value of b

(2)

The estate agent also recorded the distance, d km, of each house from the nearest train station. The results are summarised below.

$$S_{dd} = 1.02 \quad S_{qq} = 8.22 \quad S_{dq} = -2.17$$

(b) Calculate the product moment correlation coefficient between d and q

(2)

(c) Write down the value of the product moment correlation coefficient between d and p

(1)

The estate agent records the price and size of 2 additional two-bedroom houses, H and J.

House	Price (£)	Size (m^2)
H	156 400	85
J	172 900	95

(d) Suggest which house is most likely to be closer to a train station. Justify your answer.

(3)

a) $4 = \frac{1840 - a}{b} \Rightarrow 1840 - a = 4b$

$3 = \frac{1830 - a}{b} \Rightarrow 1830 - a = 3b$

$10 = b \quad \therefore \quad a = 1800$

b) $r = \frac{S_{dq}}{\sqrt{S_{dd} \times S_{qq}}} = \frac{-2.17}{\sqrt{1.02 \times 8.22}} = -0.749$

c) -0.749

d) H $p = \frac{156400}{85} = 1840$ J $p = \frac{172900}{95} = 1820$

PMCC of -0.749 suggests evidence to support negative correlation

as the distance increases, the price falls

$\therefore$ H is likely to be closer to a train station.

3. A college has 80 students in Year 12.

20 students study Biology
28 students study Chemistry
30 students study Physics
7 students study both Biology and Chemistry
11 students study both Chemistry and Physics
5 students study both Physics and Biology
3 students study all 3 of these subjects

(a) Draw a Venn diagram to represent this information.
(5)

A Year 12 student at the college is selected at random.

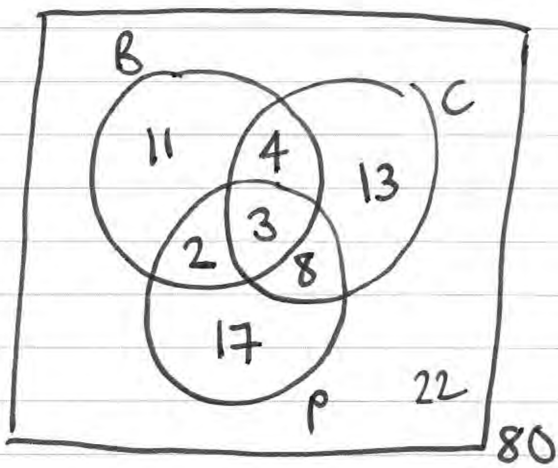
(b) Find the probability that the student studies Chemistry but not Biology or Physics.
(1)

(c) Find the probability that the student studies Chemistry or Physics or both.
(2)

Given that the student studies Chemistry or Physics or both,

(d) find the probability that the student does not study Biology.
(2)

(e) Determine whether studying Biology and studying Chemistry are statistically independent.
(3)



b) $\frac{13}{80}$

c) $\frac{47}{80}$

d) $\frac{38}{47}$

$P(B \cap C) = \frac{7}{80}$ $\quad P(B) \times P(C) = \frac{20}{80} \times \frac{28}{80} = \frac{7}{80}$

$\therefore P(B \cap C) = P(B) \times P(C)$ $\therefore$ they are Independent.

4. Statistical models can provide a cheap and quick way to describe a real world si

(a) Give two other reasons why statistical models are used.

(2)

A scientist wants to develop a model to describe the relationship between the average daily temperature, x °C, and her household's daily energy consumption, y kWh, in winter.

A random sample of the average daily temperature and her household's daily energy consumption are taken from 10 winter days and shown in the table.

x	–0.4	–0.2	0.3	0.8	1.1	1.4	1.8	2.1	2.5	2.6
y	28	30	26	25	26	27	26	24	22	21

[You may use $\sum x^2 = 24.76$ $\sum y = 255$ $\sum xy = 283.8$ $S_{xx} = 10.36$]

(b) Find S_{xy} for these data.

(3)

(c) Find the equation of the regression line of y on x in the form $y = a + bx$

Give the value of a and the value of b to 3 significant figures.

(4)

(d) Give an interpretation of the value of a

(1)

(e) Estimate her household's daily energy consumption when the average daily temperature is 2 °C

(2)

The scientist wants to use the linear regression model to predict her household's energy consumption in the summer.

(f) Discuss the reliability of using this model to predict her household's energy consumption in the summer.

(2)

a) To help understanding of a real world situation
Simplify a more complex situation
To help make predictions

b) $S_{xy} = \Sigma xy - \frac{(\Sigma x)(\Sigma y)}{n} = 283.8 - \frac{(12)(255)}{10}$

$\Sigma x = 12$
$\Sigma y = 255$

$\underline{\underline{S_{xy} = -22.2}}$

c) $b = \frac{S_{xy}}{S_{xx}} = \frac{-22.2}{10.36} = -2.142857...$

$a = \bar{y} - b\bar{x} = \left(\frac{255}{10}\right) - b\left(\frac{12}{10}\right) = 28.0714...$

$\therefore y = 28.1 - 2.14x$

d) a = energy cosumption at a temp of $0°c$

e) $x = 2$ $\quad y = 28.1 - 2.14(2) = \underline{\underline{23.8}}$

f) unreliable, extrapolation as you would expect temperatures to be much higher.

5. In a quiz, a team gains 10 points for every question it answers correctly and loses 5 for every question it does not answer correctly. The probability of answering a question correctly is 0.6 for each question. One round of the quiz consists of 3 questions.

The discrete random variable X represents the total number of points scored in one round. The table shows the incomplete probability distribution of X

x	30	15	0	-15
$\mathrm{P}(X=x)$	0.216			0.064

(a) Show that the probability of scoring 15 points in a round is 0.432 **(2)**

(b) Find the probability of scoring 0 points in a round. **(1)**

(c) Find the probability of scoring a total of 30 points in 2 rounds. **(3)**

(d) Find $\mathrm{E}(X)$ **(2)**

(e) Find $\mathrm{Var}(X)$ **(3)**

In a bonus round of 3 questions, a team gains 20 points for every question it answers correctly and loses 5 points for every question it does not answer correctly.

(f) Find the expected number of points scored in the bonus round. **(3)**

a) ✓✓✗, ✓✗✓, ✗✓✓ $\}$ $3 \times 0.6 \times 0.6 \times 0.4 = 0.432$

b) ✓✗✗, ✗✓✗, ✗✗✓ $\}$ $3 \times 0.6 \times 0.4 \times 0.4 = 0.288$

c)

x	30	15	0	-15
p	0.216	0.432	0.288	0.064

$P(30,0) = 0.216 \times 0.288 = 0.062208$

$P(0,30) = \nearrow$

$P(15,15) = 0.432^2 = 0.186624$ +

0.31104

$\underline{\underline{0.311}}$

d) $E(x) = 30 \times 0.216 + 15 \times 0.432 + 0 - 15 \times 0.064$

$= \underline{\underline{12}}$

e) $E(x^2) = 30^2 \times 0.216 + 15^2 \times 0.432 + 15^2 \times 0.064 = 306$

$V(x) = E(x^2) - E(x)^2 = 306 - 12^2 = \underline{\underline{162}}$

f)

x	60	35	10	-15
p	0.216	0.432	0.288	0.064

$E(x) = 60 \times 0.216 + 35 \times 0.432 + 10 \times 0.288 - 15 \times 0.064$

$= \underline{\underline{30}}$

6. The random variable $Z \sim \mathrm{N}(0, 1)$

A is the event $Z > 1.1$
B is the event $Z > -1.9$
C is the event $-1.5 < Z < 1.5$

(a) Find

(i) $\mathrm{P}(A)$

(ii) $\mathrm{P}(B)$

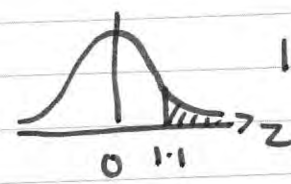
(iii) $\mathrm{P}(C)$

(iv) $\mathrm{P}(A \cup C)$ **(6)**

The random variable X has a normal distribution with mean 21 and standard deviation 5

(b) Find the value of w such that $\mathrm{P}(X > w \mid X > 28) = 0.625$ **(6)**

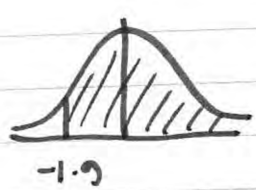
i a) $P(A) = P(Z > 1.1)$



$1 - \Phi(1.1)$
$= 0.1357$

2

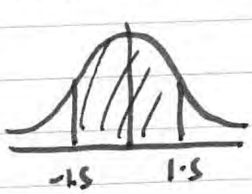
ii) $P(B) = P(Z > -1.9)$



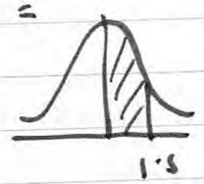
$= \Phi(1.9) = 0.9713$

2

iii) $P(C) = P(-1.5 < Z < 1.5)$



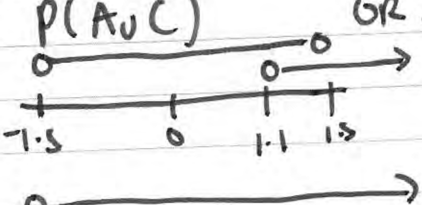
$= 2\times$ =



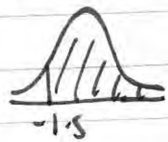
$= 2 \times [\Phi(1.5) - 0.5]$
$= 2 \times 0.4332$
$= 0.8664$

2

iv) $P(A \cup C)$ OR.



$= P(Z > -1.5) =$



$= \Phi(1.5) = 0.9332$

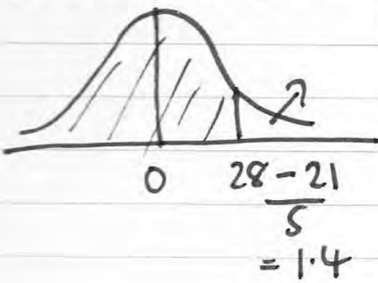
 GradeMax

b) $X \sim N(21, 5^2)$

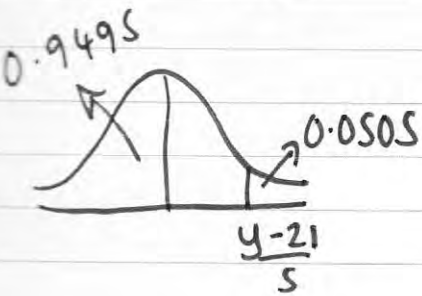
$$P(X > w \mid X > 28) = \frac{P(X > w \cap X > 28)}{P(X > 28)}$$

$P(X > 28) = P\left(Z > \frac{28-21}{5}\right) = P(Z > 1.4)$

$1 - \Phi(1.4) = 0.0808$



$\therefore \frac{P(X > w \cap X > 28)}{0.0808} = 0.625 \Rightarrow P(X > w \cap X > 28) = 0.0505$



$\Phi(y) = 0.9495$

$\therefore \frac{y-21}{5} = 1.64$

$y - 21 = 8.2$

$y = 29.2$

Since $29.2 > 28$

$\therefore w = 29.2$