3. The table shows the price of a bottle of milk, m pence, and the price of a loaf of bread, b pence, for 8 different years.

m	29	29	35	39	41	43	44	46
b	75	83	91	121	120	126	119	126

(You may use $S_{bb} = 3083.875$ and $S_{mm} = 305.5$)

(a) Find the exact value of $\sum bm$

(1)

(b) Find S_{bm}

(3)

(c) Calculate the product moment correlation coefficient between b and m

(2)

(d) Interpret the value of the correlation coefficient.

(1)

A ninth year is added to the data set. In this year the price of the bottle of milk is 46 pence and the price of a loaf of bread is 175 pence.

(e) Without further calculation, state whether the value of the product moment correlation coefficient will increase, decrease or stay the same when all nine years are used. Give a reason for your answer.

(2)

3. The weights of women boxers in a tournament are normally distributed with mean 64 kg and standard deviation 8 kg.

(a) Find the probability that a randomly chosen woman boxer in the tournament weighs less than 51 kg.

(3)

In the tournament, women boxers who weigh less than 51 kg are classified as lightweight. Ren weighs 49 kg and she has a match against another randomly selected, lightweight woman boxer.

(b) Find the probability that Ren weighs less than the other boxer.

(4)

In the tournament, women boxers who weigh more than H kg are classified as heavyweight. Given that 10% of the women boxers in the tournament are classified as heavyweight,

(c) find the value of H.

(3)

3. Morgan is investigating the body length, b centimetres, of squirrels.

A random sample of 8 squirrels is taken and the data for each squirrel is coded using

$$x = \frac{b - 21}{2}$$

The results for the coded data are summarised below

$$\sum x = -1.2 \qquad \sum x^2 = 5.1$$

(a) Find the mean of b

(3)

(b) Find the standard deviation of b

(3)

A 9th squirrel is added to the sample. Given that for all 9 squirrels $\sum x = 0$

(c) find

(i) the body length of the 9th squirrel,

(2)

(ii) the standard deviation of x for all 9 squirrels.

(2)

4. The weights of packages delivered to Susie are normally distributed with a mean of 510 grams and a standard deviation of 45 grams.

(a) Find the probability that a randomly selected package delivered to Susie weighs less than 450 grams.

(3)

The heaviest 5% of packages delivered to Susie are delivered by Rav in his van, the others are delivered by Taruni on foot.

(b) Find the weight of the lightest package that Rav would deliver to Susie.

(3)

Susie randomly selects a package from those delivered by Taruni.

(c) Find the probability that this package weighs more than 450 grams.

(4)

On Tuesday there are 5 packages delivered to Susie.

(d) Find the probability that 4 are delivered by Taruni and 1 is delivered by Rav.

(3)

6. *Ranpose* hospital offers services to a large number of clinics that refer patients to a range of hospitals.

The manager at *Ranpose* hospital took a random sample of 16 clinics and recorded

- the distance, x km, of the clinic from *Ranpose* hospital
- the percentage, y%, of the referrals from the clinic who attend *Ranpose* hospital.

The data are summarised as

$$\bar{x} = 8.1 \qquad \bar{y} = 20.5 \qquad \sum y^2 = 8266 \qquad S_{xx} = 368.16 \qquad S_{xy} = -630.9$$

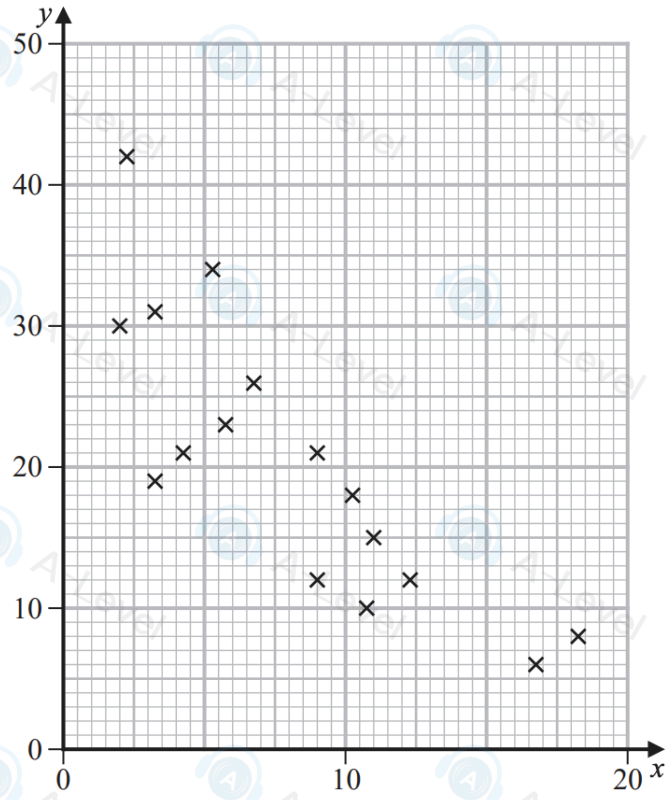
(a) Find the product moment correlation coefficient for these data.

(4)

(b) Give an interpretation of your correlation coefficient.

(1)

The manager at *Ranpose* hospital believes that there may be a linear relationship between the distance of a clinic from the hospital and the percentage of the referrals who attend the hospital. She drew the following scatter diagram for these data.



(c) State, giving a reason, whether or not these data support the manager's belief.

(1)

(d) Find the equation of the regression line of y on x, giving your answer in the form $y = a + bx$

(4)

(e) Give an interpretation of the gradient of your regression line.

(1)

(f) Draw your regression line on the scatter diagram.

(1)

The manager believes that *Ranpose* hospital should be attracting an "above average" percentage of referrals from clinics that are less than 5 km from the hospital. She proposes to target one clinic with some extra publicity about the services *Ranpose* offers.

(g) On the scatter diagram circle the point representing the clinic she should target.

(1)

2. The discrete random variable X has the following probability distribution.

x	-2	-1	0	1	3
$P(X = x)$	0.15	a	b	a	0.4

(a) Find $E(X)$.

(2)

Given that $E(X^2) = 4.54$

(b) find the value of a and the value of b.

(5)

The random variable $Y = 3 - 2X$

(c) Find $Var(Y)$.

(3)

5. The resting heart rate, h beats per minute (bpm), and average length of daily exercise, t minutes, of a random sample of 8 teachers are shown in the table below.

t	20	35	40	25	45	70	75	90
h	88	85	77	75	71	66	60	54

(a) State, with a reason, which variable is the response variable.

(2)

The equation of the least squares regression line of h on t is

$$h = 93.5 - 0.43t$$

(b) Give an interpretation of the gradient of this regression line.

(1)

(c) Find the value of $\bar{t}$ and the value of $\bar{h}$

(2)

(d) Show that the point $(\bar{t}, \bar{h})$ lies on the regression line.

(1)

(e) Estimate the resting heart rate of a teacher with an average length of daily exercise of 1 hour.

(1)

(f) Comment, giving a reason, on the reliability of the estimate in part (e).

(2)

The resting heart rate of teachers is assumed to be normally distributed with mean 73 bpm and standard deviation 8 bpm.

The middle 95% of resting heart rates of teachers lies between a and b

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

(4)

4: Two events C and D are such that

$$P(C \cup D) = 0.59 \qquad P(D) = 0.45 \qquad P(C \mid D) = 0.2$$

Find the value of

(a) $P(C \cap D)$

(2)

(b) $P(C)$

(2)

3. The parking times, t hours, for cars in a car park are summarised below.

Time (t hours)	Frequency (f)	Time midpoint (m)
$0 \leqslant t < 1$	10	0.5
$1 \leqslant t < 2$	18	1.5
$2 \leqslant t < 4$	15	3
$4 \leqslant t < 6$	12	5
$6 \leqslant t < 12$	5	9

(You may use $\sum fm = 182$ and $\sum fm^2 = 883$)

A histogram is drawn to represent these data.

The bar representing the time $1 \leqslant t < 2$ has a width of 1.5 cm and a height of 6 cm.

(a) Calculate the width and the height of the bar representing the time $4 \leqslant t < 6$

(3)

(b) Use linear interpolation to estimate the median parking time for the cars in the car park.

(2)

(c) Estimate the mean and the standard deviation of the parking time for the cars in the car park.

(3)

(d) Describe, giving a reason, the skewness of the data.

(2)

One of these cars is selected at random.

(e) Estimate the probability that this car is parked for more than 75 minutes.

(3)