Leave blank

1. The discrete random variable X has probability function p(x) and cumulative distribution function F(x) given in the table below.

x	1	2	3	4	5
p(x)	0.10	a	0.28	c	0.24
F(x)	0.10	0.26	b	0.76	d

(a) Write down the value of d **(1)**

(b) Find the values of a, b and c **(3)**

(c) Write down the value of P($X > 4$) **(1)**

Two independent observations, X_1 and X_2, are taken from the distribution of X.

(d) Find the probability that X_1 and X_2 are both odd. **(2)**

Given that X_1 and X_2 are both odd,

(e) find the probability that the sum of X_1 and X_2 is 6
Give your answer to 3 significant figures. **(3)**

a) $d = 1$

b) $a = 0.26 - 0.10 = 0.16$

$b = 0.26 + 0.28 = 0.54$

$c = 0.76 - 0.54 = 0.22$

c) $P(X>4) = P(X=5) = 0.24$

d) $P(X \text{ is odd}) = P(X=1) + P(X=3) + P(X=5)$

$= 0.1 + 0.28 + 0.24 = 0.62$

$P(\text{both odd}) = 0.62 \times 0.62 = 0.3844$

Leave blank

Question 1 continued

e) Let event A be 'sum = 6', B be 'both odd'

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$P(A \cap B) = P(1,5) + P(3,3) + P(5,1)$$

$$= (0.1 \times 0.24) + (0.28 \times 0.28) + (0.24 \times 0.1)$$

$$= 0.1264$$

$$\Rightarrow P(A|B) = \frac{0.1264}{0.3844}$$

$$= 0.3288$$

Leave blank

2. A sports teacher recorded the number of press-ups done by his students in two minutes. He recorded this information for a Year 7 class and for a Year 11 class.

The back-to-back stem and leaf diagram shows this information.

Totals	Year 7 class		Year 11 class	Totals
(6)	8 7 6 5 5 4	1		
(10)	9 7 7 6 5 4 4 4 2 2	2	0 5 6 9	(4)
(7)	8 7 5 4 3 3 0	3	3 4 5 8 8	(5)
(5)	9 9 7 2 2	4	0 5 6 7 9	(5)
(3)	8 4 0	5	0 3 5 5 6 6 7 7 7 9 9	(11)
		6	0 3 3 3 3 4 8	(7)

Key: 2|4|0 means 42 press-ups for a Year 7 student and 40 press-ups for a Year 11 student

(a) Find the median number of press-ups for each class.
(2)

For the Year 11 class, the lower quartile is 38 and the upper quartile is 59

(b) Find the lower quartile and the upper quartile for the Year 7 class.
(2)

(c) Use the medians and quartiles to describe the skewness of each of the two distributions.
(3)

(d) Give two reasons why the normal distribution should not be used to model the number of press-ups done by the Year 11 class.
(2)

a) Yr 7 total $= 6+10+7+5+3 = 31$

$\therefore$ median is $\frac{31+1}{2} = 16^{th}$ score $= 29$

Yr 11 total $= 4+5+5+11+7 = 32$

$\therefore$ median is $\frac{32+1}{2} = 16.5^{th}$ score

$= \frac{53+55}{2} = 54$

Leave blank

Question 2 continued

b) $Q_1 = \frac{31+1}{4} = 8^{th}$ score $= 22$

$Q_3 = \frac{(31+1)3}{4} = 24^{th}$ score $= 42$

c) Yr 7: $Q_3 - Q_2 = 42 - 29 = 13$

$Q_2 - Q_1 = 29 - 22 = 7$

$(Q_3 - Q_2) > (Q_2 - Q_1) \Rightarrow$ positive skew

Yr 11: $Q_3 - Q_2 = 59 - 54 = 5$

$Q_2 - Q_1 = 54 - 38 = 16$

$(Q_3 - Q_2) < (Q_2 - Q_1) \Rightarrow$ negative skew

d) Distribution for no. of press-ups is skewed and discrete

Leave blank

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)**

a) $\sum bm = 29(75) + 29(83) + 35(91) + 39(121) + 41(120) + 43(126) + 44(119) + 46(126)$

$= 33856$

b) $\sum b = 75 + 83 + \ldots = 861$

$\sum m = 29 + 29 + 35 + \ldots = 306$

$S_{bm} = \sum bm - \frac{\sum b \sum m}{n} = 33856 - \frac{(861)(306)}{8}$

$= 922.75$

c) $r = \frac{S_{bm}}{\sqrt{S_{bb} S_{mm}}} = \frac{922.75}{\sqrt{3083.875 \times 305.5}} = 0.95$

Leave blank

Question 3 continued

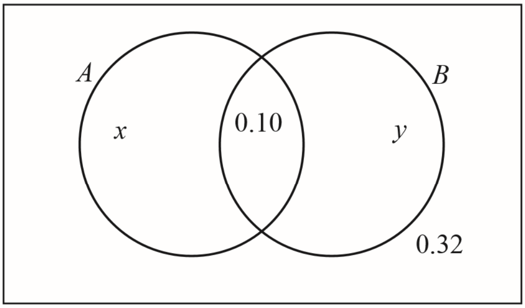
d) Strong positive correlation: The price of milk is higher in years that the price of bread is higher

e) Decrease because the new point is further away from the line of best-fit

Leave blank

4. Events A and B are shown in the Venn diagram below

where x, y, 0.10 and 0.32 are probabilities.



(a) Find an expression in terms of x for

(i) $P(A)$

(ii) $P(B \mid A)$

(3)

(b) Find an expression in terms of x and y for $P(A \cup B)$

(1)

Given that $P(A) = 2P(B)$

(c) find the value of x and the value of y

(5)

a) i) $P(A) = x + 0.10$ ii) $P(B|A) = \dfrac{0.10}{x+0.10}$

b) $P(A \cup B) = x + y + 0.10$

c) $P(A) = 2P(B)$

$x + 0.10 = 2(0.10 + y) = 0.2 + 2y$

$x = 0.1 + 2y$ (1)

Total prob. = 1

$x + 0.1 + y + 0.32 = 1$

$x = 0.58 - y$ (2)

Leave blank

Question 4 continued

$(1) = (2)$

$0.1 + 2y = 0.58 - y$

$3y = 0.48$

$y = 0.16$

In (2): $x = 0.58 - 0.16$

$= 0.42$

Leave blank

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)**

a) h, because the heart rate depends on the average length of exercise

b) For every extra minute of daily exercise, the average heart rate goes down by 0.43

Leave blank

Question 5 continued

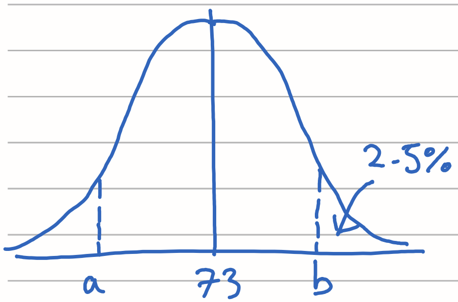
c) $\bar{t} = \frac{\Sigma t}{n} = \frac{400}{8} = 50$

$\bar{h} = \frac{\Sigma h}{n} = \frac{576}{8} = 72$

d) $93.5 - 0.43(50) = 72$

e) $h = 93.5 - 0.43(60) = 67.7$

f) It is reliable because 1 hour lies within the range of measurements

g) 

$H \sim N(73, 8^2)$

$P(H \leqslant b) = P\left(Z \leqslant \frac{b-73}{8}\right)$

$= 0.975$

From tables, $\frac{b-73}{8} = 1.96$

$b = 88.7$

By symmetry $a = 73 - (88.7 - 73)$

$= 57.3$

Leave blank

6. The random variable X has probability function

$$P(X = x) = \frac{x^2}{k} \qquad x = 1, 2, 3, 4$$

(a) Show that $k = 30$ **(2)**

(b) Find $P(X \neq 4)$ **(2)**

(c) Find the exact value of $E(X)$ **(2)**

(d) Find the exact value of $\text{Var}(X)$ **(4)**

Given that $Y = 3X - 1$

(e) find $E(Y^2)$ **(4)**

a) $\sum P(X=x) = 1$

$$\frac{1^2}{k} + \frac{2^2}{k} + \frac{3^2}{k} + \frac{4^2}{k} = 1$$

$$k = 1+4+9+16$$

$$= 30$$

b) $P(X \neq 4) = 1 - P(X=4) = 1 - \frac{4^2}{30} = \frac{7}{15}$

c) $E(X) = \sum x P(X=x)$

$$= 1 \times \frac{1^2}{30} + 2\frac{(2^2)}{30} + \frac{3(3^2)}{30} + \frac{4(4^2)}{30}$$

$$= \frac{10}{3}$$

Leave blank

Question 6 continued

d) $\mathrm{Var}(X) = E(X^2) - [E(X)]^2$

$= \sum x^2 P(X=x) - \left(\frac{10}{3}\right)^2$

$= \frac{1^2(1^2)}{30} + \frac{2^2(2^2)}{30} + \frac{3^2(3^2)}{30} + \frac{4^2(4^2)}{30} - \frac{100}{9}$

$= \frac{59}{5} - \frac{100}{9} = \frac{31}{45}$

e) $Y = 3X - 1$

$E(Y^2) = E(9X^2 - 6X + 1)$

$= 9E(X^2) - 6E(X) + E(1)$

$= 9\left(\frac{59}{5}\right) - 6\left(\frac{10}{3}\right) + 1$

$= \frac{436}{5}$

Leave blank

7. The birth weights, W grams, of a particular breed of kitten are assumed to be normally distributed with mean 99 g and standard deviation 3.6 g

(a) Find $\mathrm{P}(W > 92)$ **(3)**

(b) Find, to one decimal place, the value of k such that $\mathrm{P}(W < k) = 3\mathrm{P}(W > k)$ **(4)**

(c) Write down the name given to the value of k. **(1)**

For a different breed of kitten, the birth weights are assumed to be normally distributed with mean 120 g

Given that the 20th percentile for this breed of kitten is 116 g

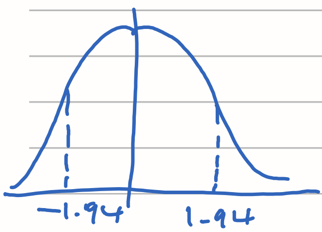
(d) find the standard deviation of the birth weight of this breed of kitten. **(3)**

a)

$W \sim N(99, 3.6^2)$

$P(W > 92) = P\left(Z > \frac{92-99}{3.6}\right)$

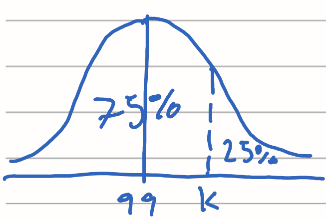
$= P(Z > -1.94)$



$= P(Z < 1.94)$

$= 0.9738$

b)



$P(W < k) = P\left(Z < \frac{k-99}{3.6}\right) = 0.75$

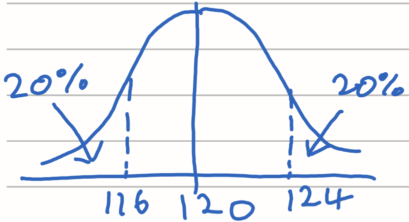
$\Rightarrow \frac{k-99}{3.6} = 0.67$

$k = 101.4$

Leave blank

Question 7 continued

d) $K \sim N(120, \sigma^2)$



$P(K > 124) = P\left(Z > \frac{124-120}{\sigma}\right)$

$= 0.2$

$\Rightarrow \frac{124-120}{\sigma} = 0.8416$

$\sigma = 4.75$