
SL Paper 1

A researcher consulted 500 men and women to see if the colour of the car they drove was independent of gender. A χ^2 test for independence was carried out.

a. Write down the null hypothesis. [1]

b. The colours of the cars were red, green, blue, black and silver. [2]

Find the number of degrees of freedom for this test.

c. At the 5 % significance level the χ^2_{calc} was found to be 8.73. [1]

Write down the critical value, χ^2_{crit} , for this test.

180 spectators at a swimming championship were asked which, of four swimming styles, was the one they preferred to watch.

The results of their responses are shown in the table.

Swimming style	Male	Female
Freestyle	20	15
Butterfly	20	30
Backstroke	10	35
Breaststroke	10	40

A χ^2 test was conducted at the 5% significance level.

a. Write down the null hypothesis for this test. [1]

b. Write down the number of degrees of freedom. [1]

c. Write down the value of χ^2_{calc} . [2]

d. The critical value, at the 5% significance level, is 7.815. [2]

State, giving a reason, the conclusion to the test.

The heights of apple trees in an orchard are normally distributed with a mean of 3.42 m and a standard deviation of 0.21 m.

- a. Write down the probability that a randomly chosen tree has a height greater than 3.42 m. [1]
- b. Write down the probability that a randomly chosen tree will be within 2 standard deviations of the mean of 3.42 m. [1]
- c. Use your graphic display calculator to calculate the probability that a randomly chosen tree will have a height greater than 3.35 m. [2]
- d. The probability that a particular tree is less than x metres high is 0.65. Find the value of x . [2]
-

Members of a certain club are required to register for one of three sports, badminton, volleyball or table tennis. The number of club members of each gender choosing each sport in a particular year is shown in the table below.

A χ^2 (Chi-squared) test at the 5% significance level is used to determine whether the choice of sport is independent of gender.

	Badminton	Volleyball	Table tennis
Male	40	20	10
Female	20	15	15

- a. Find the expected number of female volleyball players under this hypothesis. [2]
- b. Write down the p -value for the test. [2]
- c. State, with a reason, the conclusion of the test. [2]
-

Tony wants to carry out a χ^2 test to determine whether or not a person's choice of one of the three professions; engineering, medicine or law is influenced by the person's sex (gender).

- a. State the null hypothesis, H_0 , for this test. [1]
- b. Write down the number of degrees of freedom. [1]
- c. Of the 400 people Tony interviewed, 220 were male and 180 were female. 80 of the people had chosen engineering as a profession. [2]
Calculate the expected number of female engineers.
- d. Tony used a 5 % level of significance for his test and obtained a p -value of 0.0634 correct to 3 significant figures. [2]
State Tony's conclusion to the test. Give a reason for this conclusion.
-

The producer of a TV dancing show asked a group of 150 viewers their age and the type of Latin dance they preferred. The types of Latin dances in the show were Argentine tango, Samba, Rumba and Cha-cha-cha. The data obtained were organized in the following table.

	Dance			
	Argentine tango	Samba	Rumba	Cha-cha-cha
20 years old and younger	35	23	12	10
Older than 20 years old	20	17	18	15

A χ^2 test was carried out, at the 5% significance level.

- a. Write down the null hypothesis for this test. [1]
- b. Write down the observed number of viewers who preferred Rumba **and** were older than 20 years old. [1]
- c. Use your graphic display calculator to find the p -value for this test. [2]
- d. The producer claims that the type of Latin dance a viewer preferred is independent of their age. [2]

Decide whether this claim is justified. Give a reason for your decision.

A market researcher consulted males and females to determine whether the type of coffee they drink is associated with gender. The types of coffee are Cappuccino, Latte, Americano, Macchiato and Espresso. A χ^2 test was conducted, at the 5% significance level and the χ^2 value was found to be 8.73.

- a.i. Write down the null hypothesis. [1]
 - a.ii. Write down the alternative hypothesis. [1]
 - b. Write down the number of degrees of freedom for this test. [1]
 - c. Write down the critical value for this test. [1]
 - d. State whether the type of coffee drunk is independent of gender. Give a reason for your answer. [2]
-

A study was carried out to determine whether the country chosen by students for their university studies was influenced by a person's gender. A random sample was taken. The results are shown in the following table.

	Country Chosen		
	USA	Australia	UK
Male	55	26	40
Female	25	31	41

A χ^2 test was performed at the 1% significance level.

The critical value for this test is 9.210.

- a. State the null hypothesis. [1]
- b. Write down the number of degrees of freedom. [1]
- c. Write down [2]
- (i) the χ^2 statistic;
 - (ii) the associated p -value.
- d. State, giving a reason, whether the null hypothesis should be accepted. [2]
-

The number of calories a person burns during a walk depends on the time they spend walking. The table below shows the number of calories, y , burned by a person in relation to the time they spend walking, x , in minutes.

Time spent walking (x) (minutes)	10	15	20	25	30
Calories (y)	90	125	200	300	375

- a. Use your graphic display calculator to write down the equation of the regression line for y on x in the form $y = ax + b$. [2]
- b. Use your equation to estimate the number of calories that a person will burn during a 17 minute walk. [2]
- c. State whether your answer to part (b) is reliable. Give a reason for your answer. [2]
-

The mass of a certain type of Chilean corncob follows a normal distribution with a mean of 400 grams and a standard deviation of 50 grams.

A farmer labels one of these corncobs as premium if its mass is greater than a grams. 25% of these corncobs are labelled as premium.

- a. Write down the probability that the mass of one of these corncobs is greater than 400 grams. [1]
- b. Find the value of a . [2]
- c. Estimate the interquartile range of the distribution. [3]
-

A group of 100 students gave the following responses to the question of how they get to school.

	Walk	Public transport	Car	Bicycle	Total
Female	18	13	14	3	48
Male	9	17	10	16	52
Total	27	30	24	19	100

A χ^2 test for independence was conducted at the 5% significance level. The null hypothesis was defined as

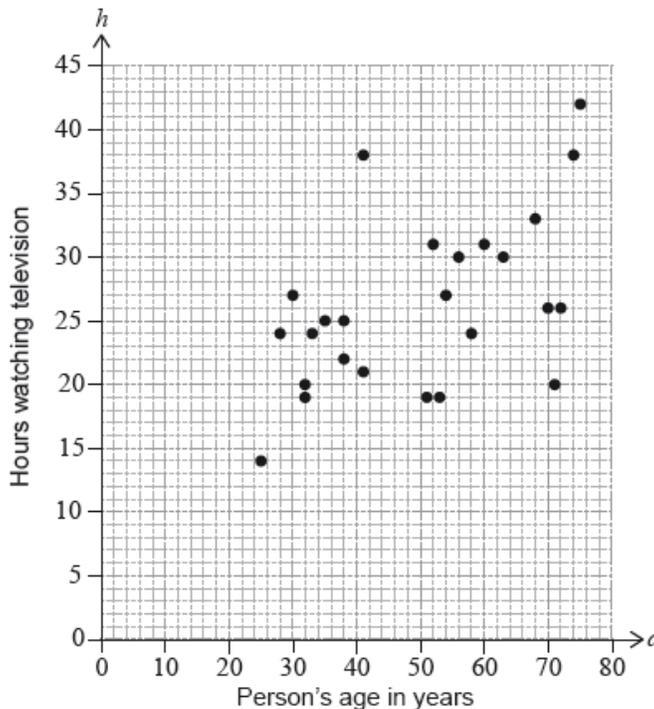
H_0 : Method of getting to school is independent of gender.

- a. Find the expected frequency for the females who use public transport to get to school. [2]
- b. Find the χ^2 statistic. [2]
- c. The χ^2 critical value is 7.815 at the 5% significance level. [2]

State whether or not the null hypothesis is accepted. Give a reason for your answer.

A survey was carried out to investigate the relationship between a person's age in years (a) and the number of hours they watch television per week (h).

h). The scatter diagram represents the results of the survey.



The mean age of the people surveyed was 50.

For these results, the equation of the regression line h on a is $h = 0.22a + 15$.

- a. Find the mean number of hours that the people surveyed watch television per week. [2]
- b. Draw the regression line on the scatter diagram. [2]

c. By placing a tick (✓) in the correct box, determine which of the following statements is true:

[1]

The correlation between h and a is positive.	
The correlation between h and a is negative.	
There is no correlation between h and a .	

d. Diogo is 18 years old. Give a reason why the regression line should not be used to estimate the number of hours Diogo watches television per week. [1]

A market researcher surveyed men and women about their preferred holiday destination. The holiday destinations were Antigua, Barbados, Cuba, Guadeloupe and Jamaica. A χ^2 test for independence was conducted at the 5 % significance level.

The χ^2 calculated value was found to be 8.73.

a. Write down the null hypothesis. [1]

b. Find the number of degrees of freedom for this test. [2]

c. Write down the critical value for this test. [1]

d. State the conclusion of this test. Give a reason for your decision. [2]

A survey investigated the relationship between the number of cleaners, n , and the amount of time, t , it takes them to clean a school.

Number of cleaners, n	Time, t (minutes)
1	193
2	172
3	118
5	112
6	87

a. Use your graphic display calculator to write down the equation of the regression line t on n . [2]

b. Write down the value of the Pearson's product-moment correlation coefficient, r . [2]

c. Use your regression equation to find the amount of time 4 cleaners take to clean the school. [2]

The marks obtained by 8 candidates in Physics and Chemistry tests are given below.

Physics (x)	6	8	10	11	10	5	4	12
Chemistry (y)	8	11	14	13	11	7	5	15

- a. Write down the product moment correlation coefficient, r . [1]
- b. Write down, in the form $y = mx + c$, the equation of the regression line y on x for the 8 candidates. [2]
- c. A ninth candidate obtained a score of 7 in the Physics test but was absent for the Chemistry test. [2]
Use your answer to (b) to estimate the score he would have obtained on the Chemistry test.
- d. Give a reason why it is valid to use this regression line to estimate the score on the Chemistry test in part (c). [1]

Applicants for a job had to complete a mathematics test. The time they took to complete the test is normally distributed with a mean of 53 minutes and a standard deviation of 16.3. One of the applicants is chosen at random.

For 11% of the applicants it took longer than k minutes to complete the test.

There were 400 applicants for the job.

- a. Find the probability that this applicant took at least 40 minutes to complete the test. [2]
- b. Find the value of k . [2]
- c. Estimate the number of applicants who completed the test in less than 25 minutes. [2]

Tania wishes to see whether there is any correlation between a person's age and the number of objects on a tray which could be remembered after looking at them for a certain time.

She obtains the following table of results.

Age (x years)	15	21	36	40	44	55
Number of objects remembered (y)	17	20	15	16	17	12

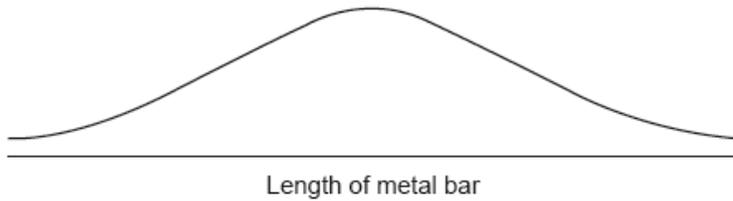
- a. Use your graphic display calculator to find the equation of the regression line of y on x . [2]
- b. Use your equation to estimate the number of objects remembered by a person aged 28 years. [1]
- c. Use your graphic display calculator to find the correlation coefficient r . [1]

d. Comment on your value for r .

[2]

A factory makes metal bars. Their lengths are assumed to be normally distributed with a mean of 180 cm and a standard deviation of 5 cm.

a. On the following diagram, shade the region representing the probability that a metal bar, chosen at random, will have a length less than 175 cm. [2]



b. A metal bar is chosen at random.

[4]

- (i) The probability that the length of the metal bar is less than 175 cm is equal to the probability that the length is greater than h cm. Write down the value of h .
- (ii) Find the probability that the length of the metal bar is greater than one standard deviation above the mean.

Minta surveyed students from her school about their preferred morning snack from a choice of an apple, a fruit salad or a smoothie.

She surveyed 350 students, of whom 210 are female.

She performed a χ^2 test at the 5% significance level to determine whether there is a relationship between the choice of morning snack and gender.

a. State Minta's null hypothesis.

[1]

b. State the number of degrees of freedom.

[1]

c. 150 students showed a preference for a smoothie.

[2]

Calculate the expected number of female students who chose a smoothie.

d. Minta found that the calculated value of the χ^2 test was 3.576. The critical value at the 5% significance level is 5.99.

[2]

State Minta's conclusion. Give a reason for your answer.

In a school, students in grades 9 to 12 were asked to select their preferred drink. The choices were milk, juice and water. The data obtained are organized in the following table.

	Milk	Juice	Water	Total
Grade 9	25	34	15	74
Grade 10	31	x	13	74
Grade 11	18	35	17	70
Grade 12	9	36	26	71
Total	83	135	71	289

A χ^2 test is carried out at the 5% significance level with hypotheses:

H_0 : the preferred drink is independent of the grade

H_1 : the preferred drink is not independent of the grade

The χ^2 critical value for this test is 12.6.

- Write down the value of x . [1]
- Write down the number of degrees of freedom for this test. [1]
- Use your graphic display calculator to find the χ^2 statistic for this test. [2]
- State the conclusion for this test. Give a reason for your answer. [2]

A shop keeper recorded daily sales s of ice cream along with the daily maximum temperature t °C. The results for one week are shown below.

t	29	31	34	23	19	20	27
s	104	92	112	48	56	72	66

- Write down the equation of the regression line for s on t . [3]
- Use your equation to predict the ice cream sales on a day when the maximum temperature is 24 °C. Give your answer correct to the nearest whole number. [3]

The scores obtained by five candidates in Mathematics and Physics examinations are given below.

Mathematics (x)	62	84	47	55	32
Physics (y)	80	91	44	48	53

- Write down the correlation coefficient, r , for the examination scores. [2]
- Write down the equation of the regression line, y on x , for the examination scores of the five candidates. [2]

c. A sixth candidate scored 72 in the Mathematics examination. Use the regression line, y on x , to estimate his score on the Physics examination. [2]

Identical mosquito traps are placed at different distances from a lake. On one day the number of mosquitoes caught in 10 of the traps is recorded.

Distance, m (x)	8	15	22	30	34	45	50	60	74	82
Number of mosquitoes (y)	78	75	72	67	66	59	59	53	48	43

It is believed the number of mosquitoes caught varies linearly with the distance, in metres, of the trap from the lake.

a. Find [4]

- (i) Pearson's product-moment correlation coefficient, r ;
- (ii) the equation of the regression line y on x .

b. Use the equation of the regression line y on x to estimate the number of mosquitoes caught in a trap that is 28 m from the lake. [2]

200 people of different ages were asked to choose their favourite type of music from the choices Popular, Country and Western and Heavy Metal.

The results are shown in the table below.

Age/Music choice	Popular	Country and Western	Heavy Metal	Totals
11 – 25	35	5	50	90
26 – 40	30	10	20	60
41 – 60	20	25	5	50
Totals	85	40	75	200

It was decided to perform a chi-squared test for independence at the 5% level on the data.

a. Write down the null hypothesis. [1]

b. Write down the number of degrees of freedom. [1]

c. Write down the chi-squared value. [2]

d. State whether or not you will reject the null hypothesis, giving a clear reason for your answer. [2]

A questionnaire was given to all members of a school community to find out which drink was the most popular to have with breakfast. The results are given in the table below, classified by age.

	Hot Chocolate	Tea	Coffee	Milk
Children aged 12 years and less	55	10	1	34
Teenagers aged from 13 to 19 years	25	35	20	10
Adults aged 20 years and over	20	40	79	6

A χ^2 test was conducted to decide whether the type of drink was independent of age.

- Find the number of degrees of freedom for the χ^2 test. [2]
- Write down the null hypothesis for the χ^2 test. [1]
- The critical value for the χ^2 test at the 5% significance level is 12.59. The χ^2 test statistic is calculated to be 146 with a p -value of 6.62×10^{-29} [2]
(both numbers given correct to 3 significant figures).

Write down the conclusion reached at the 5 % significance level. Give a clear reason for your answer.

A hospital collected data from 1000 patients in four hospital wards to review the quality of its healthcare. The data, showing the number of patients who became infected during their stay in hospital, was recorded in the following table.

	Ward				Total
	Pasteur	Nightingale	Jenner	Fleming	
Patients infected	44	27	13	16	100
Patients not infected	281	303	182	134	900
Total	325	330	195	150	1000

A χ^2 -test was performed at the 5% significance level.

The critical value for this test is 7.815.

The null hypothesis for the test is

H_0 : Becoming infected during a stay in the hospital is independent of the ward.

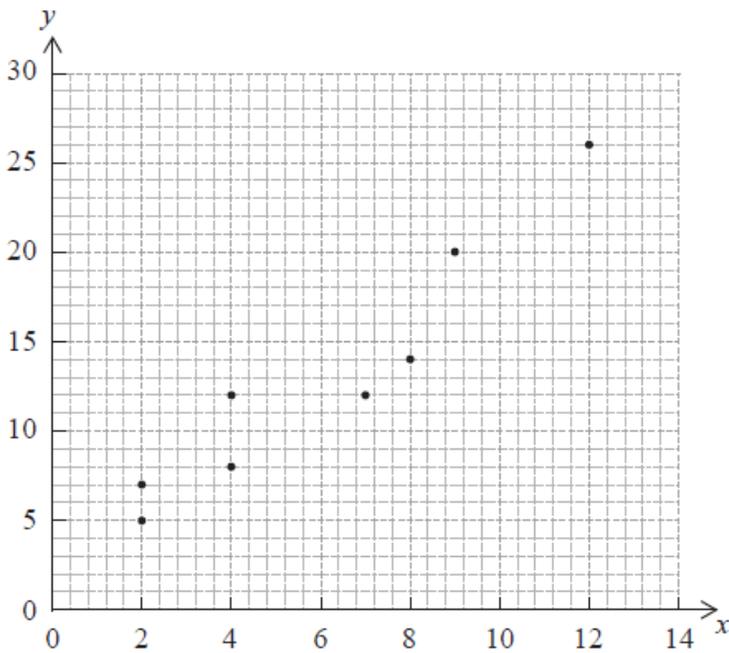
- Find the expected frequency of the patients who became infected whilst in Nightingale ward. [2]
- For this test, write down the χ^2 statistic. [2]
- State, giving a reason, whether the null hypothesis should be rejected. [2]

The daily January temperature of Cairns is normally distributed with a mean of 34°C and a standard deviation of 3.

- a. Calculate the probability that the temperature on a randomly chosen day in January is less than 39°C . [2]
- b. Calculate the expected number of days in January that the temperature will be more than 39°C . [2]
- c. On a randomly chosen day in January, the probability that the temperature is above $T^{\circ}\text{C}$ is 0.7. [2]
- Find the value of T .

Consider the following set of data which is plotted on the scatter diagram below.

x	2	4	7	12	4	8	9	2
y	5	8	12	26	12	14	20	7



- a. Write down the coordinates of the mean point (\bar{x}, \bar{y}) . [2]
- b. Write down the value of r , the Pearson's product-moment correlation coefficient for this set of data. [2]
- c. Draw the regression line for y on x on the set of axes above. [2]

- a. The manager of a travel agency surveyed 1200 travellers. She wanted to find out whether there was a relationship between a traveller's age and their preferred destination. The travellers were asked to complete the following survey. [1]

Traveller survey

My age is:

25 or younger	26–40	41–60	61 or older

My preferred destination is:

New York	Tokyo	Melbourne	Dubai	Marrakech

A χ^2 test was carried out, at the 5% significance level, on the data collected.

Write down the null hypothesis.

b. Find the number of degrees of freedom. [2]

c. The critical value of this χ^2 test is 21.026. [1]

Use this information to write down the values of the χ^2 statistic for which the null hypothesis is rejected.

d. From the travellers taking part in the survey, 285 were 61 years or older and 420 preferred Tokyo. [2]

Calculate the expected number of travellers who preferred Tokyo and were 61 years or older.

A scientist measures the concentration of dissolved oxygen, in milligrams per litre (y), in a river. She takes 10 readings at different temperatures, measured in degrees Celsius (x).

The results are shown in the table.

Temperature, °C (x)	20	24	25	28	29	32	27	25	23	21
Dissolved Oxygen, mg l ⁻¹ (y)	10.9	9.7	9.2	7.6	7.3	6.4	7.9	8.4	9.4	9.9

It is believed that the concentration of dissolved oxygen in the river varies linearly with the temperature.

a.i. For these data, find Pearson's product-moment correlation coefficient, r . [2]

a.ii. For these data, find the equation of the regression line y on x . [2]

b. Using the equation of the regression line, estimate the concentration of dissolved oxygen in the river when the temperature is 18 °C. [2]

a. A survey was conducted among a random sample of people about their favourite TV show. People were classified by gender and by TV show preference (Sports, Documentary, News and Reality TV). [2]

The results are shown in the contingency table below.

	Sports	Documentary	News	Reality TV	Total
Male	20	24	32	11	87
Female	18	30	20	25	93
Total	38	54	52	36	180

Find the expected number of females who prefer documentary shows.

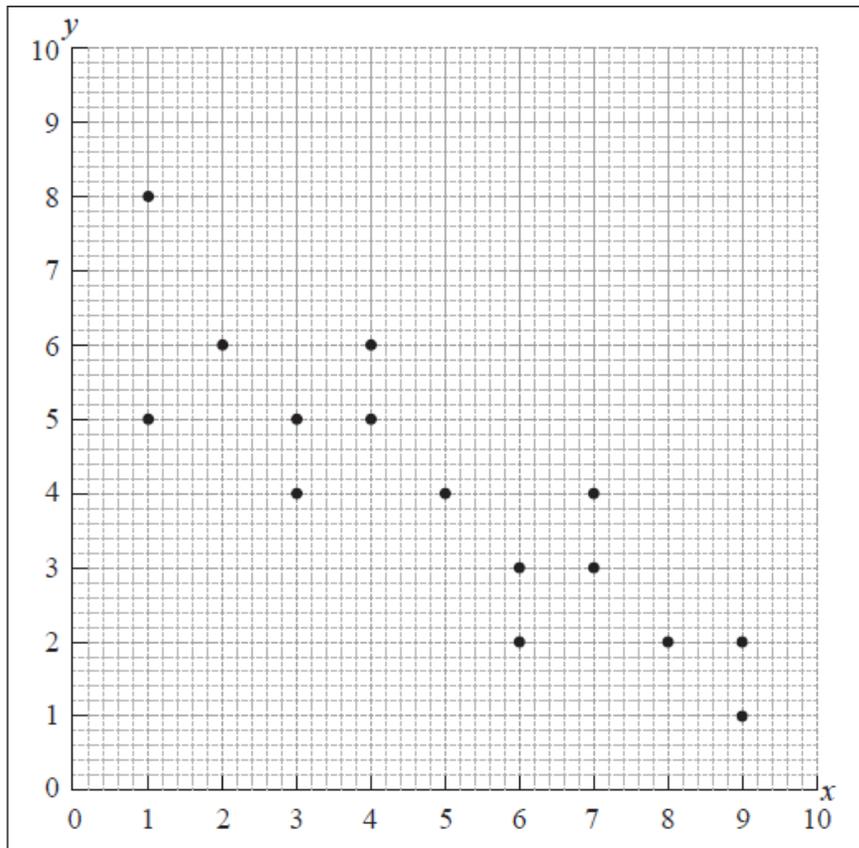
b. A χ^2 test at the 5% significance level is used to determine whether TV show preference is independent of gender. [2]

Write down the p -value for the test.

c. State the conclusion of the test. Give a reason for your answer. [2]

Consider the following values of x and y and the scatter diagram which represents the information given in the table.

x	1	1	2	3	3	4	4	b	6	6	7	7	8	9	9
y	5	a	6	4	5	5	6	4	2	3	3	4	2	1	2



a. Write down the value of

(i) a ;

(ii) b .

[2]

b. The mean of the x values is 5 and the mean of the y values is 4. Draw the line of best fit on the scatter diagram above. [2]

c. Use your line of best fit to estimate the value of y when $x = 6.5$. [2]

The local park is used for walking dogs. The sizes of the dogs are observed at different times of the day. The table below shows the numbers of dogs present, classified by size, at three different times last Sunday.

	Small	Medium	Large
Morning	9	18	21
Afternoon	11	6	13
Evening	7	8	9

a. Write a suitable null hypothesis for a χ^2 test on this data. [1]

b. Write down the value of χ^2 for this data. [2]

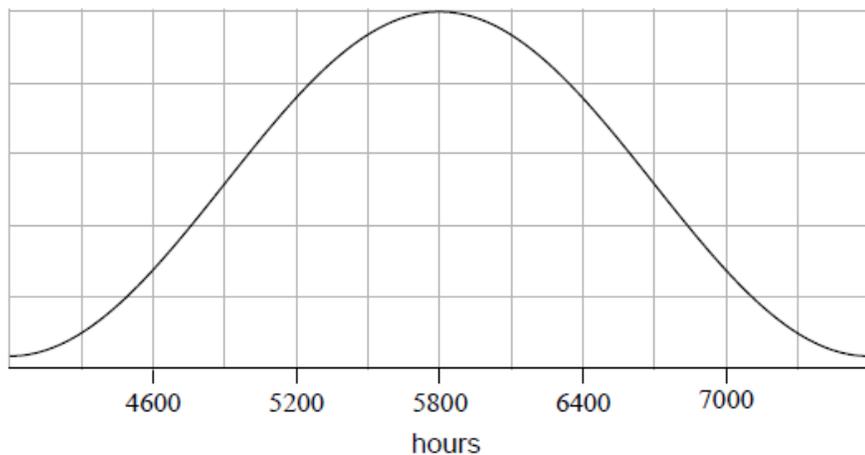
c. The number of degrees of freedom is 4. Show how this value is calculated. [1]

d. The critical value, at the 5% level of significance, is 9.488. [2]

What conclusion can be drawn from this test? Give a reason for your answer.

a. The lifetime, L , of light bulbs made by a company follows a normal distribution. [1]

L is measured in hours. The normal distribution curve of L is shown below.



Write down the mean lifetime of the light bulbs.

b. The standard deviation of the lifetime of the light bulbs is 850 hours. [2]

Find the probability that $5000 \leq L \leq 6000$, for a randomly chosen light bulb.

c. The company states that 90% of the light bulbs have a lifetime of at least k hours. [3]

Find the value of k . Give your answer correct to the nearest hundred.

The weight, W , of bags of rice follows a normal distribution with mean 1000 g and standard deviation 4 g.

a. Find the probability that a bag of rice chosen at random weighs between 990 g and 1004 g. [2]

b. 95% of the bags of rice weigh less than k grams. [2]

Find the value of k .

c. For a bag of rice chosen at random, $P(1000 - a < W < 1000 + a) = 0.9$. [2]

Find the value of a .

Malthouse school opens at 08:00 every morning.

The daily arrival times of the 500 students at Malthouse school follow a normal distribution. The mean arrival time is 52 minutes after the school opens and the standard deviation is 5 minutes.

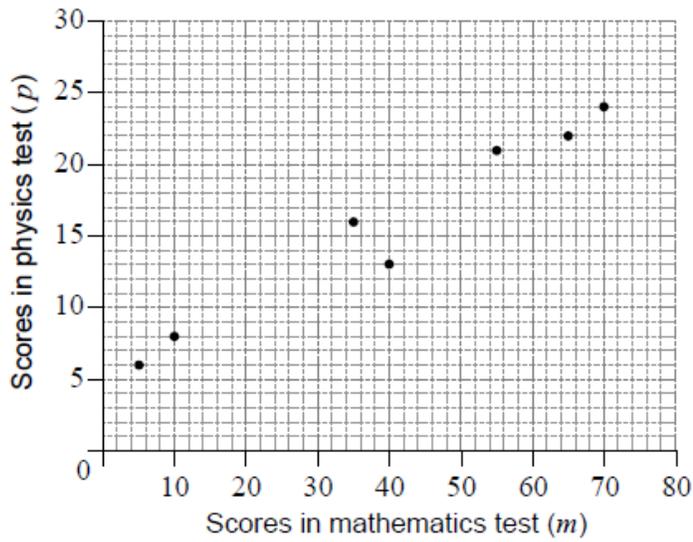
a.i. Find the probability that a student, chosen at random arrives at least 60 minutes after the school opens. [2]

a.ii. Find the probability that a student, chosen at random arrives between 45 minutes and 55 minutes after the school opens. [2]

b. A second school, Mulberry Park, also opens at 08:00 every morning. The arrival times of the students at this school follows exactly the same distribution as Malthouse school. [2]

Given that, on one morning, 15 students arrive at least 60 minutes after the school opens, estimate the number of students at Mulberry Park school.

The following scatter diagram shows the scores obtained by seven students in their mathematics test, m , and their physics test, p .

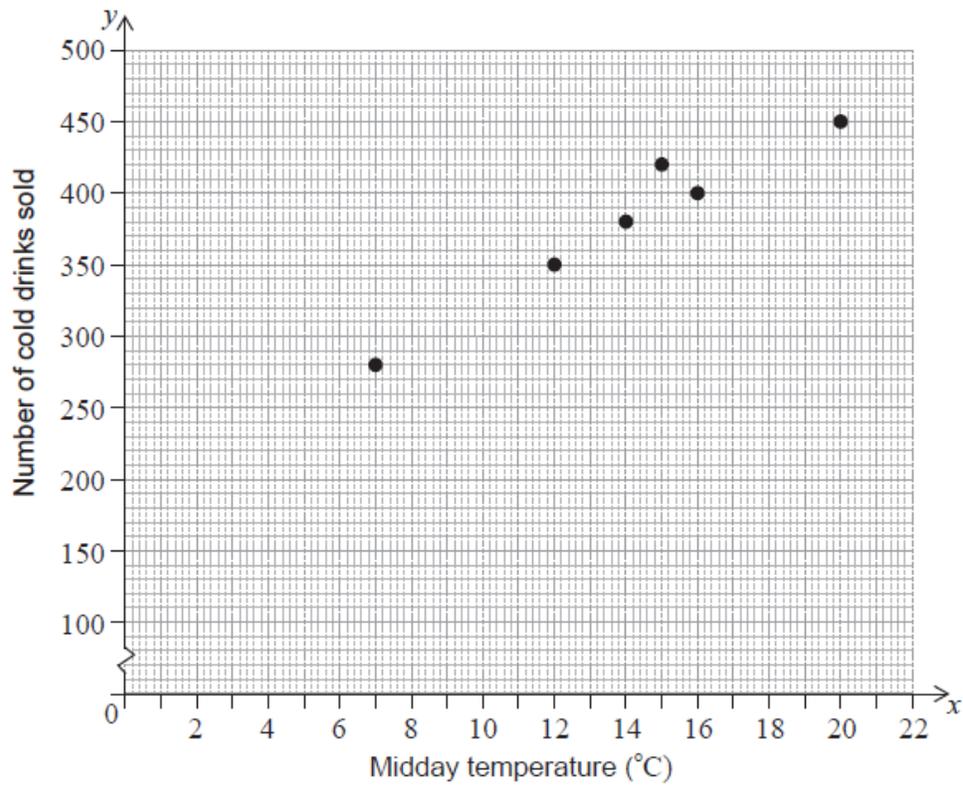


The mean point, M , for these data is $(40, 16)$.

- Plot and label the point $M(\bar{m}, \bar{p})$ on the scatter diagram. [2]
- Draw the line of best fit, by eye, on the scatter diagram. [2]
- Using your line of best fit, estimate the physics test score for a student with a score of 20 in their mathematics test. [2]

- Each day a supermarket records the midday temperature and how many cold drinks are sold on that day. The following table shows the supermarket's data for the last 6 days. This data is also shown on a scatter diagram. [2]

Midday temperature, °C (x)	7	12	14	15	16	20
Number of cold drinks sold (y)	280	350	380	420	400	450



Write down

- i) the mean temperature, \bar{x} ;
- ii) the mean number of cold drinks sold, \bar{y} .

b. Draw the line of best fit on the scatter diagram. [2]

c. Use the line of best fit to estimate the number of cold drinks that are sold on a day when the midday temperature is 10 °C. [2]