Bronze Questions

Calculator

The total mark for this section is 27

Q1

A scientist is researching whether or not birds of prey exposed to pollutants lay eggs with thinner shells. He collects a random sample of egg shells from each of 6 different nests and tests for pollutant level, p, and measures the thinning of the shell, t. The results are shown in the table below.

p	3	8	30	25	15	12
t	1	3	9	10	5	6

[You may use $\sum p^2 = 1967$ and $\sum pt = 694$]

(a) On graph paper, draw a scatter diagram to represent these data.

(2)

(b) Explain why a linear regression model may be appropriate to describe the relationship between p and t.

(1)

(c) The scientist reviews similar studies and finds that pollutant levels above 16 are likely to result in an increased risk of infection soon after hatching.

Given that $t = a + bp$.

Estimate the minimum thinning of the shell that is likely to result in an increased risk of infection.

(2)

(Total for Question 1 is 5 marks)

Q2

A random sample of 15 days is taken from the large data set for Perth in June and July 1987.

The scatter diagram in Figure 1 displays the values of two of the variables for these 15 days.

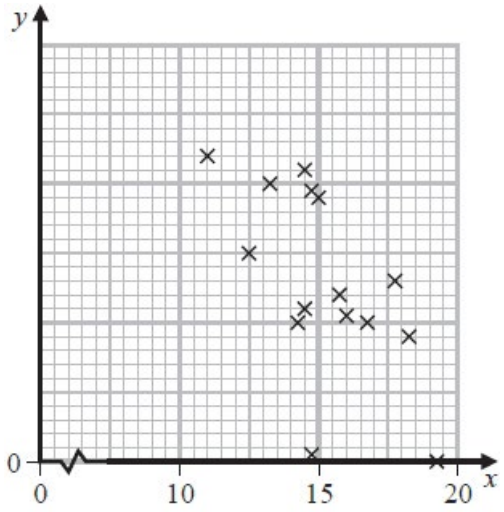


Figure 1

(a) Describe the correlation.

(1)

The variable on the x-axis is Daily Mean Temperature measured in °C.

(b) Using your knowledge of the large data set,

(i) suggest which variable is on the y-axis,
(ii) state the units that are used in the large data set for this variable.

(2)

Stav believes that there is a correlation between Daily Total Sunshine and Daily Maximum Relative Humidity at Heathrow.

He calculates the product moment correlation coefficient between these two variables for a random sample of 30 days and obtains $r = -0.377$.

(c) Carry out a suitable test to investigate Stav's belief at a 5% level of significance.

State clearly

- your hypotheses
- your critical value

(3)

On a random day at Heathrow the Daily Maximum Relative Humidity was 97%

(d) Comment on the number of hours of sunshine you would expect on that day,

giving a reason for your answer.

(1)

(Total for Question 2 is 7 marks)

Q3

A meteorologist believes that there is a relationship between the daily mean windspeed, w kn, and the daily mean temperature, t °C. A random sample of 9 consecutive days is taken from past records from a town in the UK in July and the relevant data is given in the table below.

t	13.3	16.2	15.7	16.6	16.3	16.4	19.3	17.1	13.2
w	7	11	8	11	13	8	15	10	11

The meteorologist calculated the product moment correlation coefficient for the 9 days and obtained $r = 0.609$

(a) Explain why a linear regression model based on these data is unreliable on a day when the mean temperature is 24 °C

(1)

(b) State what is measured by the product moment correlation coefficient.

(1)

(c) Stating your hypotheses clearly test, at the 5% significance level, whether or not the product moment correlation coefficient for the population is greater than zero.

(3)

Using the same 9 days a location from the large data set gave $\overline{t} = 27.2$ and $\overline{w} = 3.5$.

(d) Using your knowledge of the large data set, suggest, giving your reason, the location that gave rise to these statistics.

(1)

(Total for Question 3 is 6 marks)

Q4

Barbara is investigating the relationship between average income (GDP per capita), x US dollars, and average annual carbon dioxide (CO_2) emissions, y tonnes, for different countries.

She takes a random sample of 24 countries and finds the product moment correlation coefficient between average annual CO_2 emissions and average income to be 0.446.

(a) Stating your hypotheses clearly, test, at the 5% level of significance, whether or not the product moment correlation coefficient for all countries is greater than zero.

(3)

Barbara believes that a non-linear model would be a better fit to the data.
She codes the data using the coding $m = \log_{10} x$ and $c = \log_{10} y$ and obtains the model $c = -1.82 + 0.89m$.

The product moment correlation coefficient between c and m is found to be 0.882.

(b) Explain how this value supports Barbara's belief.

(1)

(c) Show that the relationship between y and x can be written in the form $y = ax^n$ where a and n are constants to be found.

(5)

(Total for Question 4 is 9 marks)

End of Questions

Silver Questions

Calculator

The total mark for this section is 27

Q1

In a controlled experiment, the number of microbes, N, present in a culture T days after the start of the experiment were counted.

N and T are expected to satisfy a relationship of the form

$$N = aT^b, \quad \text{where } a \text{ and } b \text{ are constants.}$$

(a) Show that this relationship can be expressed in the form

$$\log_{10}N = m\log_{10}T + c$$

giving m and c in terms of the constants a and/or b.

(2)

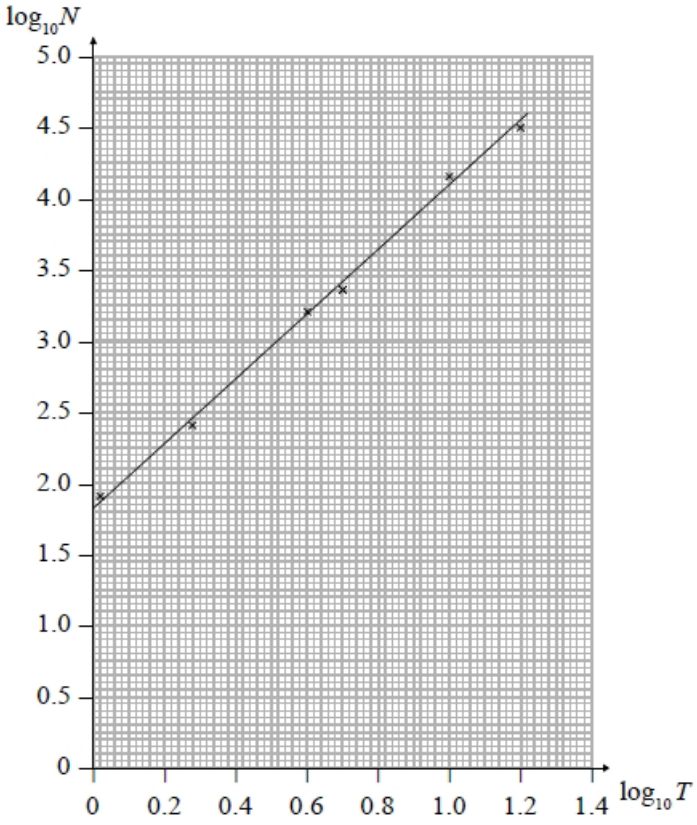


Figure 3

Figure 3 shows the line of best fit for values of $\log_{10}N$ plotted against values of $\log_{10}T$.

(b) Use the information provided to estimate the number of microbes present in the culture 3 days after the start of the experiment.

(4)

(c) Explain why the information provided could not reliably be used to estimate the day when the number of microbes in the culture first exceeds 1 000 000.

(2)

(d) With reference to the model, interpret the value of the constant a.

(1)

(Total for Question 1 is 9 marks)

Q2

The value, £V, of a vintage car t years after it was first valued on 1st January 2001, is modelled by the equation

$$V = Ap^t \qquad \text{where } A \text{ and } p \text{ are constants.}$$

Given that the value of the car was £32 000 on 1st January 2005 and £50 000 on 1st January 2012

(a) (i) find p to 4 decimal places,

(ii) show that A is approximately 24 800.

(4)

(b) With reference to the model, interpret

(i) the value of the constant A,

(ii) the value of the constant p.

(2)

Using the model,

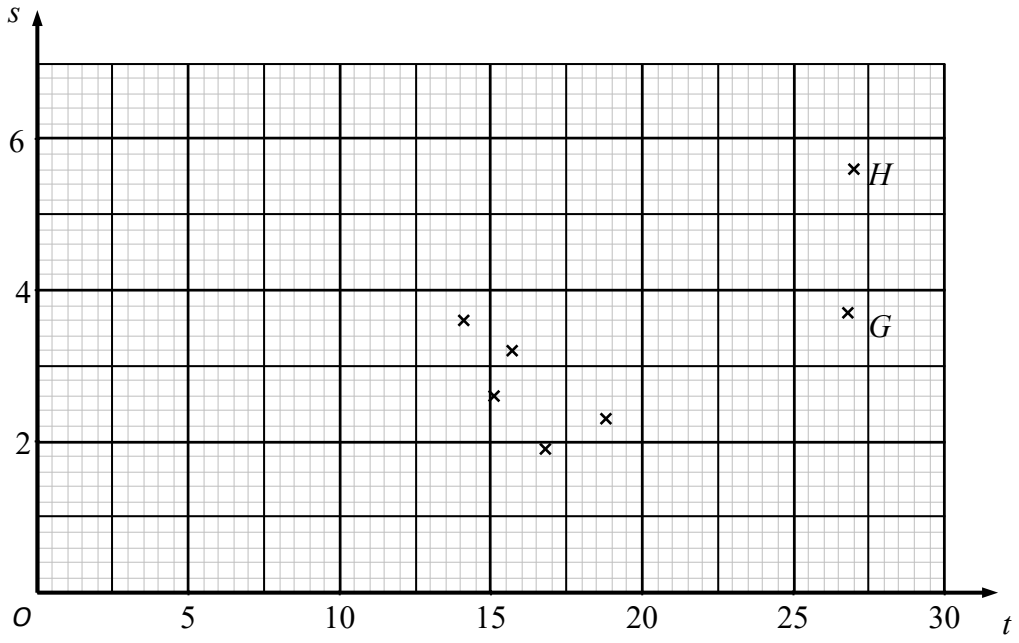
(c) find the year during which the value of the car first exceeds £100 000.

(4)

(Total for Question 2 is 10 marks)

Q3

A researcher believes that there is a linear relationship between daily mean temperature and daily total rainfall. The 7 places in the northern hemisphere from the large data set are used. The mean of the daily mean temperatures, t °C, and the mean of the daily total rainfall, s mm, for the month of July in 2015 are shown on the scatter diagram below.



(a) With reference to the scatter diagram, explain why a linear regression model may not be suitable for the relationship between t and s.

(1)

The researcher calculated the product moment correlation coefficient for the 7 places and obtained $r = 0.658$.

(b) Stating your hypotheses clearly, test at the 10% level of significance, whether or not the product moment correlation coefficient for the population is greater than zero.

(3)

(c) Using your knowledge of the large data set, suggest the names of the 2 places labelled G and H.

(1)

(d) Using your knowledge from the large data set, and with reference to the locations of the two places labelled G and H, give a reason why these places have the highest temperatures in July.

(2)

(e) Suggest how you could make better use of the large data set to investigate the relationship between daily mean temperature and daily total rainfall.

(1)

(Total for Question 3 is 7 marks)

Gold Questions

Calculator

The total mark for this section is 24 marks

Q1

An ornithologist believes that there is a relationship between the tail length, t mm, and the wing length, w mm, of female hook-billed kites. A random sample of size 10 is taken from a database of these kites and the relevant data is given in the table below.

***t* (mm)**	191	197	208	180	188	210	196	191	179	208
***w* (mm)**	284	285	288	273	280	283	288	271	257	289

The ornithologist plans to use a linear regression model based on these data and interpolate or extrapolate as necessary to estimate the wing length of other female hook-billed kites from their tail length.

(a) (i) Explain what is meant by extrapolation. **(1)**

(ii) Explain the dangers of extrapolation. **(1)**

The ornithologist attempts to calculate the product moment correlation coefficient, r, and obtains a value of 1.3.

(b) Explain how she should be able to identify that this is incorrect without carrying out any further calculations.

(1)

(c) Use your calculator to find the correct value of the product moment correlation coefficient, r.

(1)

(d) Stating your hypotheses clearly test, at the 1% significance level, whether or not there is evidence that the product moment correlation coefficient for the population is positive.

(3)

(e) Explain what your test in part (d) suggests about female hook-billed kites.

(1)

(Total for Question 1 is 8 marks)

Q2

Tessa owns a small clothes shop in a seaside town. She records the weekly sales figures, £ w, and the average weekly temperature, t °C, for 8 weeks during the summer.
The product moment correlation coefficient for these data is −0.915.
(a) Stating your hypotheses clearly and using a 5% level of significance, test whether or not the correlation between sales figures and average weekly temperature is negative.

(3)

(b) Suggest a possible reason for this correlation.

(1)

Tessa suggests that a linear regression model could be used to model these data.

(c) State, giving a reason, whether or not the correlation coefficient is consistent with Tessa's suggestion.

(1)

(d) State, giving a reason, which variable would be the explanatory variable.

(1)

Tessa calculated the linear regression equation as $w = 10\ 755 - 171t$

(e) Give an interpretation of the gradient of this regression equation.

(1)

(Total for Question 2 is 7 marks)

Q3

A research engineer is testing the effectiveness of the braking system of a car when it is driven in wet conditions.

The engineer measures and records the braking distance, d metres, when the brakes are applied from a speed of V km h^{-1}.

Graphs of d against V and $\log_{10} d$ against $\log_{10} V$ were plotted.

The results are shown below together with a data point from each graph.

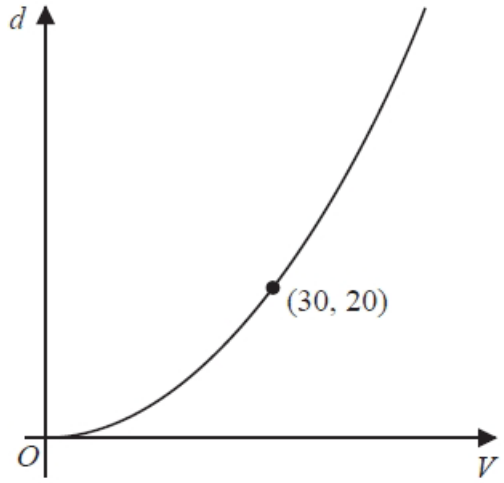


Figure 5

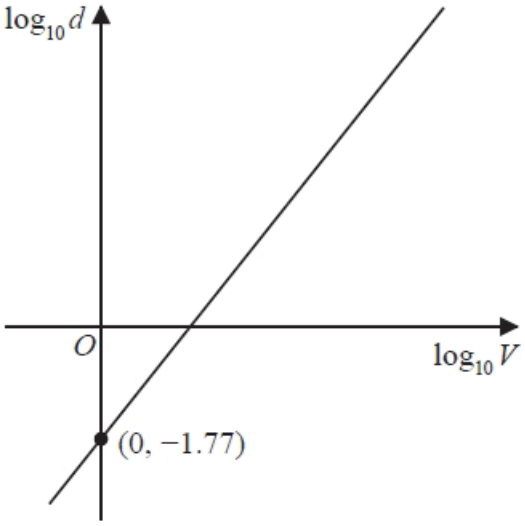


Figure 6

(a) Explain how Figure 6 would lead the engineer to believe that the braking distance should be modelled by the formula

$$d = kV^n \qquad \text{where } k \text{ and } n \text{ are constants}$$

with $k \approx 0.017$.

(3)

Using the information given in Figure 5, with $k = 0.017$,

(b) find a complete equation for the model giving the value of n to 3 significant figures.

(3)

Sean is driving this car at 60 km h^{-1} in wet conditions when he notices a large puddle in the road 100 m ahead. It takes him 0.8 seconds to react before applying the brakes.

(c) Use your formula to find out if Sean will be able to stop before reaching the puddle.

(3)

(Total for Question 3 is 9 marks)

End of questions

Bronze Questions

Calculator

The total mark for this section is 37

Q1

Three bags, *A*, *B* and *C*, each contain 1 red marble and some green marbles.

Bag *A* contains 1 red marble and 9 green marbles only
Bag *B* contains 1 red marble and 4 green marbles only
Bag *C* contains 1 red marble and 2 green marbles only

Sasha selects at random one marble from bag *A*.
If he selects a red marble, he stops selecting.
If the marble is green, he continues by selecting at random one marble from bag *B*.
If he selects a red marble, he stops selecting.
If the marble is green, he continues by selecting at random one marble from bag *C*.

(a) Draw a tree diagram to represent this information.

(2)

(b) Find the probability that Sasha selects 3 green marbles.

(2)

(c) Find the probability that Sasha selects at least 1 marble of each colour.

(2)

(d) Given that Sasha selects a red marble, find the probability that he selects it from bag *B*.

(2)

(Total for Question 1 is 8 marks)

Q2

A survey of the reading habits of some students revealed that, on a regular basis, 25% read quality newspapers, 45% read tabloid newspapers and 40% do not read newspapers at all.

(a) Find the proportion of students who read both quality and tabloid newspapers.

(3)

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

(3)

A student is selected at random. Given that this student reads newspapers on a regular basis,

(c) find the probability that this student only reads quality newspapers.

(3)

(Total for Question 2 is 9 marks)

Q3

On a randomly chosen day the probability that Bill travels to school by car, by bicycle or on foot is $\frac{1}{2}$, $\frac{1}{6}$ and $\frac{1}{3}$ respectively. The probability of being late when using these methods of travel is $\frac{1}{5}$, $\frac{2}{5}$ and $\frac{1}{10}$ respectively.

(a) Draw a tree diagram to represent this information.

(3)

(b) Find the probability that on a randomly chosen day

(i) Bill travels by foot and is late,

(ii) Bill is not late.

(4)

(c) Given that Bill is late, find the probability that he did not travel on foot.

(4)

(Total for Question 3 is 11 marks)

Q4

A and B are two events such that

$$P(B) = \frac{1}{2} \quad P(A \mid B) = \frac{2}{5} \quad P(A \cup B) = \frac{13}{20}$$

(a) Find $P(A \cap B)$.

(2)

(b) Draw a Venn diagram to show the events A, B and all the associated probabilities.

(3)

Find

(c) $P(A)$

(1)

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

(2)

(e) $P(A' \cap B)$

(1)

(Total for Question 4 is 9 marks)

End of questions

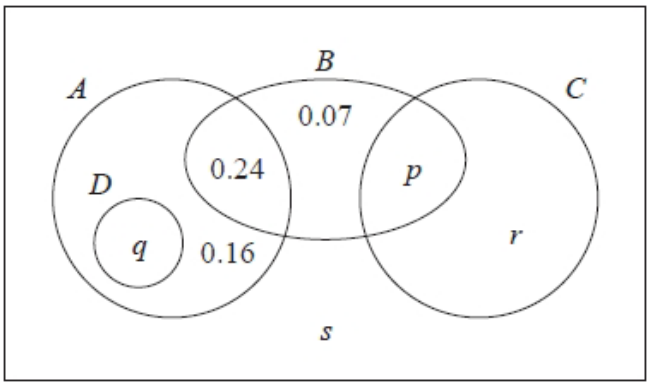
Silver Questions

Calculator

The total mark for this section is 34

Q1

The Venn diagram shows the probabilities associated with four events, A, B, C and D.



(a) Write down any pair of mutually exclusive events from A, B, C and D.

(1)

Given that P(B) = 0.4

(b) find the value of p.

(1)

Given also that A and B are independent

(c) find the value of q.

(2)

Given further that P(B' | C) = 0.64

(d) find

(i) the value of r,
(ii) the value of s.

(4)

(Total for Question 1 is 8 marks)

Q2

For the events A and B,

$$P(A' \cap B) = 0.22 \text{ and } P(A' \cap B') = 0.18 .$$

(a) Find $P(A)$.

(1)

(b) Find $P(A \cup B)$.

(1)

Given that $P(A \mid B) = 0.6$

(c) find $P(A \cap B)$.

(3)

(d) Determine whether or not A and B are independent.

(2)

(Total for Question 2 is 7 marks)

Q3

A disease is known to be present in 2% of a population. A test is developed to help determine whether or not someone has the disease.

Given that a person has the disease, the test is positive with probability 0.95.

Given that a person does not have the disease, the test is positive with probability 0.03.

(a) Draw a tree diagram to represent this information.

(3)

A person is selected at random from the population and tested for this disease.

(b) Find the probability that the test is positive.

(3)

A doctor randomly selects a person from the population and tests him for the disease. Given that the test is positive,

(c) find the probability that he does not have the disease.

(2)

(d) Comment on the usefulness of this test.

(1)

(Total for Question 3 is 9 marks)

Q4

The Venn diagram in Figure 1 shows the number of students in a class who read any of 3 popular magazines A, B and C.

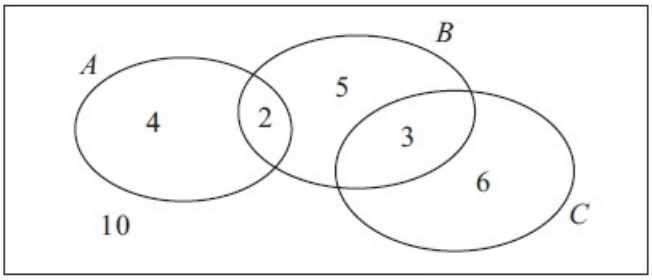


Figure 1

One of these students is selected at random.

(a) Show that the probability that the student reads more than one magazine is $\frac{1}{6}$.

(2)

(b) Find the probability that the student reads A or B (or both).

(2)

(c) Write down the probability that the student reads both A and C.

(1)

Given that the student reads at least one of the magazines,

(d) find the probability that the student reads C.

(2)

(e) Determine whether or not reading magazine B and reading magazine C are statistically independent.

(3)

(Total for Question 4 is 10 marks)

End of questions

Gold Questions

Calculator

The total mark for this section is 30

Q1

Given that

$$P(A) = 0.35 \qquad P(B) = 0.45 \qquad \text{and} \qquad P(A \cap B) = 0.13$$

find

(a) $P(A' \mid B')$.

(2)

(b) Explain why the events A and B are not independent.

(1)

The event C has $P(C) = 0.20$.

The events A and C are mutually exclusive and the events B and C are statistically independent.

(c) Draw a Venn diagram to illustrate the events A, B and C, giving the probabilities for each region.

(5)

(d) Find $P([B \cup C]')$.

(2)

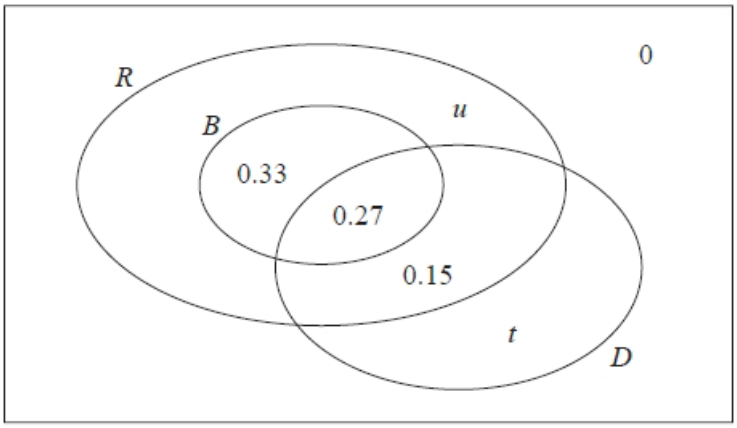
(Total for Question 1 is 10 marks)

Q2

The Venn diagram shows the probabilities of customer bookings at Harry's hotel.

R is the event that a customer books a room
B is the event that a customer books breakfast
D is the event that a customer books dinner

u and t are probabilities.



(a) Write down the probability that a customer books breakfast but does not book a room.

(1)

Given that the events B and D are independent

(b) find the value of t,

(4)

(c) hence find the value of u.

(2)

(d) Find

(i) $P(D|R \cap B)$,
(ii) $P(D|R \cap B')$.

(4)

A coach load of 77 customers arrive at Harry's hotel.

Of these 77 customers

40 have booked a room and breakfast
37 have booked a room without breakfast

(e) Estimate how many of these 77 customers will book dinner.

(2)

(Total for Question 2 is 13 marks)

Q3

(a) Given that P(A) = a and P(B) = b express P($A \cup B$) in terms of a and b when

(i) A and B are mutually exclusive,
(ii) A and B are independent.

(2)

Two events R and Q are such that

P($R \cap Q'$) = 0.15, P(Q) = 0.35 and P($R|Q$) = 0.1

Find the value of

(b) P($R \cup Q$),

(1)

(c) P($R \cap Q$),

(2)

(d) P(R).

(2)

(Total for Question 3 is 7 marks)

End of questions

Bronze Questions

Calculator

The total mark for this section is 33

Q1

The random variable X has a normal distribution with mean 20 and standard deviation 4.

(a) Find P(X >25).

(3)

(b) Find the value of d such that P(20 < X < d) = 0.4641.

(4)

(Total for Question 1 is 7 marks)

Q2

The time taken for a randomly selected person to complete a test is M minutes, where $M \sim N(14, \sigma^2)$

Given that 10% of people take less than 12 minutes to complete the test,

(a) find the value of σ.

(3)

Graham selects 15 people at random.

(b) Find the probability that fewer than 2 of these people will take less than 12 minutes to complete the test.

(3)

Jovanna takes a random sample of n people.

Using a normal approximation, the probability that fewer than 9 of these n people will take less than 12 minutes to complete the test is 0.3085 to 4 decimal places.

(c) Find the value of n.

(8)

(Total for Question 2 is 14 marks)

Q3

A machine cuts strips of metal to length L cm, where L is normally distributed with standard deviation 0.5 cm.

Strips with length either less than 49 cm or greater than 50.75 cm **cannot** be used.

Given that 2.5% of the cut lengths exceed 50.98 cm,

(a) find the probability that a randomly chosen strip of metal **can** be used.

(5)

Ten strips of metal are selected at random.

(b) Find the probability fewer than 4 of these strips **cannot** be used.

(2)

A second machine cuts strips of metal of length X cm, where X is normally distributed with standard deviation 0.6 cm

A random sample of 15 strips cut by this second machine was found to have a mean length of 50.4 cm

(c) Stating your hypotheses clearly and using a 1% level of significance, test whether or not the mean length of all the strips, cut by the second machine, is greater than 50.1 cm

(5)

(Total for Question 3 is 12 marks)

End of questions

Silver Questions

Calculator

The total mark for this section is 37

Q1

A manufacturer fills jars with coffee. The weight of coffee, W grams, in a jar can be modelled by a normal distribution with mean 232 grams and standard deviation 5 grams.

(a) Find P($W < 224$).

(3)

(b) Find the value of w such that P($232 < W < w$) = 0.20

(4)

Two jars of coffee are selected at random.

(c) Find the probability that only one of the jars contains between 232 grams and w grams of coffee.

(3)

(Total for Question 1 is 10 marks)

Q2

A shopkeeper knows, from past records, that 15% of customers buy an item from the display next to the till. After a refurbishment of the shop, he takes a random sample of 30 customers and finds that only 1 customer has bought an item from the display next to the till.

(a) Stating your hypotheses clearly, and using a 5% level of significance, test whether or not there has been a change in the proportion of customers buying an item from the display next to the till.

(6)

During the refurbishment a new sandwich display was installed. Before the refurbishment 20% of customers bought sandwiches. The shopkeeper claims that the proportion of customers buying sandwiches has now increased. He selects a random sample of 120 customers and finds that 31 of them have bought sandwiches.

(b) Using a suitable approximation and stating your hypotheses clearly, test the shopkeeper's claim. Use a 10% level of significance.

(8)

(Total for Question 2 is 14 marks)

Q3

A machine puts liquid into bottles of perfume. The amount of liquid put into each bottle, D ml, follows a normal distribution with mean 25 ml.

Given that 15% of bottles contain less than 24.63 ml

(a) find, to 2 decimal places, the value of k such that $P(24.63 < D < k) = 0.45$.

(5)

A random sample of 200 bottles is taken.

(b) Using a normal approximation, find the probability that fewer than half of these bottles contain between 24.63 ml and k ml.

(3)

The machine is adjusted so that the standard deviation of the liquid put in the bottles is now 0.16 ml.

Following the adjustments, Hannah believes that the mean amount of liquid put in each bottle is less than 25 ml.

She takes a random sample of 20 bottles and finds the mean amount of liquid to be 24.94 ml.

(c) Test Hannah's belief at the 5% level of significance.

You should state your hypotheses clearly.

(5)

(Total for Question 3 is 13 marks)

End of questions

Gold Questions

Calculator

The total mark for this section is 35

Q1

The lifetimes of bulbs used in a lamp are normally distributed.
A company X sells bulbs with a mean lifetime of 850 hours and a standard deviation of 50 hours.

(a) Find the probability of a bulb, from company X, having a lifetime of less than 830 hours.

(3)

(b) In a box of 500 bulbs, from company X, find the expected number having a lifetime of less than 830 hours.

(2)

A rival company Y sells bulbs with a mean lifetime of 860 hours and 20% of these bulbs have a lifetime of less than 818 hours.

(c) Find the standard deviation of the lifetimes of bulbs from company Y.

(4)

Both companies sell the bulbs for the same price.

(d) State which company you would recommend. Give reasons for your answer.

(2)

(Total for Question 1 is 11 marks)

Q2

The length of time, L hours, that a phone will work before it needs charging is normally distributed with a mean of 100 hours and a standard deviation of 15 hours.

(a) Find P ($L > 127$).

(3)

(b) Find the value of d such that P($L < d$) = 0.10.

(3)

Alice is about to go on a 6 hour journey.
Given that it is 127 hours since Alice last charged her phone,

(c) find the probability that her phone will not need charging before her journey is completed.

(4)

(Total for Question 2 is 10 marks)

Q3.

The lifetime, L hours, of a battery has a normal distribution with mean 18 hours and standard deviation 4 hours.

Alice's calculator requires 4 batteries and will stop working when any one battery reaches the end of its lifetime.

(a) Find the probability that a randomly selected battery will last for longer than 16 hours.

(1)

At the start of her exams Alice put 4 new batteries in her calculator.
She has used her calculator for 16 hours, but has another 4 hours of exams to sit.

(b) Find the probability that her calculator will not stop working for Alice's remaining exams.

(5)

Alice only has 2 new batteries so, after the first 16 hours of her exams, although her calculator is still working, she randomly selects 2 of the batteries from her calculator and replaces these with the 2 new batteries.

(c) Show that the probability that her calculator will not stop working for the remainder of her exams is 0.199 to 3 significant figures.

(3)

After her exams, Alice believed that the lifetime of the batteries was more than 18 hours. She took a random sample of 20 of these batteries and found that their mean lifetime was 19.2 hours.

(d) Stating your hypotheses clearly and using a 5% level of significance, test Alice's belief.

(5)

(Total for Question 3 is 14 marks)

End of questions