

IB Maths: AA HL

Statistics Toolkit

Topic Questions

These practice questions can be used by students and teachers and is Suitable for IB Maths AA HL Topic Questions

| | |
|------------|-----------------------------|
| Course | IB Maths |
| Section | 4. Statistics & Probability |
| Topic | 4.1 Statistics Toolkit |
| Difficulty | Medium |

Level: IB Maths

Subject: IB Maths AA HL

Board: IB Maths

Topic: Statistics Toolkit

Question 1

Every week an orangutan sanctuary measures the weight of each of its orangutans.

The weights, to the nearest kg, of **ALL** their 18 adult males are listed below:

52, 57, 63, 80, 56, 66, 101, 68, 55, 96, 70, 62, 66, 64, 99, 91, 55, 92

(a) Using a convenience sample of size six, calculate the mean weight of the male orangutans from the data set above.

[2 marks]

(b) Starting from the third data value, take a systematic sample of size six and re-calculate the mean weight of the male orangutans from the data set above.

[2 marks]

(c) Compare your results from parts (a) and (b) and state, with a reason, which sampling method is more reliable.

[2 marks]

Question 2

A supermarket wants to gather data from its shoppers on how far they have travelled to shop there. One lunchtime an employee is stationed at the door of the shop for half an hour and instructed to ask every customer how far they have travelled.

(a) (i) State the sampling method the employee is using.

(ii) Give one advantage and one disadvantage of using this method.

[3 marks]

(b) State and briefly describe an alternative method of non-random sampling that the employee could use to obtain the required data for a sample of 30 customers.

[2 marks]

Question 3

A pharmacy sells face masks in a variety of sizes. Their sales over a week are recorded in the table below:

| | Kids | | Adults | | | |
|---------------|-------|-------|--------|----|----|----|
| Size | Small | Large | S | M | L | XL |
| Frequency f | 29 | 4 | 8 | 24 | 15 | 4 |

- (i) Write down the mode for this data.
- (ii) Explain why, in this case, the mode from part (i) would not be particularly helpful to the shop owner when reordering masks.
- (iii) Given that the shop is open every day of the week, calculate the mean number of masks sold per day.

[4 marks]

Question 4

The lengths (l cm) of a sample of nine otters, measured to the nearest centimetre by a wildlife research team, are:

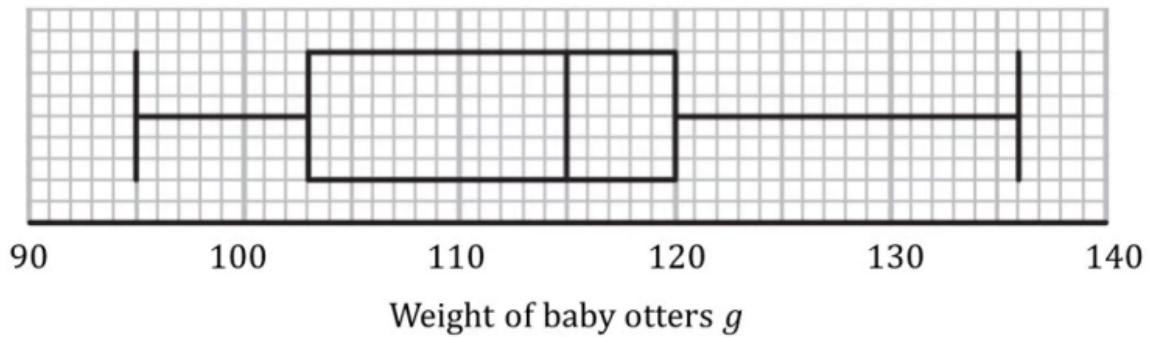
76 77 91 65 63 83 92 61 88

Calculate the mean and standard deviation of the nine recorded lengths.

[3 marks]

Question 5

Jeanette works for a conservation charity who rescue orphaned otters. Over many years she records the weight (g) of each otter when it first arrives. The data is illustrated in the following box and whisker diagram:



(a) Using the box plot above:

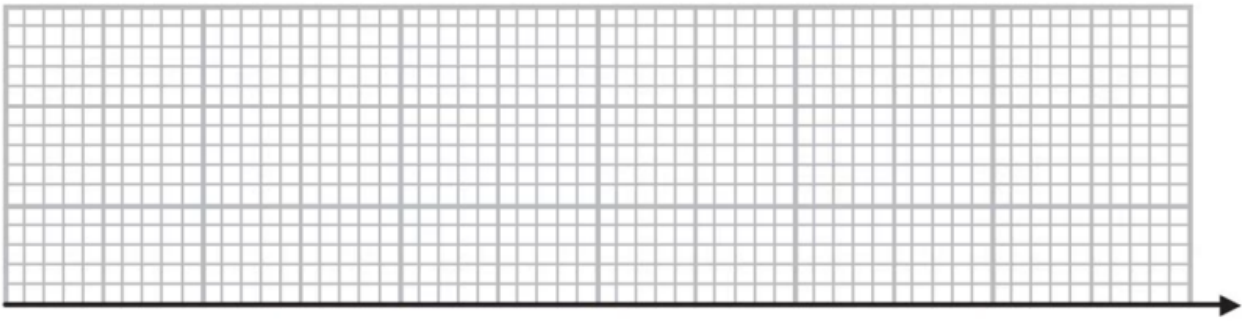
- (i) Write down the median weight of the otters.
- (ii) Write down the lower quartile.
- (iii) Find the interquartile range.

[4 marks]

Otters are then weighed weekly to track their growth. Summary data on the weights (g) of otters after one month is shown in the table below:

| | Weight g |
|---------------------|------------------------------|
| Smallest weight | 125 |
| Range | 48 |
| Median | 152 |
| Upper Quartile | 164 |
| Interquartile Range | 33 |

(b) On the grid, draw a box plot for the information given above.



[3 marks]

Question 6

The heights, in metres, of a flock of 20 flamingos are recorded and shown below:

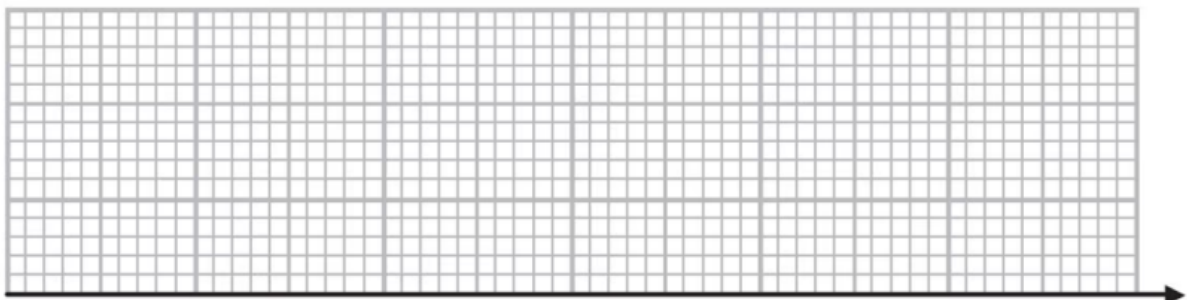
| | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 0.4 | 0.9 | 1.0 | 1.0 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 |
| 1.3 | 1.3 | 1.3 | 1.4 | 1.4 | 1.4 | 1.4 | 1.5 | 1.5 | 1.6 |

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

- (a) (i) Find the values of Q_1 , Q_2 and Q_3 .
(ii) Find the interquartile range.
(iii) Identify any outliers.

[4 marks]

- (b) Using your answers to part (a), draw a box plot for the data.



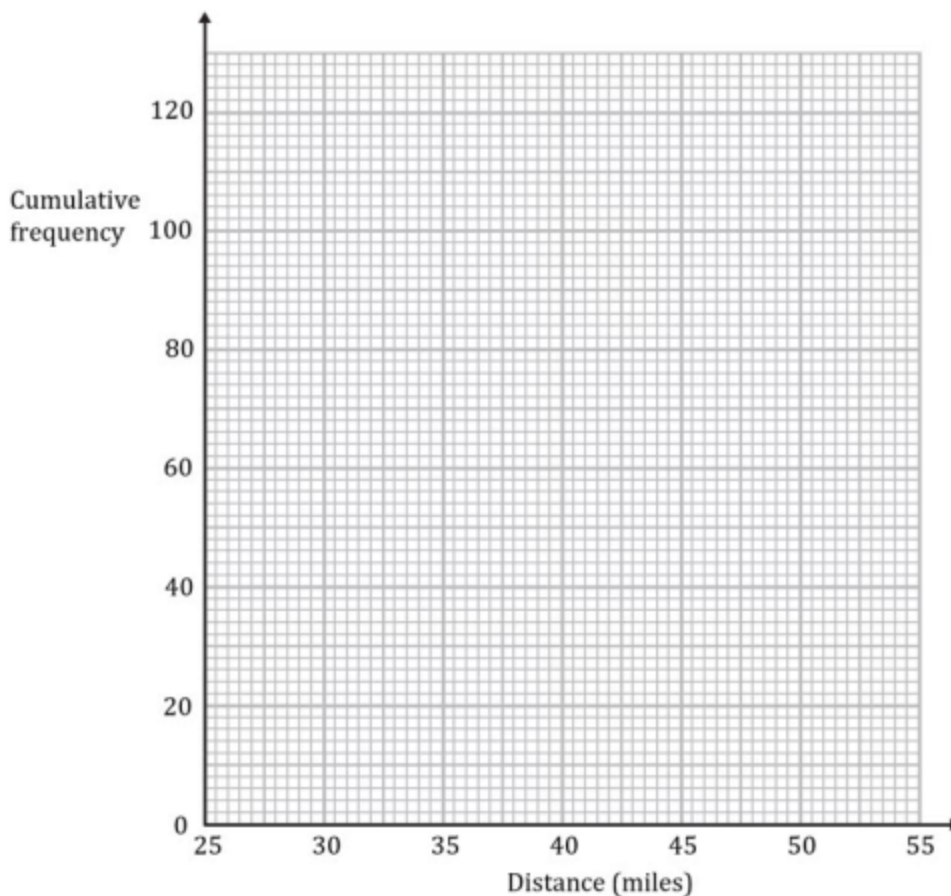
[3 marks]

Question 7

120 competitors enter an elimination race for charity. Runners set off from the same start running as many laps of the course as possible. Their total distance is tracked and the competitor who runs the furthest over a 6-hour period is the winner. The distances runners achieved are recorded in the table below:

| Distance d (miles) | Frequency f |
|----------------------|---------------|
| $25 \leq d < 30$ | 8 |
| $30 \leq d < 35$ | 10 |
| $35 \leq d < 40$ | 32 |
| $40 \leq d < 45$ | 54 |
| $45 \leq d < 50$ | 10 |
| $50 \leq d < 55$ | 6 |

(a) On the grid below, draw a cumulative frequency graph for the information in the table.



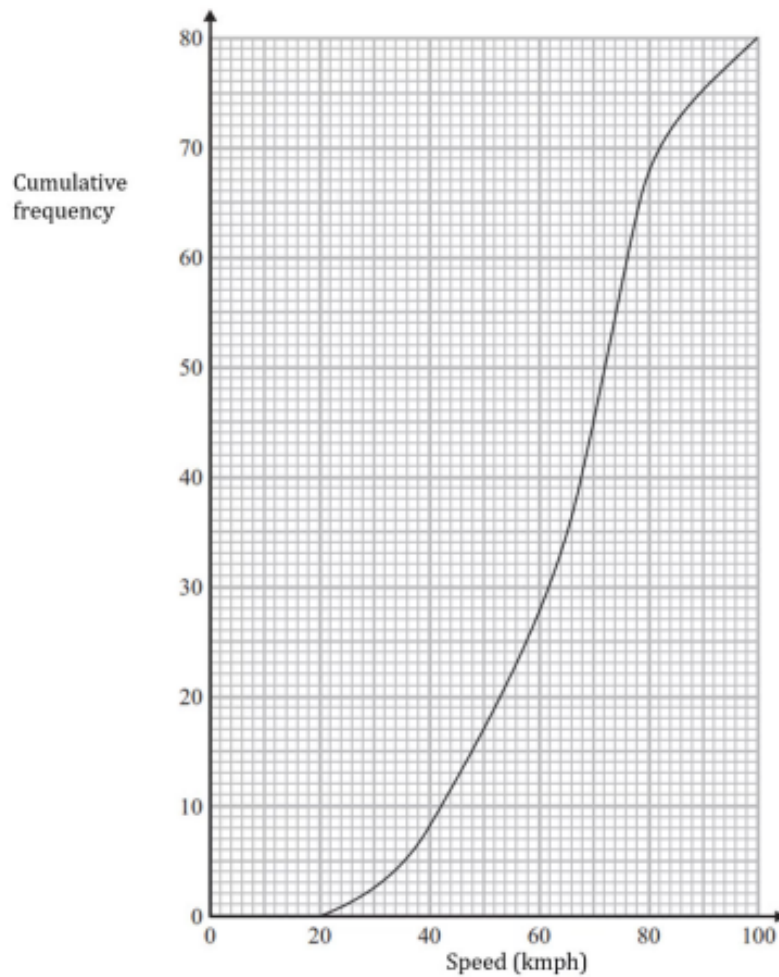
[3 marks]

(b) Use your graph to find an estimate for the median and interquartile range.

[3 marks]

Question 8

Police check the speed of vehicles travelling along a stretch of highway. The cumulative frequency curve below summarises the data for the speeds, in kmph, of 80 vehicles:



(a) Use the graph to find an estimate for the median speed.

[2 marks]

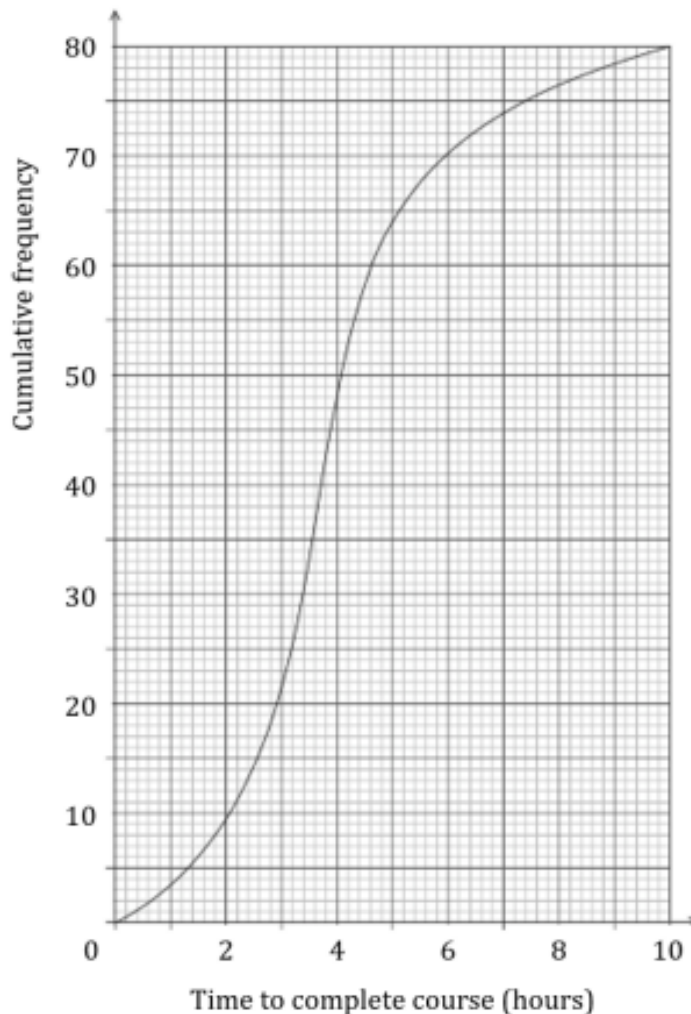
The speed limit for this section of road is 80 kmph.

(b) Vehicles travelling above the speed limit are issued with a speeding ticket. Those travelling more than 10% over the speed limit are pulled over. Use the graph to estimate the percentage of vehicles that the police pull over.

[3 marks]

Question 9

The following cumulative frequency curve shows the number of hours, h , students took to complete their online driving course. The data is taken from 80 students, randomly selected from a large sample over a 12 month period.



(a) Find the median number of hours spent completing the online driving course.

[2 marks]

(b) Find the number of students whose online course time was within 1 hour of the median.

[2 marks]

(c) Calculate the interquartile range.

[2 marks]

The same information is represented by the following table.

| | | | | |
|------------|----------------|----------------|----------------|-----------------|
| Hours, h | $0 < h \leq 2$ | $2 < h \leq 4$ | $4 < h \leq 7$ | $7 < h \leq 10$ |
| Frequency | 9 | p | q | 6 |

(d) Find the value of p and the value of q .

[3 marks]

It is known that 10% of students take longer than d hours to complete the online driving course.

(e) Find the value of d .

[3 marks]

It is known that over a 12 month period, 4000 students in total sat the online driving course.

(f) Estimate the number of students over a 12 month period who took less than 3 hours to complete the course.

[3 marks]