# Probability and Statistics


PROBABILITY ..... 2

SETS AND VENN DIAGRAMS ..... 4

STATISTICS SUMMARY ..... 7

CONTINUOUS AND DISCRETE DATA ..... 8

MEAN MODE AND MEDIAN ..... 9

THE MEAN OF GROUPED DATA ..... 10

SPREAD ..... 11

CUMULATIVE FREQUENCY ..... 12

CUMULATIVE FREQUENCY GRAPHS ..... 13

HISTOGRAMS ..... 14

## Probability

Probabilities must always add up to

Eg If the chance of rain today is 0.7, what is the chance that it will stay dry?

***Tree Diagrams***

Tree diagrams are a way of working out the probabilities of two things happening.

Eg The chance of rain today is $\frac{1}{3}$. The chance of rain tomorrow is $\frac{4}{7}$. Draw a tree diagram to represent this information.

What is the chance of two wet days?

What is the chance of *at least* one wet day?

Ann has a bag containing 8 chocolates. 5 are dark chocolate and 3 are milk chocolate. She takes two chocolates from the bag at random and eats them. Draw a tree diagram and find the chance of drawing;

2 dark chocolates
1 dark and 1 milk chocolate (in either order)
at least one dark chocolate.

## Sets and Venn Diagrams



Write the meaning of the following elements of set notation;

| | |
|---|---|
| $\in$ | $\notin$ |
| $\subseteq$ | $\subset$ |
| $\cup$ | $\cap$ |
| $\varnothing$ | n**A** |
| U | **B**' |

If U = {a, b, c, d, e, f, g, h, i}
A = {a, c, e, g, i}
B = {a, d, g, h}

Draw the Venn diagram and List

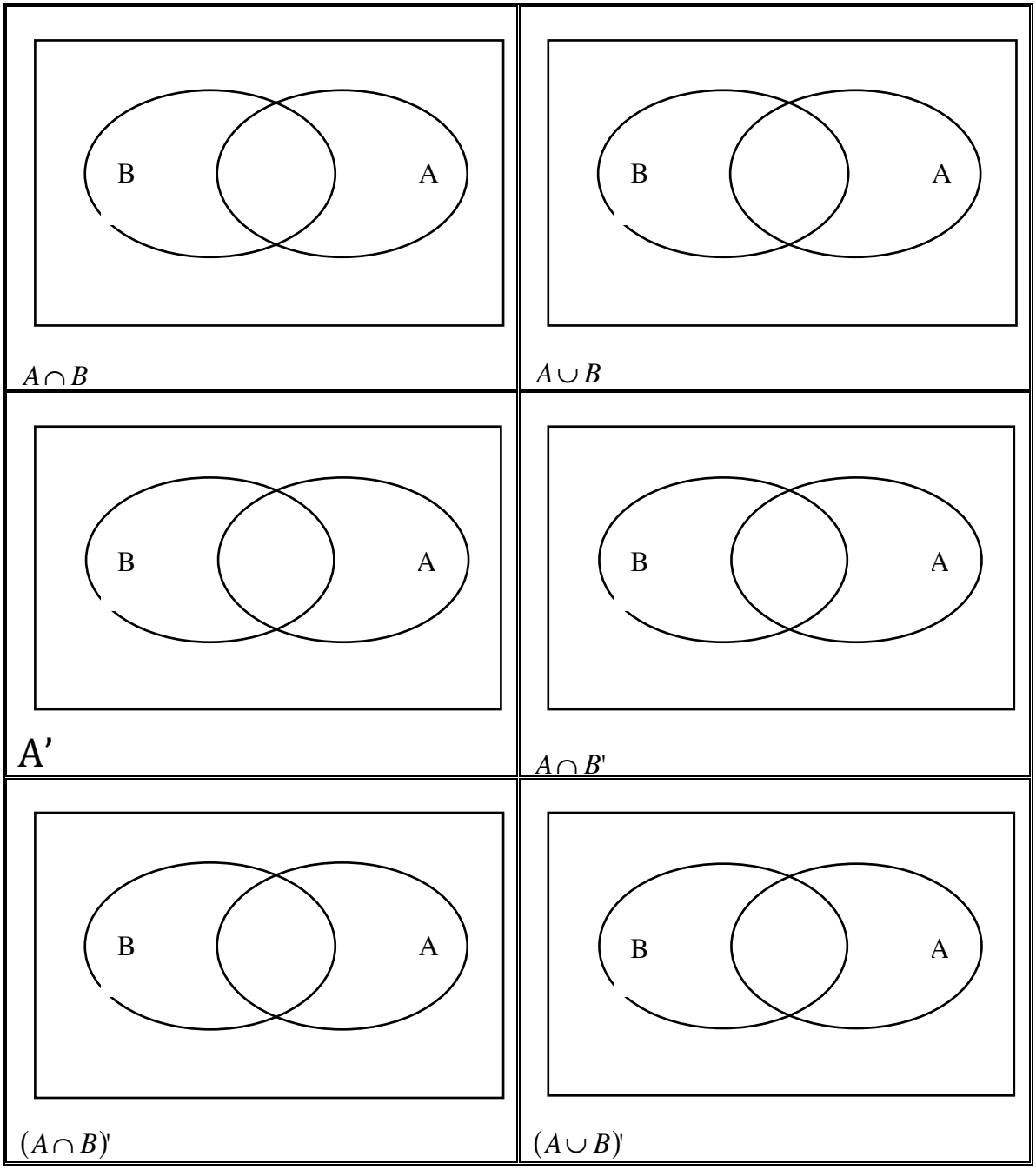
$A \cap B$

$A \cup B$

$n(A \cap B)$

$n(A' \cap B)$

Shade the appropriate regions of the diagrams below;



## Statistics Summary

Statistics is only about 3 things;

You will complete the following diagram as we work through the syllabus, it will help you categorise the techniques you learn.

| | | |
|---|---|---|
| | | |

## Continuous and Discrete Data

Discrete data comes in _____ like people at a football match. The number can only go up by a whole person. You can't have 45,214.43 people watching a game.

Give 4 other examples of discrete data

1

2

3

4

Continuous data can take _____ value - like your height. You don't suddenly grow from 1.45m to 1.46m in a sudden jump!

Give 4 other examples of Continuous data

1

2

3

4

## Mean Mode and Median

There are 3 types of averages, which you use in different situations. They are

1

2

3

Mean

The mean is the _____ average. You _____ all the data and then divide by the _____.

Example – find the mean of 10, 12, 17, 24

Median

The median is the _____ number. You put the data _____ and then find the middle number. You use the median when the data has one very _____ or _____ value which would influence the mean.

Example – find the median of 10, 8, 17, 5, 20

Mode

The mode is the _____ number. You use the mode when you want to find out what happens most often.

Give two examples when you would use the mode.

1

2

## The Mean of Grouped Data

When your data is grouped you don't know its exact value so you have to estimate that each data point lies in the middle of its group. You then find the mean as before.

Calculate the mean height of each of these classes

| Height | $1.10 \leq h < 1.15$ | $1.15 \leq h < 1.20$ | $1.20 \leq h < 1.25$ | $1.25 \leq h < 1.30$ | $1.30 \leq h < 1.35$ | $1.35 \leq h < 1.40$ |
|---|---|---|---|---|---|---|
| Class A (f) | 5 | 15 | 23 | 27 | 12 | 8 |
| | | | | | | |

| Height | $1.10 \leq h < 1.15$ | $1.15 \leq h < 1.20$ | $1.20 \leq h < 1.25$ | $1.25 \leq h < 1.30$ | $1.30 \leq h < 1.35$ | $1.35 \leq h < 1.40$ |
|---|---|---|---|---|---|---|
| Class B | 7 | 18 | 19 | 23 | 15 | 8 |
| | | | | | | |

Which class, on average, is taller ?

## Spread

**Range**
The range gives you a rough measure of how spread out the data are. Range is calculated from…

**Quartiles**
If you put a data set in order, the lowest data point is called the

_____ value. The middle point is called the _____

value and the highest point is called the _____ value. Quartiles

are just a way of adding to that. The lower quartile is _____ of the

way through the data whilst the upper quartile is _____

of the way through the data.

*Calculate the above datapoints for this series;*

*-1 2 4 6 7 8 8 9 10 12 14 15 16 18 19 20 20*

What proportion of the data points must fall between the two quartiles?

What is the inter-quartile range ?

What does it tell us?

*Compare the following data series using the methods above.*

*1 2 4 6 7 8 8 9 10 12 14 15 16 18 19 20 20*

*1 2 3 4 4 5 5 9 10 12 14 17 18 18 19 20 20*

## Cumulative Frequency

This is simply the _____.

*Calculate the cumulative frequencies for the classes we investigated earlier*

| Height | 1.10≤ h <1.15 | 1.15≤ h <1.20 | 1.20≤ h <1.25 | 1.25≤ h <1.30 | 1.30≤ h <1.35 | 1.35≤ h <1.40 |
|---|---|---|---|---|---|---|
| Class A (f) | 5 | 15 | 23 | 27 | 12 | 8 |
| C-Freq | | | | | | |

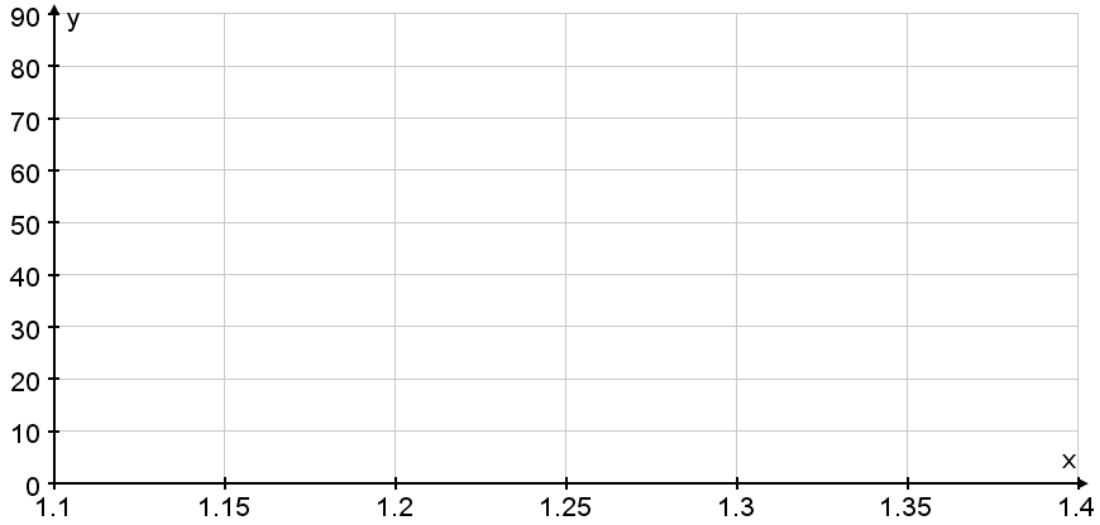
| Height | 1.10≤ h <1.15 | 1.15≤ h <1.20 | 1.20≤ h <1.25 | 1.25≤ h <1.30 | 1.30≤ h <1.35 | 1.35≤ h <1.40 |
|---|---|---|---|---|---|---|
| Class B | 7 | 18 | 19 | 23 | 15 | 8 |
| C-Freq | | | | | | |

## Cumulative Frequency Graphs

They show you the same as the box plot but with _____.

**The important point to remember about cumulative frequency curves is to always _____**

*Plot on the graph below, both of the cumulative frequencies that you worked out. Don't forget to label the axes.*



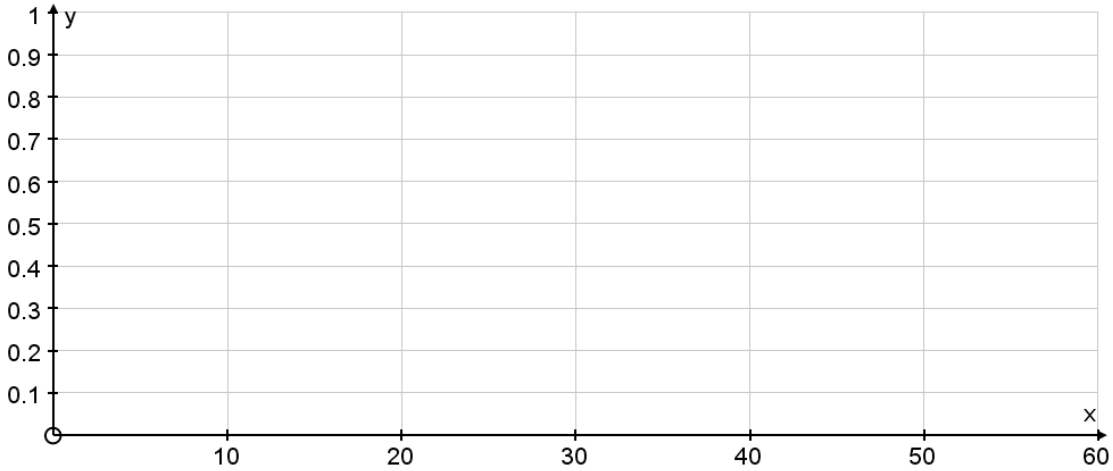

You can also work out the lower quartile, median and upper quartile on a cumulative frequency curve as they are just the values which are _____, _____ and _____ of the way through the data.

## Histograms

Histograms are a very important type of graph as they allow you to use _____ in the regions where the data is most packed and less detail in other areas.

| Group | $0 \leq x < 30$ | $30 \leq x < 35$ | $35 \leq x < 40$ | $40 \leq x < 50$ | $50 \leq x < 60$ |
|---|---|---|---|---|---|
| Frequency | 3 | 4 | 2 | 4 | 2 |
| Group Width | | | | | |
| Frequency Density | | | | | |

Now draw the histogram on the graph below. Don't forget to label the axes with the appropriate titles.




Online histogram worksheet here
Online worksheet here