GCSE Maths – Probability

Samples

Notes

WORKSHEET

Samples

In statistics, a **population** is the overall group of people, values or items that can be **grouped by a common characteristic**. For example, a population could contain every student in a school.

Populations are analysed to collect data and build trends. Often, a population contains such a large number of items that it can be hard to gather data from every individual. Instead, the population can be **sampled** – or split into smaller **sub-groups**. The sample should be **representative** of the population – meaning that any data collected from the sub-group also applies to the entire population.

- **Population** – a group of items with a common characteristic.
- **Sample** – a smaller sub-group of a population.
- **Representative** – data from the sub-group also applies to the population. Any trends or patterns shown in the data reflect the trends or patterns shown in the entire population.

We can make sure that a sample is representative of a population by choosing it in a **non-biased**, or 'fair', way.

Methods of Sampling

Sampling methods make sure that a study **avoids bias** by choosing the items fairly. There are three main types of sampling:

1. **Random Sampling** – each item in the population has an equal chance of being chosen for the sample. This is un-biased as one item is not favoured over another.

   A random sample could be carried out by numbering the items in a population, then using a random number generator to pick which numbers are included in the sample. Repeated numbers are ignored and the process continues until the sample is the required size.

2. **Systematic Sampling** – the population is ordered and items are chosen at regular intervals. The interval should be a factor of the population. One way to decide on the interval is to divide the population size by the sample size.

   A population of 50 towns is put into alphabetical order. 50 ÷ 10 = 5 so every 5^{th} town is added to the sample, so that the sample contains 10 towns.

3. **Stratified Sampling** – used when the population can be split into sub-categories (or strata) with the same characteristic. Random or systematic sampling can then be applied to each strata. Note: the number of items chosen from each strata should be proportional to the size of the strata.

   A stratified sample could be used when Mr. Miller wants to find out whether the students in his school like maths. He splits the population of the school into year groups, then applied the random sampling method to each year group to choose his population.

Sample Size

When we take a sample of a population, the **size of the sample** will affect how **representative** it is. Take a population of 5000 people:

- A sample of 5000 will be **very** representative as everyone's views are included.
- A sample of 1000 will be **quite** representative – it is likely that the data collected will show the views of the whole population.
- A sample of 10 **will not** be representative of the population. There are not enough people in this sample to correctly show trends in the population, so the data collected would probably be **biased**.

If your population size is less than 100, your sample should contain all items.

If your population size is more than 100, your sample should contain at least 100 items OR 10% of your population, if this is higher.

Example: Using systematic sampling, take a sample of 6 towns from the following population: Aberdeen, Bristol, Birmingham, Cardiff, Corby, Coventry, Durham, Edinburgh, Essex, Halton, Havant, Hereford, Leicester, Newport, St. Alban's, Warwick, Worchester, Worksop.

Determine the interval for systematic sampling.

The population contains 18 towns. 18 ÷ 6 = 3 so every 3rd town is added to the sample.

The sample contains Aberdeen, Cardiff, Durham, Halton, Leicester and Warwick.

Example: A school has 500 students split into 5 year groups. Each student has either packed lunches or school dinners. The table below shows how many students are in each year group.

Year Group	7	8	9	10	11
Number of Students	50	180	70	100	100

Joseph wants to find out how many students have packed lunches by taking a sample of 50 students. Decide on the best method of sampling for his study and give a reason why.

The best sampling method would be stratified sampling.

Stratified sampling is where items are taken from each group (or strata), in proportion to the population.

This is because…

- each year group has a different number of students.
- each year group forms a strata.
- it will be most representative of the population.

It is most representative because the number of items in each strata is proportional to the population.

Stratified Sampling - Calculations

Items from each strata are put into the sample. It is important that the number of items from each strata is **proportional** to the size of the strata within the population.

For example, take the following situation:

- A town has a population of 100.
- Within the town there are 50 women, 30 men and 20 children.

A researcher wants to find out how many people like living in the town by surveying a sample of 10 townsfolk. The researcher would use stratified sampling to choose their sample. The men, women and children form three distinct strata.

How does the researcher work out how many from each strata should be included in their sample?

Within the sample of 10 townsfolk:

- $\frac{50}{100} = \frac{5}{10}$ of the sample should be women. $\frac{5}{10}$ of 10 is 5, so **5 women** should be randomly chosen from the strata for the sample.
- $\frac{30}{100} = \frac{3}{10}$ of the sample should be women. $\frac{3}{10}$ of 10 is 3, so **3 men** should be randomly chosen from the strata for the sample.
- $\frac{20}{100} = \frac{2}{10}$ of the sample should be women. $\frac{2}{10}$ of 10 is 2, so **2 children** should be randomly chosen from the strata for the sample.

The sample now contains 5 women, 3 men and 2 children, which is **proportional** to the population. So, the number of items chosen from each strata is determined by the following equation:

$$\frac{Size\ of\ Strata}{Size\ of\ Population} \times Sample\ size$$

Example: A school has 500 students split into 5 year groups. Each student has either packed lunches or school dinners.

Year Group	7	8	9	10	11
Number of Students	50	180	70	100	100

Joseph wants to find out how many students have packed lunches by taking a sample of 50 students. Work out how many students from year 9 should be included in his sample.

1. Identify the size of the **strata**, size of **population** and **sample** size.

   Size of strata = 70, Size of population = 500, Size of sample = 50

2. Apply the formula to calculate sample size.

$$Number\ of\ items = \frac{Size\ of\ Strata}{Size\ of\ Population} \times Sample\ Size = \frac{70}{500} \times 50 = \mathbf{7\ students}$$

Samples – Practice Questions

1. Find the interval size using systematic sampling in each of the following scenarios:

   a) Tom wants a sample of 15 people from a population of 150 people.
   b) Alfie needs to interview 20 students from a school of 180 students.

2. Select the best method of sampling for each of the following situations:

   a) Ethan wants to get a 50 people from 2000 strangers who visited the market today.
   b) Amina has a list of 200 towns in England and needs a sample of 20 towns.
   c) Jasmine wants to interview 15 of the 150 girls at her rugby club, who are split into 5 different age groups.

3. Complete the tables below:

   a) Alice visits 100 people in local cafés. She takes a sample of 10 people to investigate what people like to drink.

	Café A	Café B	Café C	Total
Number of people in the cafe	30	20	40	100
Number of people in sample				10

   b) Emmilé wants to interview people at a football game. She needs a sample of 40 people from a population of 800.

	Men	Women	Children	Total
Number of people in the cafe	480	120	200	800
Number of people in sample				10

Worked solutions for the practice questions can be found amongst the worked solutions for the corresponding worksheet file.