

Inferring properties of populations or distributions from a sample

Check that you can:

- criticise questionnaires and surveys
- understand and use proportion as equality of ratios.

Population and Sample

The **population** is the entire group that we want to draw conclusions about.

Advantages

- All opinions considered.
- Data is reliable.

Disadvantages

- Takes a long time to conduct.
- Can be costly.

A **sample** is the subset of the population we'll use to draw our conclusions from. A sample is always smaller in size than the population.

Advantages

- Quick to conduct.
- Cost-effective.

Disadvantages

- Sampling technique could lead to biased results.
- Only a cross-section of opinions.

Simple random sampling is a technique that is used to choose the subset of a population. All members have the same probability of being selected.

Bias

If something is biased, it means that the population is not fairly represented. This could happen in the following situations.

Sample size – a small sample can lead to misleading results, as the sample might only include people of the same gender, age, income, etc.

Timing of the survey – if you conduct a survey of train passengers at 8am, you will bias the sample towards working people.

Location of the survey – if you want information on people's favourite sport, and you conduct a survey outside a football ground, you will bias the sample towards people whose favourite sport is football.

REMEMBER!

Sample is a subset of the population; however, the sample must be selected using appropriate techniques in order to avoid any bias.

Making inferences regarding the population

To estimate the number of a type of bird in a forest, a scientist captures 50 birds, rings their legs and lets them go. The following day, the scientist captures 30 birds and 5 of them are ringed. What inferences can we make regarding the size of the population?

Answer:

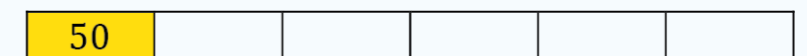
Knowing that on the following day 5 of the 30 captured birds are ringed, we can infer that a sixth of the population is ringed:

$$\frac{5}{30} = \frac{1}{6}$$

Therefore, if the original sample size is 50 birds, then:

$$\frac{1}{6} = 50$$

Or using the bar method:



Thus, the population size is $50 \times 6 = 300$ birds.

There are some **assumptions** that need to be considered when making inferences as they may have an impact on the population size. For example, was the sample chosen in a random way? The population may also have increased or decreased between the collection of the samples due to births or deaths.