



Section 2.3 Populations and sampling

Definitions of

- Simple random sampling
- Cluster sampling
- Quota sampling
- Stratified sampling

are required to be known.



In a statistical enquiry, you often need information about a particular group. This group is known as the population and it could be small, large or even infinite.

Examples of populations include

- (i) all second-level pupils in Ireland
- (ii) paid-up members of golf clubs
- (iii) people entitled to vote in a general election.

If information is obtained from all members of a population, the survey is called a census.

Sample survey

When a population is large, taking a census can be very time-consuming and difficult to do with accuracy. So when a census is ruled out as being impractical, information is normally taken from a small part of the population. The chosen members of the population are called a **sample** and an investigation using a sample is called a **sample survey**. Data from a sample can be obtained relatively cheaply and quickly. If the data is representative of the population, a sample survey can give an accurate indication of the population characteristic that is being studied.

The **size** of a sample is important. If the sample is too small, the results may not be very reliable.

If the sample is too large, the data may take a long time to collect and analyse.

However, large samples are more likely to give reliable information than small ones.

Bias in sampling

The sample you select for your study is very important. If the sample is not properly selected, the results may be **biased**. If **bias** exists, the results will be distorted and so may not be representative of the population as a whole.

Bias in a sample may arise from any of the following:

› **Choosing a sample which is not representative**

Example Cara is doing a survey on people's attitude towards gambling. If she stands outside a casino and questions people as they enter or leave, the results will be biased as these people are already involved in gambling.

› **Not identifying the correct population**

Example The school principal wants to find out about students' attitudes to school uniforms. She questions ten Leaving Certificate students only. This may lead to biased results as the opinions of the younger students (from 1st year to 5th year) are not included.

› **Failure to respond to a survey**

Many people do not fill in responses to questionnaires sent through the post. Those who do respond may not be representative of the population being surveyed.

› **Dishonest answers to questions**

Sampling methods

The purpose of sampling is to gain information about the whole population by selecting a sample from that population. If you want the sample to be representative of the population, you must give every member of the population an equal chance of being included in the sample. This is known as **random sampling**. Before a random sample is selected, a **sampling frame** must be used to identify the population. A sampling frame consists of all the items in the population to ensure that every item has a chance of being selected in the sample.

Some of the most commonly-used sampling methods are given below.

1. Simple random sampling

A sample of size n is called a **simple random sample** if every possible sample of size n has an equal chance of being selected. In practice, this means that each member of the population has an equal chance of being selected.

There are many ways of doing this.

Methods for choosing a **simple random sample** could involve giving each member of the population a number and then selecting the numbers for the sample in one of these ways:

- › putting the numbers into a hat and then selecting however many you need for the sample
- › using a random number table
- › using a random number generator on your calculator or computer

Any of these methods are suitable only if the population is relatively small and the sampling frame is clearly identified.

2. Stratified sampling

Stratified sampling is used when the population can be split into separate groups or strata that are quite different from each other. The number selected from each group is proportional to the size of the group. Separate random samples are then taken from each group.

3. Systematic sampling

A sample which is obtained by choosing items at regular intervals from an unordered list is called a **systematic sample**. For example, if you wish to choose 20 students from 200 students, you could take every tenth student from the register. Select a random number between 1 and 10, e.g., 4. Thus you could select the 4th, 14th, 24th, 34th ... until you get 20 students

4. Quota sampling

Quota sampling is widely used in market research and in opinion polls. First the population is divided into groups in terms of age, general education levels, social class, etc. The interviewer is then told how many people (the quota) to interview in each of these groups, but the interviewer makes the choice of who exactly is asked. A disadvantage of quota sampling is that the actual people or items chosen are left to the discretion of the interviewer which could lead to bias. An advantage of quota sampling is that no sampling frame is required.

Example 2

Simon wanted to investigate whether people in Ireland measured their height in metric or imperial units. He went to his local supermarket and asked the first twenty people he met how tall they were.

- (i) For this survey, state the sampling frame, the sampling method used and why it might be biased.
- (ii) Outline a better method of choosing a sample.

i	Sampling frame : people in local supermarket at that time
	Sampling method : convenience sampling
ii	A better method would : <ul style="list-style-type: none">• include more people and not just those in local supermarket at this particular time.• include an element of randomness• perhaps include proportionate sub-groups based on age