Random Sampling

Ordinarily we would want each thing in our sampling frame to have an **equal chance of being chosen**, in order to **avoid bias**.
This is known as *random sampling*. There are a few ways to do this...

Simple Random Sampling

What is it?

Every sample has an equal chance of being selected.

Method:

In sampling frame each item has an identifying number. Use a random number generator, or '*lottery sampling*' (eg. picking names from a hat)

Advantages

- Bias free
- Easy and cheap to implement
- Each number has a known equal chance of being selected

Disadvantages

- Not suitable when the population size is large
- Sampling frames needed.

Systematic Sampling

What is it?

Required elements are chosen at regular intervals in an ordered list. ie. Take every kth elements where: $k = \frac{PopulationSize(N)}{SampleSize(n)}$ starting at random items between 1 and k.

Advantages

- Simple and quick to use.
- Suitable for large samples / populations

Disadvantages

- Sampling frame again needed.
- Can introduce bias if sampling frame is not adequately random.

Stratified Sampling

What is it?

Population divided into groups (strata) and a ***simple random sample*** is carried out in each group/strata.
Same proportion ($\frac{SampleSize(n)}{PopulationSize(N)}$) is sampled from each strata.
Often used when the sample is large, and the population naturally divides into groups.

Advantages

- Reflects population structure
- Guarantees proportional representation of groups withing populations

Disadvantages

- Population must be clearly identified into distinct strata
- Selection from within each stratum suffers from same disadvantages as ***simple random sampling***