College of Applied Medical Sciences

Department of Environmental Health

Quality Control

Lecture4

Population :

a set which includes all measurements of interest to the researcher (The collection of responses, measurements, or counts that are of interest) .

Sample:

A subset of the population

Why sampling?

1-Get information about large populations .

2-Less costs .

3-Less field time .

4-More accuracy .

5-When it's impossible to study the whole population .

Target Population:

The population to be studied/ to which the investigator wants to generalize his result

Sampling Unit:

smallest unit from which sample can be selected

Sampling frame

List of all the sampling units from which sample is drawn

Sampling scheme

Method of selecting sampling units from sampling frame

Types of sampling

1-Simple random sampling

2-Systematic sampling

3-Stratified sampling

4-Multi-stage sampling

5-Cluster sampling

1-Simple random sampling

Principle

Equal chance/probability of each unit being drawn

Procedure

Take sampling population

Need listing of all sampling units (“sampling frame”)

Number all units

Randomly draw units

Advantages

Simple Sampling error easily measured

Disadvantages

Need complete list of units

Units may be scattered and poorly accessible

Heterogeneous population

important minorities might not be taken into account

2-Systematic sampling

Principle

Select sampling units at regular intervals (e.g. every 20^{th}unit)

Procedure

Arrange the units in some kind of sequence

Divide total sampling population by the designated sample size (eg 1200/60=20)

Choose a random starting point (for 20, the starting point will be a random number between 1 and 20)

Select units at regular intervals (in this case, every 20^{th} unit), i.e. 4^{th}, 24^{th}, 44^{th} etc.

Advantages

Ensures representativity across list

Easy to implement

Disadvantages

Need complete list of units

Periodicity-underlying pattern may be a problem
(characteristics occurring at regular intervals)

3-Stratified sampling

When to use

Population with distinct subgroups

Procedure

Divide (stratify) sampling frame into homogeneous subgroups (strata) e.g. age-group, urban/rural areas, regions, occupations

Draw random sample within each stratum

Advantages

Can acquire information about whole population and individual strata

Precision increased if variability within strata is smaller (homogenous) than between strata

Disadvantages

Sampling error is difficult to measure

Different strata can be difficult to identify

Loss of precision if small numbers in individual strata (resolved by sampling proportional to stratum population

4-Multiple stage sampling

Principle:

Consecutive sampling

Example :

sampling unit = household

1st stage: draw neighbourhoods

2nd stage: draw buildings

3rd stage: draw households

5-Cluster sampling

Principle

Cluster: a group of sampling units close to each other .Whole population divided into groups .

A type of multi-stage sampling where all units at the lower level are included in the sample

Random sample taken of these groups (“clusters”)

Within selected clusters, all units e.g. households included (or random sample of these units)

Provides logistical advantage

Advantages

Simple as complete list of sampling units within population not required

Less travel/resources required

Disadvantages

Cluster members may be more alike than those in another cluster (homogeneous)

this “dependence” needs to be taken into account in the sample size and in the analysis (“design effect”)

sample size

Sample size calculations are important to ensure that estimates are obtained with required precision or confidence.

$n=z^2/e^2*p(1-p)$

n= the sample size

z= stander error associated with the chosen level of confidence

z is the z-score, e.g. 1.645 for a 90% confidence interval, 1.96 for a 90% confidence interval, 2.58 for a 99% confidence interval

e= acceptable sample error

p = is the estimated proportion of an attribute that is present in the population, Assume there is a large population but that we do not know the variability in the proportion that will adopt the practice; therefore, assume p=.5 (maximum variability).