Dimensions
Research & Marketing Consultancy
Empowerment through knowledge

Research Sampling

Objectives:

- Discuss methods of sampling.
- Identify major sources of measurement error.
- Critically evaluate the instrumentation used in a nursing research study.
- Identify types of information suitable for self-report instruments.
- Identify type of semi-structured self-report techniques
- Compare the advantages and disadvantages of interviews and questionnaires.

Lesson: Research Sampling

Sampling: is a process used to study a response to an intervention in a small population that can be applied to a larger population. Some terms to become familiar with are listed and explained below.

1. Element:
   An element is the most basic unit on which information will be collected - individuals, chart records, etc.. For example: In a study of nursing delivery system in hospital that has three different units using 3 different delivery systems, the elements are the first floor case method, the second floor team method, and the third floor partnership model
2. Population

   A poulation is a set of individuals that meet sampling criteria

   The target population is the entire set of population that the researcher would like to make generalizations about

   The accessible population is the one that meets the criteria established and is also accessible, considering constraints of time, money, researcher availability

3. Generalizability

   Generalizability is extending findings from the sample to the larger population

4. Sampling Criteria

   A well defined set that meets very specific criteria

1. criteria must be very well defined
2. must have limiting factors so that persons not meeting the criteria will be excluded
3. must be able to control for homogeneity by excluding from the desired population anyone who would bring in a confounding variable

5. Representativeness

   The extent to which the sample and the population are alike

6. Sampling Unit

   The selection of a portion of the target population that will represent the entire population

Types of Sampling

NONPROBABILITY SAMPLING
Uses a non random method to select the sample - you cannot be assured that every element available is fairly represented in the sample

1. CONVENIENCE SAMPLING

   Uses the most readily available subjects and is the easy method to obtain subjects

   Example: all students enrolled in a nursing program; first 200 patients admitted to a nursing unit

   Problem: risk of bias is very great
   sample tends to be self selecting:

   what motivated people to volunteer?
   what sample of the population is missed because they were not available?

2. QUOTA SAMPLING
   Knowledge about the population is used to build some design into the sample
   Each stratum of the population is represented proportionally
   Must base sampling on previous knowledge:from a literature review

   Example: you want to study attitudes of nurses about use of nursing diagnosis what type of samples would you think would be important to include? level of education ; years in practice as a nurse

   Problem: Even these techniques do not assure that no bias may be present - in the above example, what variable could affect a nurse's willingness to participate in the study?

4. PURPOSIVE SAMPLING

Researcher handpicks subjects to participate in the study based on identified variables under consideration. Used when the population for study is highly unique

Example: Parents of children with Tay Sach's disease

Problems: Must assume that errors of judgment in ranges of the sample will tend to even out - as many subjects who are at the far ends of the population will cancel each other out

Uses for purposive sampling

1. validation of a test or instrument with a known population
2. collection of exploratory data from an unusual population
3. use in qualitative studies to study the lived experience of a specific population

How does purposive and quota sampling differ?

Purposive restrict the sample population to a very specific population and then tends to use all of the subjects available

PROBABILITY TESTING
Random selection of subjects from a specific population

SIMPLE RANDOM SAMPLING

population if defined listing all of the descriptors identify all populations that meet the descriptors and give each a number use a table of random numbers to select population for study, read off numbers in any fixed direction

Advantages: researcher bias cannot operate representation of the desired population is maximized probability of selecting a nonrepresentative sample is decreased as the sample size is increased

Disadvantages: very time consuming
it may be impossible to obtain a list of every person eligible to be part of the population under study

1. STRATIFIED RANDOM SAMPLING

uses a quota for subsets to ensure that all subgroups are fairly represented similar to proportional quota sampling except that a random approach is used to select the sub populations

Example: see diagram of study on registered nurses

Questions to be addressed

1. what is the logical basis for selecting the subsets?
2. do you have sufficient information available to divide population into subsets

3. should each subset be equal in size or should the size be based on the frequency in the population?
4. are there enough subjects to get meaningful groups into each subset?
5. have random procedures ben used to select subjects for each of the subsets?

Problems: similar to a simple random design in terms of stability to identify appropriate subjects

greater because of the need for greater numbers of subjects to fill each of the subsets

2. CLUSTER SAMPLING
used to break up large groups into smaller workable models

Example: The researcher wants to examine nursing practices in county health departments

Stage 1: identify all states - each will be a sampling group - randomly select a certain percentage of states

Stage 2: select a random sample of subjects from the first sample: a random sample of county health departments within the states selected

Stratified random sampling technique could be used by looking at counties based on rural vs urban, etc.

Advantages: more economical of time and money

Disadvantage: sampling error can creep in

3. SYSTEMATIC SAMPLING

select every nth subject from a list of all possible subjects - example: every 5th patient admitted to the hospital

the population listing must be random - example a list of nurses by alphabetical order

the sample selection of the population must start at a random point - if you had an alphabetical listing of all subjects, you would not start with the "A" - but rather with a random point in the list and then go by the nth interval

Sampling interval - determined by the size of the group

n = <u>total population</u> = size of the desired sample

Problems: be sure geographic or cyclic events are not introduced

Example: use of 7 as an interval size when looking at use of a facility

Geographic regions that happen to vary with the interval size: every 3rd room being a private room as compared to double rooms

4. MATCHED SAMPLING
   used to obtain equivalent comparison groups: match on characteristics such as age, sex, schooling, etc.

SAMPLE SIZE

Power Analysis

In quantitative studies, the larger the sample the greater likelihood will it be non biased

In qualitative studies, the sample size is generally very small

The sample size will be indicated by the type of statistical tools that will be used

The degree of precision needed will help to determine sample size

The smaller the expected differences in subject response to the intervention, the large the sample size needed to demonstrate a significantly different response

If the study has been well designed, a smaller sample size can produce good results

Population and Sample

The major use of inferential statistics is to use information from a **sample** to infer something about a **population**.

A **population** is a collection of data whose properties are analyzed. The population is the *complete* collection to be studied, it contains *all* subjects of interest.

A **sample** is a *part* of the population of interest, a sub-collection selected from a population.

A **parameter** is a numerical measurement that describes a characteristic of a population, while a **sample** is a numerical measurement that describes a characteristic of a sample. In general, we will use a statistic to infer something about a parameter.

Ex. Joe D. Politician is running for President. He calls you on the phone and asks you to find out what percentage of the registered voters in the country will vote for him. There are a few things you could try.

Option I : Call *all* registered voters on the phone and ask them who they will vote for.

Although this would provide a very accurate result, it would be a very tedious and time consuming project. All registered voters represent the population of interest here, and a better approach would be to use a sample.

Option II : Call 4 registered voters, 1 in each time zone, and ask them who they will vote for.

Although this is a very easy task, the results would not be very reliable. To use a sample to make inferences about a population, the sample should be representative of the population. How likely is it that these 4 registered voters would represent the population of all registered voters? Not very! The sample needs to look just like the population, but smaller.

Option III : Somewhere between Option I and Option II.

We want to use a method that will be easier than Option I, but more reliable than Option II.

So, you randomly select 2000 registered voters and poll them. 1,120 (56%) tell you that they will vote for Joe.

The population of interest here is all registered voters, and the parameter is the percentage of them that will vote for Joe.

The sample is the 2000 registered voters that were polled, and the statistic is the percentage of them that will vote for Joe.

You can tell Joe that approximately 56% of all registered voters will vote for him.

Ex. In a Statistics class of 40 students, 24 had a credit card with them.

The statement "60% of the students in this Statistics class had a credit card with them" is a descriptive statement. The population is the 40 students in this Statistics class. The 60% represents a parameter.

The statement "60% of the students in all classes have a credit card with them" is an inferential statement. The 40 students in this Statistics class represent a sample of students in all classes. The 60% represents a statistic.