# Sampling

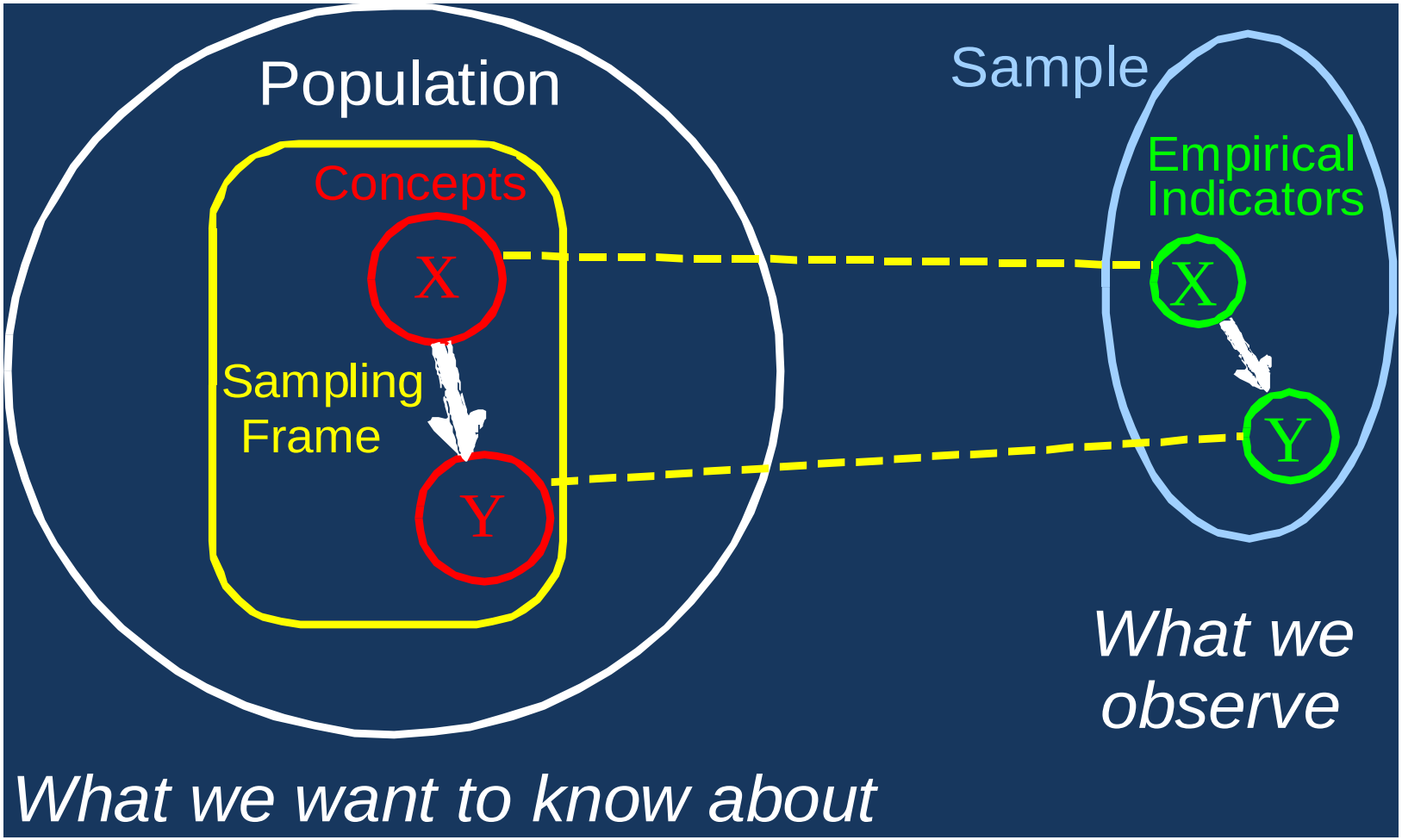
Dr. Sahar Hassan

# A sampling plan

Specifies in advance how participants are to be selected and how many to include.

- It is usually discussed in a report's Method section, sometimes in a subsection called "Sample" or "Study participants.

# The Logic of Sampling & Measurement



# Basic Sampling Concepts

## Population

- The aggregate of cases in which a researcher is interested
- Population is limited to human subjects but it might consists of all the hospital records on file in a particular hospital or all blood samples at a particular laboratory..
- E.g. all children in Canada with cystic fibrosis.

# Basic Sampling Concepts (cont.)

- ✓ Target population: The entire population of interest
- ✓ Accessible population: The portion of the target population that is accessible to the researcher, from which a sample is drawn

# Population Importance

- ✓ Conclusions based on data
- ✓ Data from accessible population
- ✓ Decisions made from study results

# Sampling

- Sampling: selection of a portion of the population to represent the entire population
- Eligibility criteria: The characteristics that define the population
- Inclusion criteria
- Exclusion criteria
- Strata: Subpopulations of a population (e.g., male, female)

# Example of inclusion & exclusion criteria:

- Lindgren and colleagues (2008) studied how elderly patients with ischemic coronary heart disease cluster based on presenting symptoms in the week before hospitalization.
- Patients from five medical centers were eligible for the study if they were:

✓ unpartnered adults aged 65 or older;

✓ had positive enzyme tests considered diagnostic for myocardial infarction or coronary artery bypass surgery;

✓ read and spoke English;

✓ had a telephone;

✓ lived within 50 miles of the medical center.

# Example

For example, the researcher's target population might be all diabetic patients in the United States, but, in reality, the population that is accessible might be diabetic patients in a particular clinic.

# So…

A sample is a subset of population elements.

Sampling is the process of selecting cases to represent an entire population so that inferences about the population can be made.

# Steps in the sampling process

1) Identify the target population

2) Identify the accessible population

3) Determine the size of the sample needed

4) Select the sampling technique

5) Implement the plan

# Critical Factor

- The sample needs to be representative of your population of interest

- Generalizability (external validity) of your results is dependent on this factor!

# Randomization

1) ensures representativeness

2) unbiased selection

3) to equalize characteristics across experimental & control conditions

# Terms

- Random selection: sample is representative of larger population
- Random assignment: involves equalizing experimental groupings (essential for internal validity of a study)

# Sampling Goal in Quantitative Research

## Representative sample

- A sample whose key characteristics closely approximate those of the population

# Representative sample

But there is never a guarantee of a representative sample.

More easily achieved with:

- ✓ Probability sampling
- ✓ Homogeneous populations
- ✓ Larger samples

# Sampling Problems in Quantitative Research

- ✓ Sampling bias: The systematic over- or under representation of segments of the population on key variables.
- ✓ It is affected by: Homogeneity of population
- Age, blood pressure & stress level are all attributes that reflects heterogeneity of human

# Sampling Bias (example)

- As an example of consciously biased selection, suppose we were investigating patients' responsiveness to nurses' touch & decide to use as our sample the first 50 patients meeting eligibility criteria in a specific hospital unit.
- We decide to omit Mr. Z from the sample because he has shown hostility to nurses. Mrs. X, who has just lost a spouse, is also excluded from the study because she is under stress.

# Sampling Bias (example) cont.

- We have made conscious decisions to exclude certain individuals, and the decisions do not reflect bona fide حسن النية
- Eligibility criteria
- This can lead to bias because responsiveness to nurses' touch (the dependent variable) may be affected by patients' feelings about nurses or their emotional state.

Sampling error: Differences between sample values and population values

# Strata

- Subpopulations of a population (e.g., male/female)
- QoL: 35-50; >50-65; >65
- often used in sample selection to enhance the sample's representativeness

- For instance, a population consisting of all RNs in the US could be divided into two strata based on gender (M/F)
- Alternatively, we could specify three strata (nurses < 30 years of age, nurses aged 30-45 years, & nurses 46 years or >
- Strata are often used in sample selection to enhance the sample's representativeness

# Sampling Techniques in Quantitative Studies

## Nonprobability sampling

- ✓ Does not involve selection of elements at random

## Probability sampling

- ✓ Involves random selection of elements: Each element has an equal, independent chance of being selected.

# Types of Nonprobability Sampling: Quantitative Research

- Convenience sampling
- Quota sampling
- Consecutive sampling
- Purposive sampling

# Nonprobability Sampling

Convenience Sampling entails using the most conveniently available people as participants

A group of participants to whom the researcher has access, for example, patients on a ward.

A faculty member who distributes questionnaires to nursing students in a class is using a convenience sample, or an accidental sample, as it is sometimes called

# Convenience

- ✓ Poor approach
- ✓ The right place at the right time
- ✓ Example:

A class room of students;

Pt who attend a clinic on a specific day

## Example of a convenience sample:

- Fraser and Polito (2007) compared the self-efficacy of men versus women with multiple sclerosis (MS). They used a convenience sample of 556 individuals with MS.
- Stopping people at a street corner to conduct an interview is sampling by convenience.

- The problem with convenience sampling is that available subjects might be atypical of the population, and so the price of convenience is the risk of bias

# Nonprobability Sampling

Snowball Sampling (network sampling or chain sampling); used when the population is people with characteristics who might be difficult to identify

- In a snowball sample, participants who are already part of the sample are asked to identify others who would possibly be suitable for inclusion in the study and who would be agreeable to taking part in it
- In other words, the sample gradually increases in size, like a snowball being rolled down a hill
- This type of sample is useful when the researcher is studying a subgroup who may not easily be accessible otherwise, for example drug users
- Network sampling

# Nonprobability Sampling

- Consecutive Sampling involves recruiting *all* of the people from an accessible population who meet the eligibility criteria over a specific time interval
- For example, in study of ventilated-associated pneumonia in ICU patients, if the accessible population were patients in an ICU of a specific hospital, consecutive sample might consist of all eligible patients who were admitted to that ICU over a 6-month period Or it might be the first 250 eligible patients admitted to the ICU, if 250 were the targeted sample size

- Consecutive samples can be selected either for a retrospective or prospective time period
- For example, the sample could include every patient who visited a diabetic clinic in the previous 30 days (retrospective or prospective)?
- Or, it could include all of the patients who will enroll in the clinic in the next 30 days (retrospective or prospective)??

# Nonprobability Sampling

- Purposive sampling is often used when researchers want a sample of experts, as in the case of a needs assessment using the key informant approach or in Delphi surveys
- Purposeful Sampling (judgmental sampling) is based o the belief that researchers' knowledge about the population can be used to hand-pick sample members

# Example of purposive sampling

- Van den Heede and colleagues (2007) assessed the views of an international panel of experts regarding the state of nurse staffing & patient outcomes research
- Two rounds of surveys were conducted with a purposively selected sample of researchers & nurse administrators from 10 countries.

# Nonprobability Sampling

- Quota Sampling: researchers identify population strata and determine how many participants are needed from each stratum
- By using information about population characteristics, researchers can ensure that diverse segments are adequately represented in the sample.

# Example 1: quota Sample

- For instance, if a researcher is required by a quota sampling plan to interview 10 men between the ages of 65 & 80 years, a trip to a nursing home might be the most convenient method of obtaining those subjects
- Yet this approach would fail to represent the many senior citizens who live independently in the community.

# Example 2: quota sample

- Williams et al (2000) studied mothers' expectations for children' development
- The researchers used quota sampling to ensure an equal number of urban & rural mothers, and an equal number of male and female children.

# Evaluation of Nonprobability Sampling

- Although a no-nprobability sample is often acceptable for pilot, exploratory, or in-depth qualitative research, for most quantitative studies, the use of nonprobability samples is problematic
- Nonprobability samples are rarely representative of the population
- When every element in the population does not have a chance of being included in the sample, it is likely that some segment of it will be systematically under represented.

# Evaluation of Nonprobability Sampling

- Why, then, are nonprobability samples used in most nursing studies?
- The advantage of these sampling designs lies in their convenience & economy
- Probability sampling, requires skill & resources
- There is often no option but to use a nonprobability approach or to abandon the project altogether
- Even hard-nosed research methodologists would hesitate to advocate the abandonment of an idea in the absence of a random sample.

# Probability sampling

# Types of Probability Sampling

- ✓Simple random sampling
- ✓Stratified random sampling
- ✓Systematic sampling
- ✓Cluster sampling

- Probability sampling involves the random selection of elements from a population.

- Random assignment refers to the process of allocating subjects to different treatment conditions at random.
- Random *assignment* has no bearing on how subjects in an experiment were selected in the first place.

- Simple Random Sampling: in this technique every unit of the study population has an equal and independent chance of selection.
- Randomization software

https://www.randomizer.org/

# Steps for Simple Random Sampling

- Identify the accessible population or list of elements
- Choose the method for getting the sample
- An easy example:

✓ Names of elements on slips of paper

✓ Papers are placed into a hat

✓ Individual draws a slip of paper

✓ Individual continues till sample number is met

# Example of a simple random sample

- Nachreiner and colleagues (2007) conducted a survey of registered nurses & licensed practical nurses to compare their experience with workplace violence.
- Questionnaires were mailed to a random sample of 6,300 licensed nurses in Minnesota.

# Simple Random Sampling

- Advantages:

✓ Little knowledge of population is needed

✓ Most unbiased of probability method

✓ Easy to analyze data and compute errors

- Disadvantages:

✓ Complete listing of population is necessary

✓ It is time consuming to use

# Stratified Random Sampling

- Population divided into strata, then random selection from the stratified sampling frames
- Enhances representativeness
- Can sample proportionately or disproportionately from the strata

## Variable commonly used for stratification:

- Age; gender; ethnicity; socioeconomic status; diagnosis; geographic area; type of institution; type of care; and site of care.

## Stratified Random Sampling

- The most common procedure for drawing a stratified sample is:

✓ to group together elements belonging to a stratum & to select randomly the desired number of elements

✓ You can either select an equal number of elements from each stratum or select unequal numbers

- To illustrate the procedure used in the simplest case, suppose that the list consisted of 25 men (numbers 1 through 25) and 25 women (numbers 26 through 50)
- Using gender as the stratifying variable, we could guarantee a sample of 10 men and 10 women by randomly sampling 10 numbers from the first half of the list and 10 from the second half.

- Stratifying variables usually divide the population into unequal subpopulations
- For example, if the person's race were used to stratify the population of U. S. citizens, the subpopulation of white people would be larger than that of African-American and other nonwhite people.
- The researcher might decide to select subjects in proportion to the size of the stratum in the population, using proportionate stratified sampling.

- **If the population** was students in a nursing school that had 10% African-American students, 10% Hispanic students, and 80% white students, then a proportionate stratified sample of 100 students, with racial/ ethnic background as the stratifying variable, would consist of 10, 10, and 80 students from the respective strata

- When researchers are interested in understanding differences among strata, proportionate sampling may result in insufficient numbers for making comparisons.
- In the previous example, would the researcher be justified in drawing conclusions about the characteristics of Hispanic nursing students based on only 10 cases?

- It would be unwise to do so.
- For this reason, researchers often adopt a disproportionate sampling design when comparisons are sought between strata of greatly unequal size.
- In the example, the sampling proportions might be altered to select 20 African-American students, 20 Hispanic students, and 60 white students.
- This design would ensure a more adequate representation of the two racial/ethnic minorities.
- When disproportionate sampling is used, however, it is necessary to make an adjustment to the data to arrive at the best estimate of *overall population values. This adjustment process,* known as weighting.

# Example of stratified random sampling

- Ekwall and Hallberg (2007) studied caregiver satisfaction among informal older caregivers who participated in a mail survey in Sweden. The sample was stratified on the basis of age.
- Questionnaires were mailed to:
- 2,500 elders aged 75 to 79,
- 2,500elders aged 80 to 84,
- 2,000 elders aged 85 to89,
- and 1,500 elders aged 90 and over

# Cluster sampling

- ✓ **K**nown as multistage sample because it involves taking samples in stages, from the more general to the more detailed unit
- ✓ It is particularly useful when widely geographic areas to be sampled by personal interview
- ✓ In cluster sampling, there is a successive random sampling of units
- ✓ The first unit is large groupings, or clusters

# Cluster sampling [Example1]

- ✓In drawing a sample of nursing students, we might first draw a random sample of nursing schools and then draw a random sample of students from those schools.

# Cluster sampling [Example 2]

- ✓ A researcher want to study the prenatal care at the MCH centers all over Palestine:
- ✓ The first step is to divide Palestine into clusters (North, south, center)
- ✓ Second the researcher need to decide how many centers in each cluster wants to include

# Cluster sampling [Example 2]

- ✓ Third, the MCH centers needed are chosen by using simple random sample

- ✓ In many cases the each cluster may need to be clustered again, for example, the governmental MCH & the Non-Governmental ones.

# Disadvantages of Cluster sampling

- ✓ Causes a larger sampling error
- ✓ The appropriate handling of the statistical data from cluster is very complex
- ✓ Requires each member assignment of population to cluster
- ✓ Uses a more complicated statistic analysis

# Advantages of Cluster Random Sampling

- Saves time and money
- Arrangements made with small number sampling units
- Characteristics of clusters/population can be estimated

# Systematic sampling

- Involves the selection of every *kth* case from a list, such as every 10th person on a patient list.
- Where the difference between any consecutive number is the same
- Systematic sampling designs can be applied in such a way that an essentially random sample is drawn.

# Systematic Random Sampling

Process

- ✓Obtain a listing of population
- ✓Determine the sample size
- ✓Determine the sampling interval ($k = N/n$)
- ✓Select random starting point
- ✓Select every $k$th element

# Example on systematic sampling

- For instance, if we wanted a sample of 50 from a population of 5,000, our sampling interval would be 100 (5,000/50 = 100).
- In other words, every 100th case on the sampling frame would be sampled
- Next, the first case would be selected randomly

(e.g., by using a table of random numbers)

- If the random number chosen were 73, the people corresponding to numbers 73, 173, 273, and so forth would be included in the sample.

## Example of a systematic sample:

- Houghtonand colleagues (2008) surveyed nurse anesthesia bout their practices and attitudes regarding smoking intervention
- Using the membership list of the American Association of Nurse Anesthetists, every 30th name in the alphabetized list was selected for the sample

# Systematic Random Sampling

## Advantages

- ✓ Easy to draw sample
- ✓ Economical
- ✓ Time-saving technique

## Disadvantages:

- ✓ Samples may be biased
- ✓ After first sample chosen, no longer "equal chance"

# Special Sampling Strategy

- ✓ One of the most common strategy is “Matching”
- ✓ Matching is used to construct an equivalent comparison sample group by filling it with subjects who are similar to each subject in another sample group in relation to such pre-established variables such as sex, age, marital status, education...etc.
- ✓ In some studies, in order to ensure high control and representative sample both matching & randomization are used.

# Evaluation of Probability Sampling

- The only viable method of obtaining representative samples. If all the elements in the population have an equal probability of being selected, then the resulting sample is likely to do a good job of representing the population
- Probability sampling allows researchers to estimate the magnitude of sampling error.
- Sampling error refers to differences between population values (such as the average age of the population) and sample values (such as the average age of the sample).

# Sample Size

- The number of study participants in the final sample
- Sample size adequacy is a key determinant of sample quality in quantitative research
- The larger the sample, the more representative it is likely to be
- The larger the sample, the smaller the sampling error
- There are no simple formulas that can tell you how large a sample you will need in a given study

# Sample Size

- when samples are too small, and researchers run the risk of gathering data that will not support their hypotheses—even when those hypotheses are correct
- When critiquing quantitative studies, you must assess both the sample size & the sample selection method to judge how representative the sample likely was.

# Sample Size

- Qualitative studies almost always use small, nonrandom samples.
- Qualitative researchers ask such sampling questions as:

✓ "Who would be an information-rich data source for my study?"

✓ "To whom should I talk, or what should I observe, to maximize my understanding of the phenomenon?"

# Saturation

- A guiding principle in sampling is data saturation—that is, sampling to the point at which no new information is obtained and redundancy is achieved.

# Sampling Bias

- Bias when samples are not carefully selected
- All nonprobability sampling methods have it
- It may occur in probability sampling methods
  - ✓Subjects decide not to participate when chosen
  - ✓Final sample is now not representative of population

- END