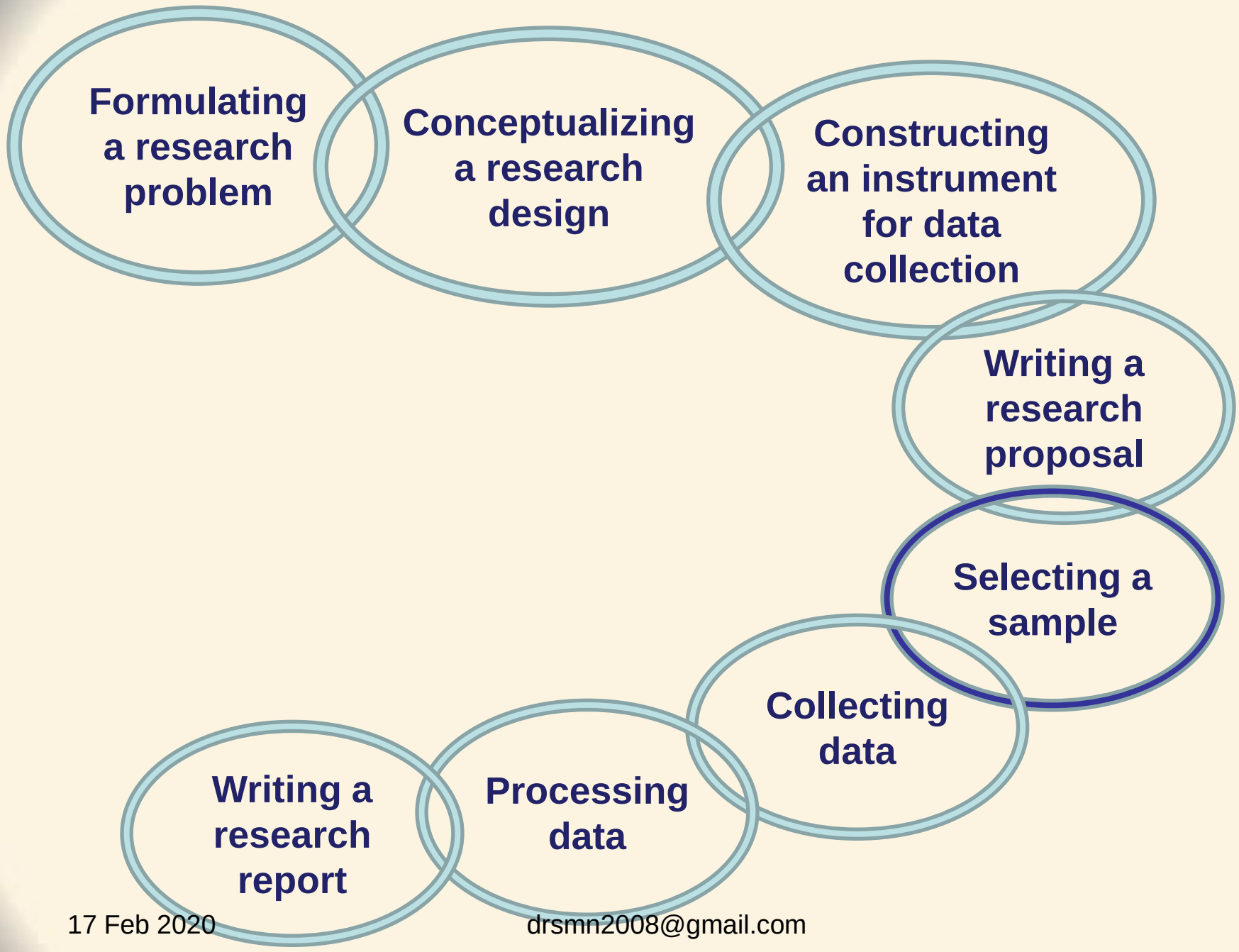
Sampling

Prof. Soe Moe Naing

Professor & Head

P & SM Department

University of Medicine 2 Yangon


Formulating a research problem
Conceptualizing a research design
Constructing an instrument for data collection
Writing a research proposal
Selecting a sample
Collecting data
Processing data
Writing a research report

Population & Samples

- "Population" in bio-statistics means "an aggregate of **ALL** persons, events, or procedures with a common characteristics

- "population = as the largest collection of entities for which we have an interest at a particular time"

- one issue is to consider whether it would be necessary or feasible to study all the population

- to study the entire population; it would not only be expensive but also not improve the quality of the study

- It is usual to study a sample of the population on the basis that such a sample is truly representative of all the characteristics of the population

- A representative sample is necessary to be selected with the appropriate sampling technique and large enough sample size

➢• Population interest of the research is defined by;

- Eligible area in “Study site/ area”
- Eligible period in “Study period”
- Eligible criteria or characteristics in “Study population”

Sampled population (Study Population)

- The sampled population is the population from which one actually draws a sample.

Target population

- The target population is the population about which one wishes to make an inference.

Population

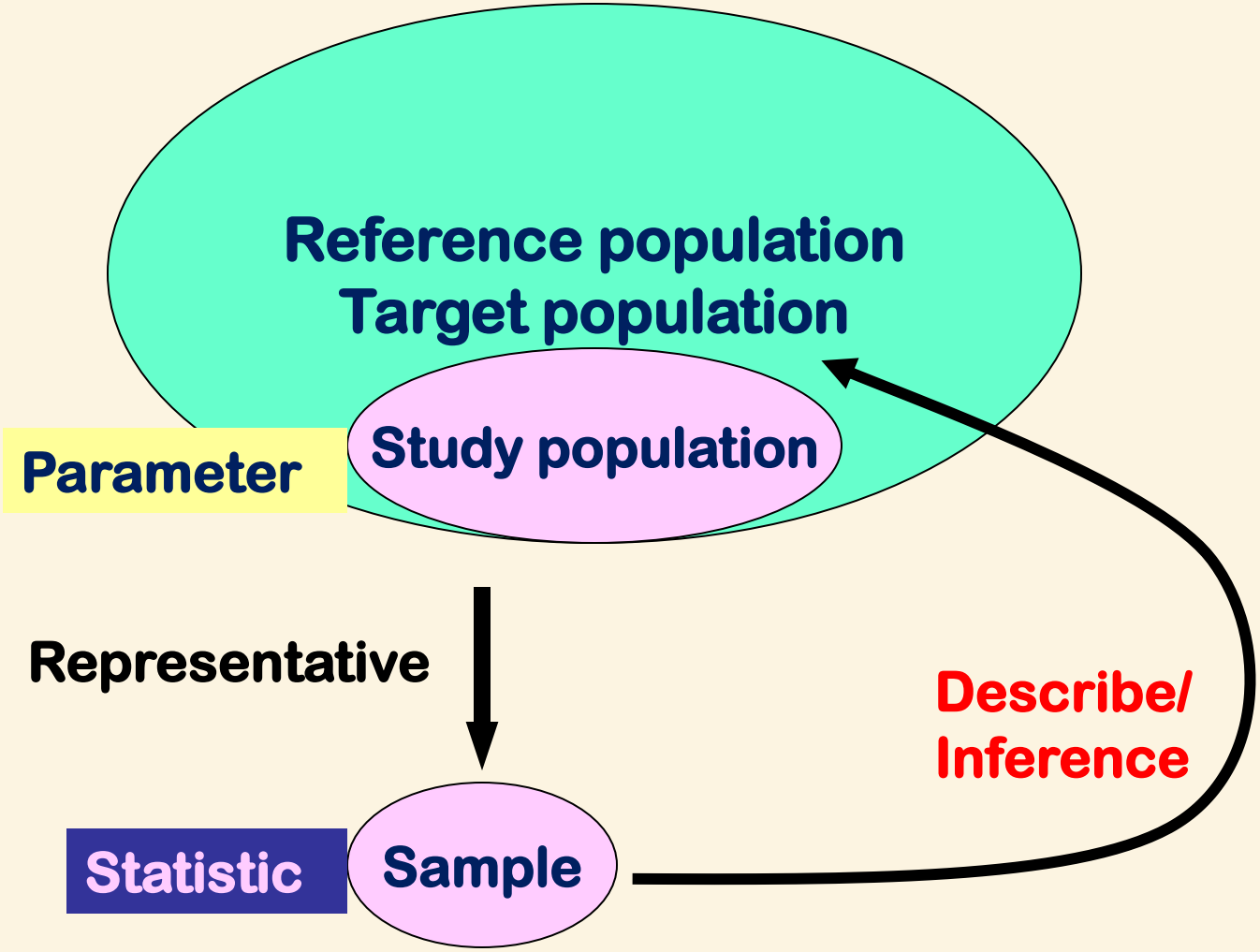
Finite population

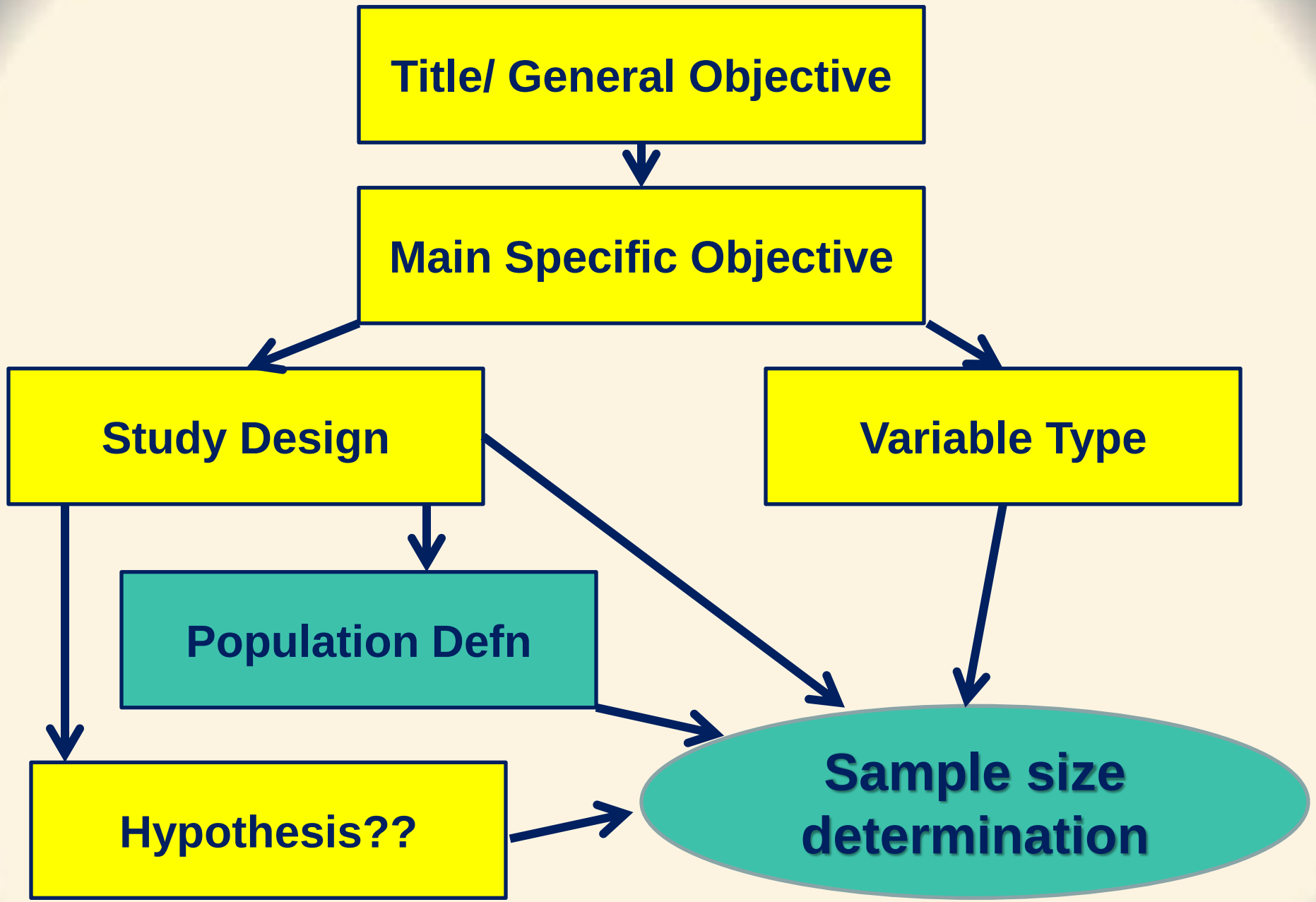
- If a population of values consists of a **fixed** number of these values, the population is said to be finite.

Infinite population

- a population consists of an **endless** succession of values, the population is an infinite one

WAYNE W. DANIEL, PH.D.
Professor Emeritus
Georgia State University


Reference population
Target population
Parameter
Study population
Representative
Describe/
Inference
Statistic
Sample


Title/ General Objective
Main Specific Objective
Study Design
Variable Type
Population Defn
Hypothesis??
Sample size determination

Example

Main specific objective

- **to compare the serum X level between patients with disease A and healthy persons**

Hypothesis

- **serum X level is associated with presence of disease A**

Study design

- Cross –sectional analytical study

Study Site

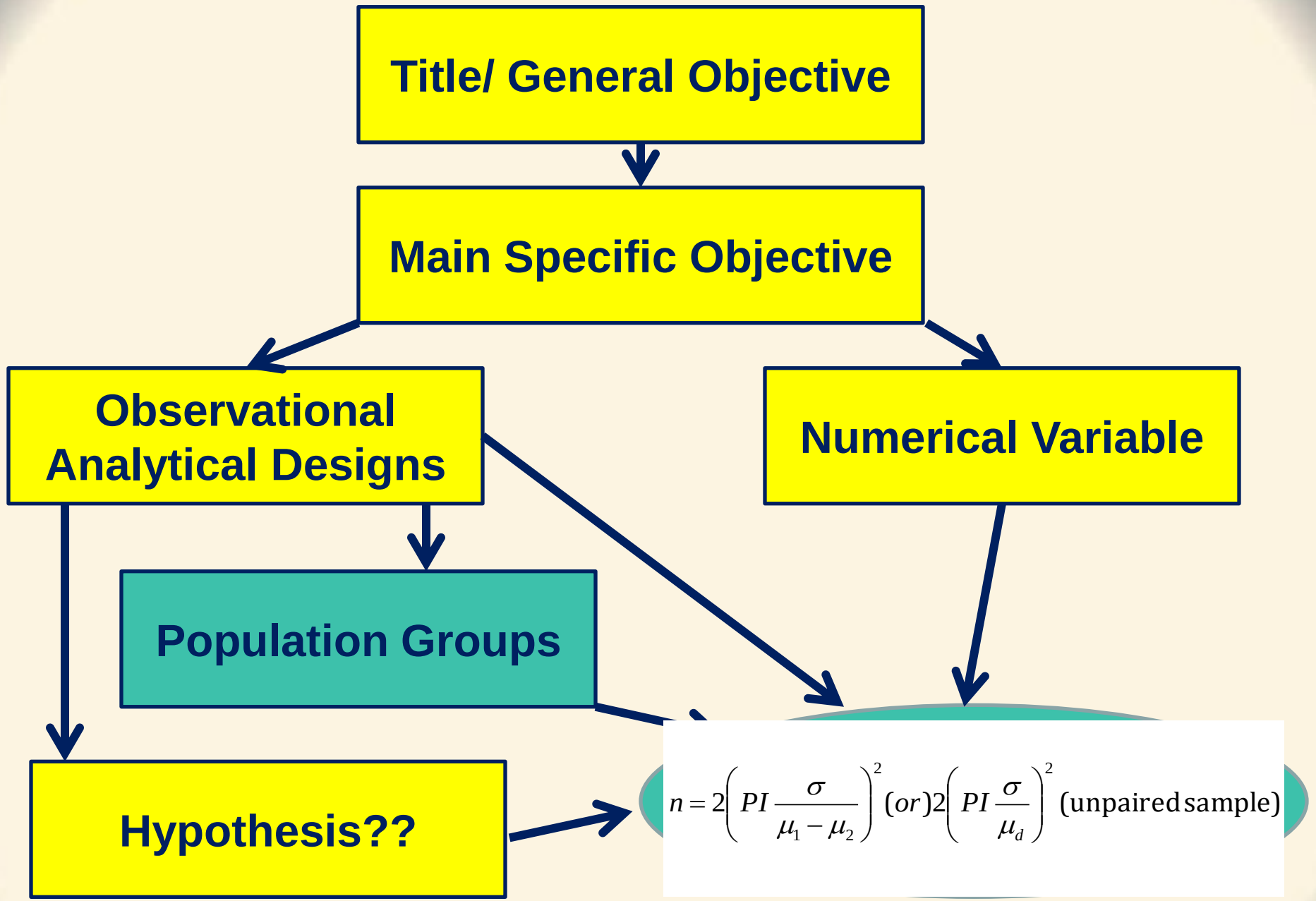
- --- Unit, ---- Hospital, selected ward in ---- Township

Study Period

- from January 2020 to December 2020

Study Population

- patient with disease X attending to study site and age-sex matched healthy persons


Title/ General Objective
Main Specific Objective
Observational Analytical Designs
Numerical Variable
Population Groups
Hypothesis??
$n = 2\left(PI\frac{\sigma}{\mu_1 - \mu_2}\right)^2 (or) 2\left(PI\frac{\sigma}{\mu_d}\right)^2 (\text{unpaired sample})$

$$n = 2\left(\frac{PI\,\sigma}{\mu_1 - \mu_2}\right)^2 (or) 2\left(\frac{PI\,\sigma}{\mu_d}\right)^2 (\text{unpaired sample})$$

Result

n = 60 (minimum for each group)

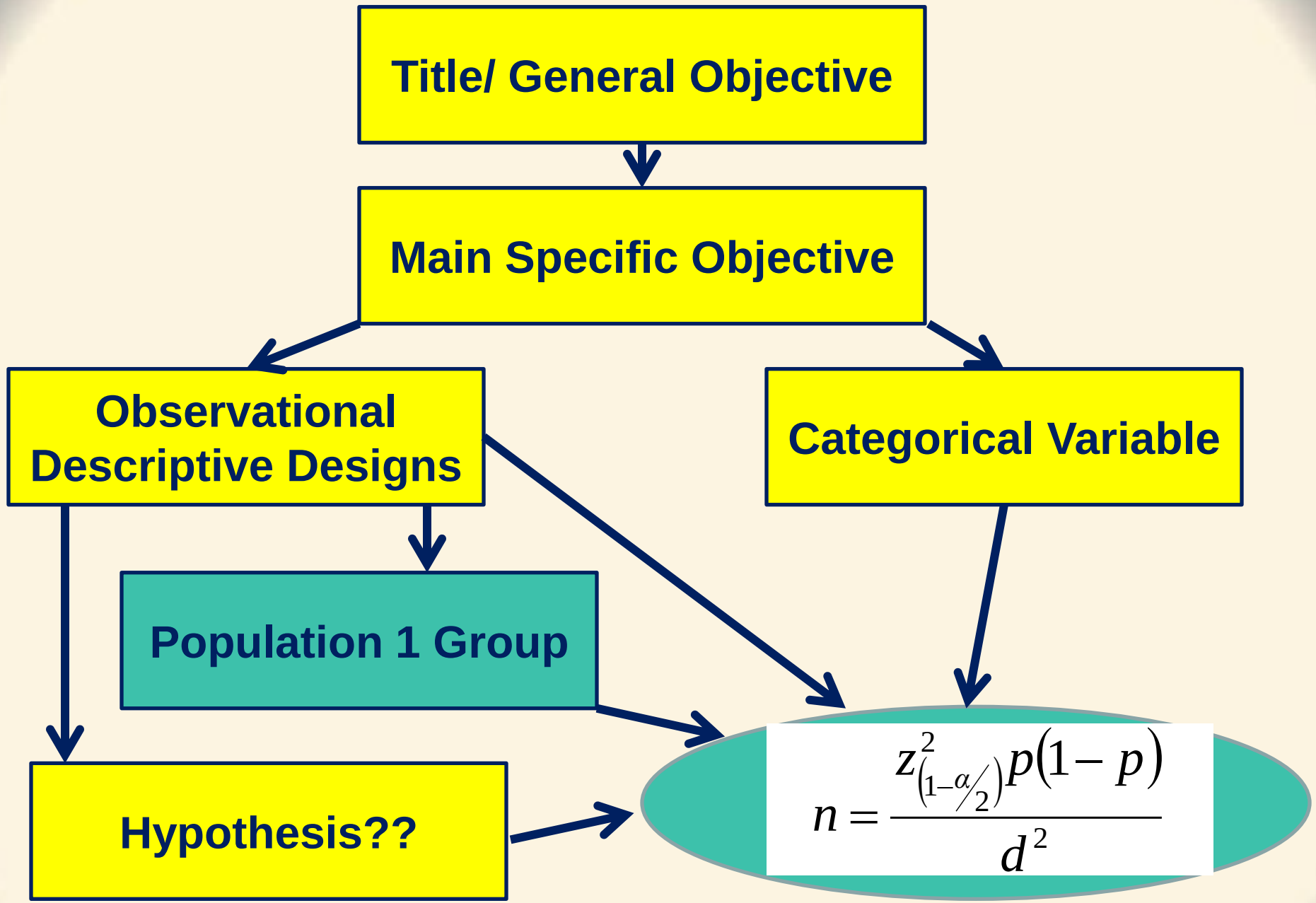
Total = 120

Effect size

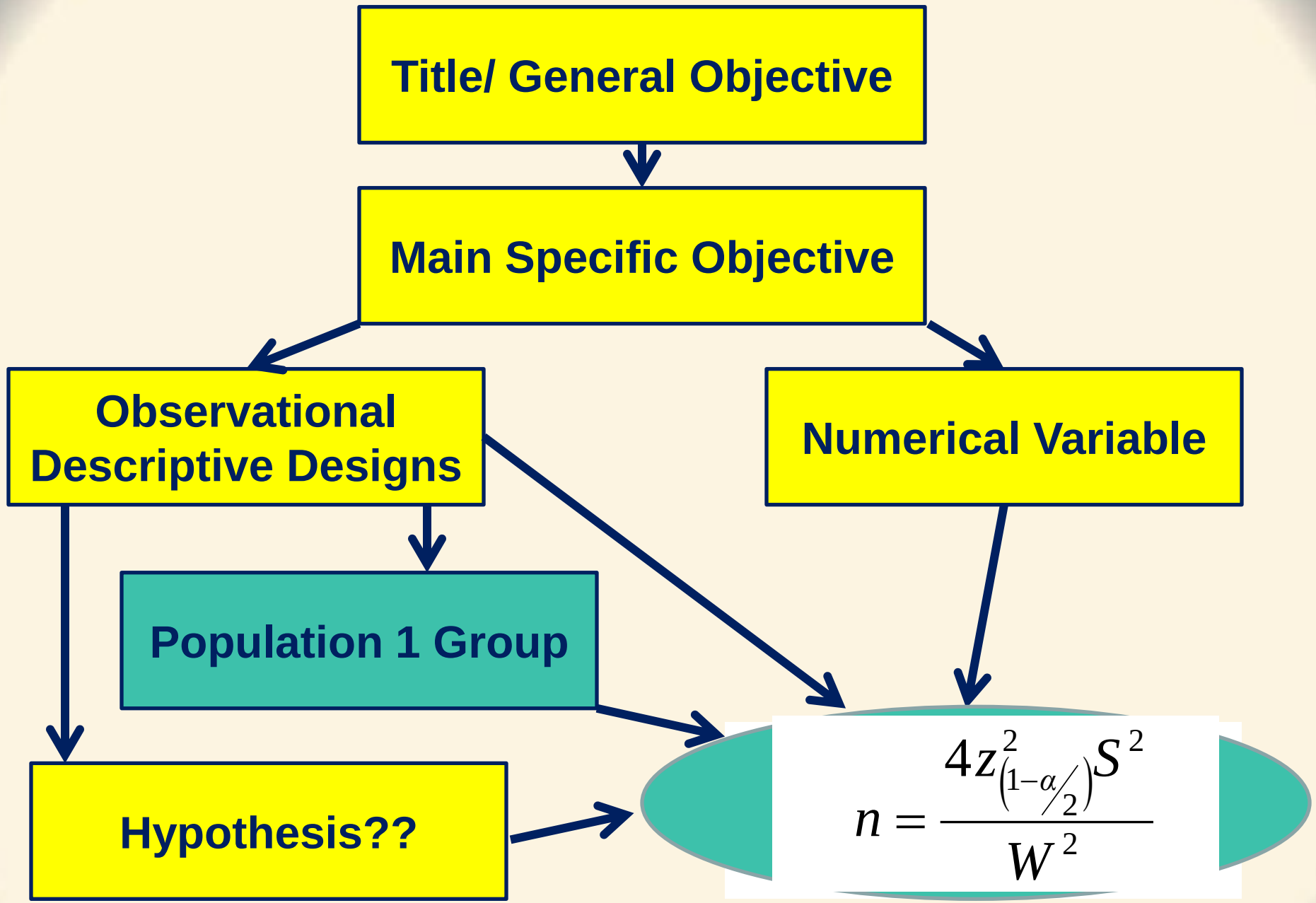
- When **a difference is statistically significant**, it does not necessarily mean that it is big, important, or helpful in decision-making.

- It **simply means** you can be confident that **there is a difference**.

- **To know if** an observed difference is not only statistically significant but also **important or meaningful**, you will **need** to calculate its **effect size**.


Title/ General Objective
Main Specific Objective
Observational Descriptive Designs
Categorical Variable
Population 1 Group
Hypothesis??
$n = \frac{z^2_{\left(1-\alpha/2\right)} p\left(1-p\right)}{d^2}$

$$n = \frac{z^2_{(1-\alpha/2)}\, p(1 - p)}{d^2}$$


Title/ General Objective
Main Specific Objective
Observational Descriptive Designs
Numerical Variable
Population 1 Group
Hypothesis??
$n = \frac{4z^2_{(1-\alpha/2)}S^2}{W^2}$

$$n = \frac{4z^2_{\left(1-\alpha/2\right)}S^2}{W^2}$$

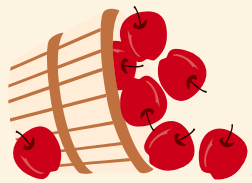
***** There are numerous factors that influence sample size calculation.*****

Sample size in quantitative studies

- The **bigger** the sample, the **better** the study becomes. This is **not necessarily true**.
- It is much better to increase **the accuracy** and richness **of data collection** than to increase sample size after a certain point.
- It is better to make extra **efforts to get a representative sample** rather **than to get** a very **large sample**.

The eventual **sample size** is usually a compromise between what is **desirable** and what is **feasible**

“Sample size determination”



Sampling Methods

1. Probability sampling
2. Non-probability sampling

Probability Sampling

- the probability or **chance** of each sampling unit being selected is the **same** for all sampling unit

Probability sampling

1. Simple Random Sampling (SRS)
2. Systematic Sampling
3. Stratified Sampling
4. Cluster Sampling
5. Multi-stages Sampling

1. Simple Random Sampling

Steps are;

- assign a serial number to each population element (sampling frame must be known)
- selection can be made by drawing serial numbers from a pool containing all serial numbers, e.g. lottery

1. Simple Random Sampling

- this is done until the sample size required is achieved
- alternative methods are table of random numbers and computer

2. Systematic Sampling

Steps are;

- number of each population element is already arranged systematically in serial manner (sampling frame must be known)
- determine population size (N) and sample size (n)
- calculate the sampling interval (K) [K = N/ n]

2. Systematic Sampling

Steps;

- randomly select the random start (r) among the first interval
- e.g. all population size (N) = 1000

  sample size (n) = 200
- K = N/ n = 1000/ 200 = 5
- select a random start (r) among serial numbers 1,2,3,4,5

- if (r) = 3, continue to select as
- 🧍(r + K) = 3 + 5 = 8
- 🧍(r + 2K)= 3+ 10 = 13
- 🧍(r + 3K)= 3 + 15 = 18
- 🧍 until 200 sample size is obtained

3. Stratified Sampling

Steps are;

- population is first divided into two or more strata
- the strata may be based on a single criterion

  e.g. by sex = two strata of 🚹 and 🚺
- in stratified sampling, a simple random sampling is taken from each stratum and the sub-samples are joined to form the whole sample

4. Cluster sampling

Steps are;

- first, formation of clusters, in terms of large grouping, usually a real aggregation of elements; e.g. village, district, township, ward
- second, clusters are selected by simple or stratified random sampling
- third, ultimate selection from the clusters is done by simple or systematic random sampling

Design effect

- the extent to which the expected sampling error with **cluster sampling** in a survey departs from the sampling error that can be expected under **simple random sampling**.

5. Multi-stages Sampling

- Combination of above two or more sampling methods are used as stage by stage for very large survey

Non-probability Sampling

- Purposive sampling
- Convenient sampling
- Snow ball method

Example

Main specific objective

- **to measure the serum A level among patients with disease X**

Research Question

- **What is the mean serum A level among patients with disease X?**

Study design

- **Cross –sectional descriptive study**

Study Site

- **---------**

Study Period

- **from January 2019 to December 2019**

Study Population

- **patient with disease X**

Sampling procedure

- All eligible patients will be selected.

Sampling procedure

- Eligible patients will be selected consecutively until required sample size is obtained.

Sampling procedure

- Two eligible patients will be selected randomly per week.

Learning sources

• Daniel, W. W., & Cross, C. L. (2013). Biostatistics: A foundation for analysis in the health sciences

• Hulley, SB, Cummings, SR, et al. (2001). Designing Clinical Research, Second Edition; Lippincott Williams and Wilkins. -- Chapter 6. Estimating the sample size

• Lwanga, S.K., & Lemeshow, S. (1991). Sample Size Determination in Health Studies: A Practical Manual

• Thabane, L. (2004). Sample Size Determination in Clinical Trials

• Thomas H.H. (1991). Understanding Biostatistics

Thank you for your attention