BIOSTATISTICS

Dr. Hamza Aduraidi

Unit One

INTRODUCTION

Biostatistics

- **It** can be defined as the application of the mathematical tools used in statistics to the fields of biological sciences and medicine.
- It is a growing field with applications in many areas of biology including epidemiology, medical sciences, health sciences, educational research and environmental sciences.

Concerns of Biostatistics

- Biostatistics is concerned with collection, organization, summarization, and analysis of data.
- We seek to draw inferences about a body of data when only a part of the data is observed.

Purposes of Statistics

- To describe and summarize information thereby reducing it to smaller, more meaningful sets of data.
- To make predictions or to generalize about occurrences based on observations.
- To identify associations, relationships or differences between the sets of observations.

Data

- Data are numbers which can be measurements or can be obtained by counting.
- Biostatistics is concerned with the interpretation of the data and the communication of information about the data.

Populations and Samples

1. A population is the collection or set of all of the values that a variable may have. The entire category under consideration.
2. A sample is a part of a population. The portion of the population that is available, or to be made available, for analysis.

Population and Sampling

- Sampling: the process of selecting portion of the population.
- Representativeness: the key characteristic of the sample is close to the population.
- Sampling bias: excluding any subject without any scientific rational. Or not based on the major inclusion and exclusion criteria.

Example

- Studying the self esteem and academic achievement among college students.
- Population: all student who are enrolled in any college level.
- Sample: students' college at the University of Jordan.

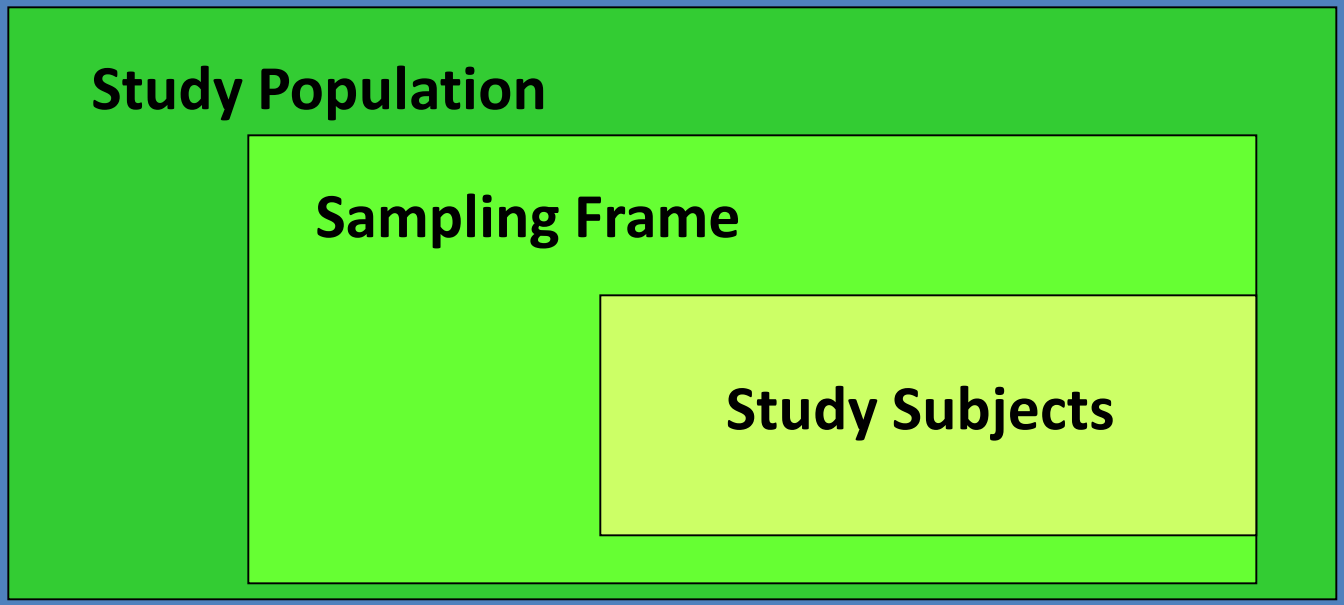
What is sampling?

- Sampling is the selection of a number of study units/subjects from a defined population.

Questions to Consider

- Reference population – to whom are the results going to be applied?
- What is the group of people from which we want to draw a sample (study population)?
- How many people do we need in our sample (Sample Size) ?
- How will these people be selected(Sampling Method)?

Sampling - Populations



Sampling

- Element: The single member of the population (population element or population member are used interchangeably)
- Sampling frame is the listing of all elements of a population, i.e., a list of all medical students at the university of Jordan, 2014-2016.

Sampling Methods

- Sampling depends on the sampling frame.
- Sampling frame: is a listing of all the units that compose the study population.

Types of Sampling Methods

- Probability Sampling Methods. Involves the use of random selection process to select a sample from members or elements of a populations.
  - Simple Random Sampling
  - Systematic sampling.
  - Stratified sampling.
  - Cluster sampling.
  - Multistage sampling.

Probability Sampling Methods

- Involves random selection procedures to ensure that each unit of the sample is chosen on the basis of chance
- All units of the study population should have an equal or at least a known chance of being included in the sample
- Requires a sampling frame
  - Listing of all study units

Simple Random Sampling

- This is the simplest of probability sampling
  - Make a numbered list of all units in the population
  - Decide on the sample size
  - Select the required number of sampling units using the lottery method or a random number table

TABLE 10–2. Random Numbers

21	71	89	96	97
82	59	22	78	12
76	93	64	79	28
20	60	70	34	51
93	58	36	93	90
68	63	19	21	91
18	32	36	27	71
58	80	(58)	67	50
66	25	(20)	31	62
17	25	(07)	94	18
02	29	(30)	15	92
55	06	(25)	09	26
38	11	(01)	47	93
42	47	(73)	25	84
82	04	(23)	08	88
37	24	(51)	98	05
94	58	85	86	71
37	92	(27)	20	58
29	64	(13)	05	24
85	48	(37)	37	21
20	56	91	53	66
33	23	13	82	54
62	11	(29)	17	37
01	57	73	53	97
34	19	(75)	62	16
81	10	(55)	36	36
92	50	32	68	82
37	33	43	20	08
10	50	18	85	27

Systematic Sampling

- Individuals are chosen at regular intervals from the sampling frame
- Ideally we randomly select a number to tell us the starting point
  - every 5th household
  - every 10th women attending ANC
- Sampling fraction = $\frac{\text{Sample size}}{\text{Study population}}$
- Interval size= $\frac{\text{study population}}{\text{Sample size}}$

Stratified Sampling

- If we have study units with different characteristics which we want to include in the study then the sampling frame needs to be divided into strata according to these characteristics
- Ensures that proportions of individuals with certain characteristics in the sample will be the same as those in the whole study population
- Random or systematic samples of predetermined sample size will have to be obtained from each stratum based on a sampling fraction for each stratum

Cluster Sampling

- Selection of study units (clusters) instead of the selection of individuals
- All subjects/units in the cluster who meet the criteria will be sampled.
  - Clusters often geographic units
  - e.g. schools, villages etc
- Usually used in interventional studies
  - E.g. assessing immunization coverage
- Advantages
  - sampling frame is not required in this case
  - Sampling study population scattered over a large area

Multistage Sampling

- Involves more than one sampling method
- Is therefore carried out in phases
- Does not require a initial sampling frame of whole population
- NEED TO KNOW SAMPLING FRAME OF CLUSTERS E.G. PROVINCES
- Require sampling frames of final clusters final clusters
- Applicable to community based studies e.g. interviewing people from different villages selected from different areas, selected from different districts, provinces

Nonprobability Sampling Methods

Nonprobability sampling: the sample elements are chosen from the population by nonrandom methods.

More likely to produce a biased sample than the random sampling.

This restricts the generalization of the study findings.

Most frequent reasons for use of nonprobability samples involve convenience and the desire to use available subjects.

Types of Sampling Methods

- Nonprobability Sampling Methods:
  - Convenience sampling.
  - Snowball sampling.
  - Quota sampling.
  - Purposive sampling.

- **Convenience sampling** (Accidental or incidental sampling):
  - People may or may not be typical of the population, no accurate way to determine their representativeness
  - Most frequently used in health research
- Advantages:
  - Saves time and money

- **Snowball sampling**: a method by which the study subjects assist in obtaining other potential subjects (networking)
- Useful in topics of research where the subjects are reluctant to make their identity known, Drug users, Aids patients, etc.

- **Quota sampling**
  - In quota sampling, the sample is selected by convenience (e.g. the first 50% of males and 50% of females)
  - A mean for securing potential subjects from these strata.
  - In a quota sampling variables of interest to the researcher (include subject attributes), such as age, gender, educational background are included in the sample

- **Purposive sampling** (handpicking, judgmental):
  - Subjects are chosen because they are typical or representative of the accessible population, or because they are experts (more knowledgeable) in the field of research topic.
  - Qualitative researchers use Purposive sampling

Variables

1. A variable is an object, characteristic, or property that can have different values.
2. A quantitative variable can be measured in some way.
3. A qualitative variable is characterized by its inability to be measured but it can be sorted into categories.