2024
SUMMARY

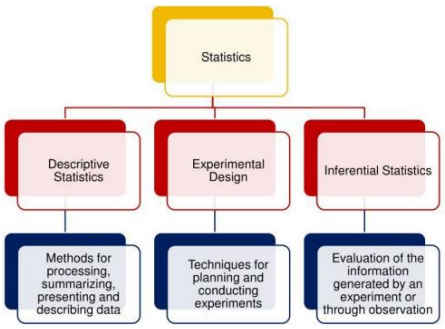
COMMUNITY MEDICINE

DR. AHMAD AL-QAWASMI

ARKAN ACADEMY

Biostatistics

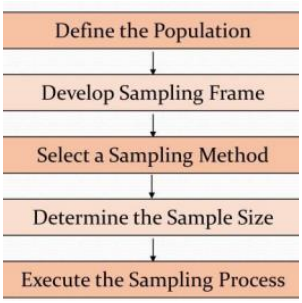
- Statistics involve data collection, data compilation (organization, classification, summarization), presentation (tables, diagrams, graphs), analysis, interpretation, reporting and dissemination
- Statistics is a *science* (a branch of mathematics) and it is an *art* (concerned with ways and means of handling data and presenting information and draw relevant conclusions)



- Functions of Statistics: Simplifies complex data and figures, presents facts in a definite form, facilitates comparison, helps in formulating and testing of hypothesis, interpretation and abstraction of evidence and it is vital for the formulation of evidence-based policies
- **Biostatics:** The application of the mathematical tools used in statistics to the fields of biological sciences and medicine (clinical ***research***, ***epidemiology***, ***health sciences***, ***public health***, ***environmental*** sciences and ***educational*** research)
  - It is important in research because it is used to describe and *summarize* information thereby reducing it to smaller, more meaningful sets of data and to identify associations, *relationships* or differences between the sets of observations
  - It is used to draw ***inferences*** about a body of data when ***only a part*** of the data is observed (generalize findings to the population based on observations from a sample of that population)
  - Statistics allows clinical researchers to draw reasonable and accurate inferences to make sound decisions in the presence of uncertainty
- Biostatistics is concerned with the interpretation of the ***quantitative data*** (which can be measured or counted) and the communication of information about the data
  - **Data:** ***Raw***, unorganized facts have to be processed (simple, ***random***, ***useless*** until processed)
  - **Information:** ***Processed***, ***organized*** data and presented into a ***useful*** context
- **Data collection:** Collecting data from all the relevant sources, via 2 main methods:
  - Primary data: Originated (collected) by the researcher to address the problem at hand
    - ✓ It requires higher costs and ***more time*** for collecting and analyzing data
    - ✓ *Observation*, *questionnaire*, *survey*, *census*, *interviews*
  - Secondary data: Data is collected by some different (not the researcher) with different purpose, so it is pre-existing data in various publications and databases or in other studies
    - ✓ *Census data*, *medical record*, *registries*, *published research*
    - ✓ **Census data:** Government collected data, provides comprehensive information about population such as demographics and employments
    - ✓ **Academic journals:** Peer-reviewed articles and studies, that contain data and findings relevant to various academic fields
- Population: collection or set of all values a variable may have (entire category under consideration)
- Sample: A **part** of a population (portion of the population that is available for analysis)

- Sampling: The selection of a number of study units/subjects from a defined population
  - Representative sample: A subset of a population that seeks to ***accurately reflect*** the characteristics of the larger group
  - Sampling bias: ***Excluding*** any subject without any scientific rational Or not based on inclusion and exclusion criteria (without providing assurance that the sample is representative)
    - ✓ Such as selecting the first 100 males in a mall to represent all males
    - ✓ To declare that a sample is representative, probability sampling must be used

- Element: The single member of the population
- Sampling frame: list of all elements of a population. Example: a list of all the registered medical students at the university of Jordan 2022-2024
- To whom are the results going to be generalized? ***Reference population***
- The group of people where 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***



- Type of sampling methods:

1. Probability Sampling Methods

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

- **Simple Random Sampling:** The simplest, make a **numbered list** of all units in the population then select a number of sampling units using the lottery method or a random number table
  - Advantages (Pros): ***Strong external validity*** (generalized to entire population), ***Relative speed*** and ***efficiency*** compared to the census
  - Disadvantages (Cons): ***Expensive***, ***Time consuming*** (slow), ***Not always possible***
  - Examples: Researchers who want to assess what Jordanians know about a particular health topic (list telephone numbers from a database of all cell phones, and randomly select numbers)

- **Systematic sampling:** Individuals are chosen from the sampling frame at **regular intervals**
  - Pros: ***External validity***, ***Relative speed***
  - Cons: ***Limited feasibility*** (not possible without a list of all members of the population)
  - Example: every 5th household, every 10th women attending ANC
  - Sampling fraction = Sample size / Study population
  - Interval Size = Study population / 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**
  - Random or systematic samples will have to be obtained from each stratum
  - Pros: ***External validity***, ***Representation*** of small and **minority** group
  - Example: Assessing the attitude of university students from different collages about smoking

- **Cluster sampling:** Selection of **clusters** instead of the selection of individuals
  - Clusters often *geographic units* (such as schools, villages)
  - Usually used in *interventional studies* such as assessing immunization coverage
  - Pros (Advantages): ***sampling frame is not required***, ***External validity***, ***Faster*** and ***more efficient*** and ***Cost reduction***
  - Cons: ***Not always possible***

- **Multistage sampling:** a version of ***cluster*** sampling (but more complex) where it contains ***2 or more stages*** in sample selection which makes primary data collection more manageable
  - It is not as effective as true random sampling being overly ***expensive*** and ***time-consuming***
  - Does **not** require initial sampling frame of whole population but **require sampling frames** of final clusters and it is applicable to community-based studies such as interviewing people from different villages selected from different districts, selected from different provinces

2. Nonprobability Sampling Methods

- The sample elements are chosen from the population by nonrandom methods
- More likely to produce a biased sample which restricts the generalization of the study findings

- **Convenience sampling (Accidental or incidental sampling)** based on who is **easy** for researchers to recruit for the study so people may or may not be representative of the population
  - Most frequently used in health research and it saves time and money

- **Snowball sampling:** begins when researchers **contact a few people** who meet a study's criteria then the researchers ask each person to recommend a few others
  - Built on each participant's social network
  - Pros: ***Effective way to study hard-to-reach groups*** and reach small or stigmatized groups where the subjects are reluctant to make their identity known (***Drug users, AIDS patients***)
  - Cons: hard to be generalized, slow

- **Quota sampling:** identify **subsets** of the population that are important to represent and then sample participants within each subset
  - Pros: ***Easy***, ***quick*** results, save ***time and money***
  - Cons: Not representative, cannot produce reliable results

- **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

ARKAN

ACADEMY

علم في كل مكان

Arkan academy

Arkanacademy

www.arkan-academy.com

+962 790408805