Biostatistics Block I
Lecture 1

Huining Kang
HuKang@salud.unm.edu

July 6, 2020

Introduction to Biostatistics

Content

- Introduction
- Some basic concepts
- Measurement and measurement scales
- The simple random sample
- The scientific method

Introduction

- Statistics and biostatistics
  - A science pertaining to the collection, analysis, interpretation or explanation, and presentation of data
  - Biostatistics refer to the application of statistics on biological and medical fields
- Course Objectives
  - To organize and summarize data (descriptive statistics)
  - To make decision about a population by examining a small part of data from the population (statistical inference)

Some basic concepts

- Data
  - Number from a measurement
  - Number from the process of counting
- Source of data
  - Routinely kept records
  - Survey
  - Experiments
  - External sources (published report, commercially available data bank, research literature, etc.)

Some basic concepts

- Variable
  - A characteristic that takes on different values in different persons, places, or things
- Quantitative variables
  - Can be measured in usual sense
  - Convey information regarding amount
- Qualitative Variables
  - Cannot be measured in quantitative sense, but can be categorized, such as gender, ethnicity.

Some basic concepts

- Random variable
  - Values of the variable obtained arise as a result of a chance factor, cannot be exactly predicated in advance
  - Observed values of a random variable is called observations or measurements
- Discrete random variable
  - Characterized by gaps and interruptions in the values that it can assume, such as those result from process of counting.
- Continuous random variable
  - Does not possess the gaps or interruptions characteristic, can assume any value within a specified relevant interval of real number.

Some basic concepts

- Population
  - The largest collection of entities for which we have an interest at a particular time
- Population of values
  - The largest collection of values of a random variable for which we have an interest at a particular time
  - May be finite or infinite
  - The members of population is usually denoted by

$$x = (x_1, x_2, \cdots, x_i, \cdots)$$

- Sample
  - A part of a population, usually denoted by

$$X = (X_1, X_2, \cdots, X_n)$$

  - The observation of a sample is usually denoted by

$$x = (x_1, x_2, \cdots, x_n)$$

Measurement and measurement scales

- Measurement
  - Assignment of numbers to objects or events according to a set of rules
  - Scale of measurement: nominal (lowest), ordinal, interval and ratio (highest)
- Nominal scale
  - Consists of “naming” observations or classifying them into various mutually exclusive and collectively exhaustive categories. Only comparisons that can be made between variables are equality and inequality.
  - Example: gender, race, religious affiliation.

Measurement and measurement scales

- Ordinal scale
  - The numbers assigned to objects represent the rank order (1^{st}, 2^{nd}, 3^{rd}, etc.). Comparison of greater or less can be made in addition to equality and inequality. But operation such as conventional addition and subtraction are still meaningless.
  - Example: Convalescing patients (unimproved, improved and much improved); children's intelligence (above average, average, below average); Risk of certain disease (High, medium and low)

Measurement and measurement scales

- Interval scale
  - Apart from having all the features of ordinal measurements, differences between arbitrary pairs of measurements can be made. Operations such as averaging and subtraction are therefore meaningful, but addition is not and a zero point on the scale is arbitrary; negative value can be used. Ratio between numbers on the scale is not meaningful, so multiplication and division cannot be carried out directly, but ratios of difference can be expressed.
  - Example: Temperature in Fahrenheit or Celsius scales, year date in many calendars.

Measurement and measurement scales

- Ratio scale
  - One in which the ratio between any two measurements is meaningful and there is a true zero. Operation such as multiplication and division become meaningful.
  - Example: height, weight, length, does of the drug

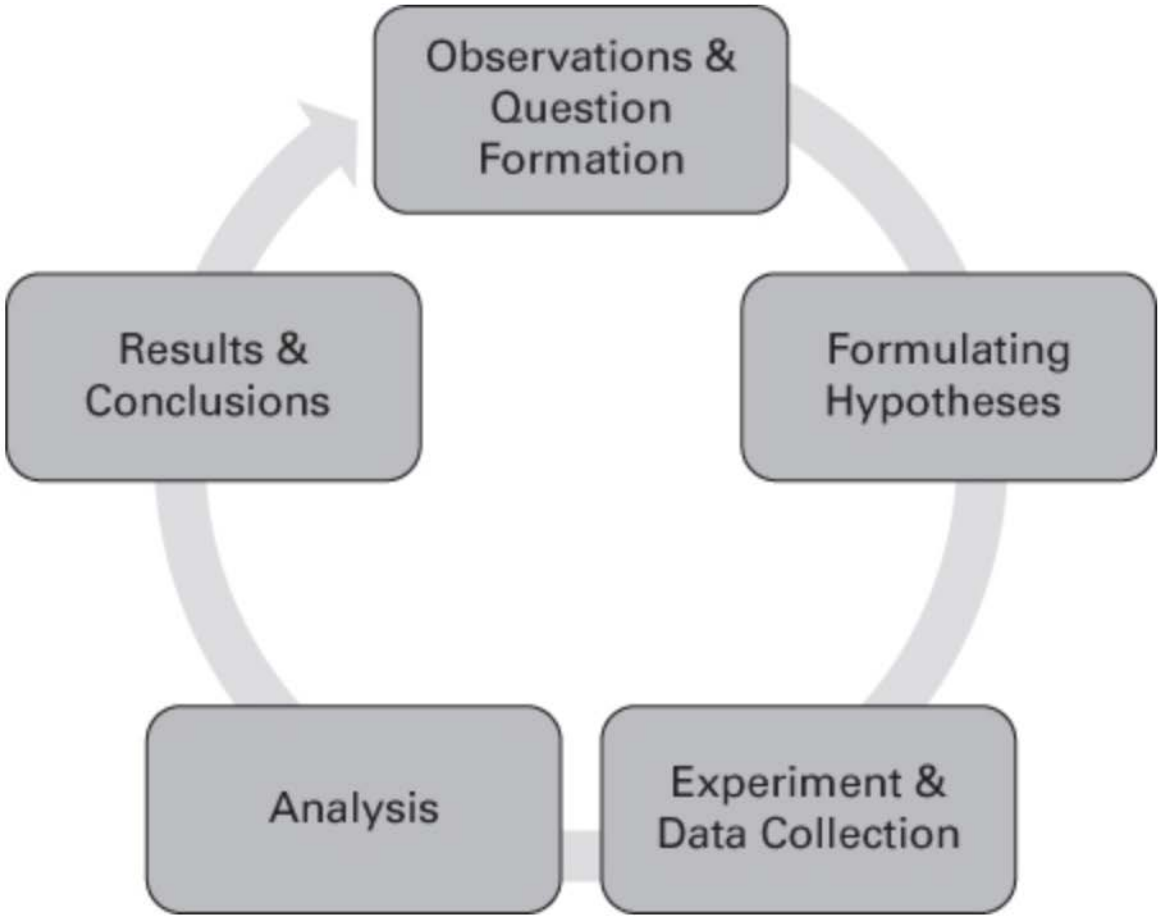
Simple random sample

- Statistical inference
  - Procedure by which we reach a conclusion about a population on the basis of the information contained in a sample that has been drawn from that population
- Simple random sample
  - A sample of size n is drawn from a population of size of N in such a way that every possible sample of size n has the same chance of being selected.
- Simple random sampling
  - Mechanics of drawing a simple random sample

Simple random sample

- Simple random sampling
  - Sample with replacement
  - Sample without replacement
  - In practice, sampling is always done without replacement
- Method of simple random sampling
  - Random number table
  - Generating random numbers in Stata
    - Stata solution for example1.4.1
      - After importing data, submit command
        sample 10, count
    - Refer do-file lecture01.do
    - Check help for command sample

The Scientific Method



Summary – what we have leanred

- Basic concepts
  - Statistics (descriptive statistics, statistical inference)
  - Data, variable (quantitative and qualitative) and random variable (discrete and continuous)
  - Population and sample
- Measurement and measurement scales
  - Measurement
  - Measurement scales (nominal, ordinal, interval and ratio)
- Simple random sample
  - Simple random sample
  - Table of random number, random number generator in software
- The scientific method