

ISCB20.05–Introduction to Biostatistics

Md. Jubayer Hossain

<https://jhossain.me/>

December 25, 2020

Lead Organizer, Introduction to Scientific Computing for Biologists

Founder, Health Data Research Organization



Section-1.2: Data Research Methods

Overview

- What is Research? Research Process
- Types of Research
- Sampling and Sampling Methods
- Data Collection Methods
- Types of Statistical Methods
- Steps in a Research Project

What is Research?

Definition-1

Research is "creative and systematic work undertaken to increase the stock of knowledge". It involves the collection, organization, and analysis of information to increase understanding of a topic or issue. A research project may be an expansion on past work in the field. (<https://en.wikipedia.org/wiki/Research>)

Definition-2

Research is defined as the creation of new knowledge and/or the use of existing knowledge in a new and creative way so as to generate new concepts, methodologies and understandings. This could include synthesis and analysis of previous research to the extent that it leads to new and creative outcomes. (<https://www.westernsydney.edu.au/research/researchers/>)

Definition-3

Simply, Research is an organized and systematic way of finding answers to questions.

Research Process



Source: <https://www.scribbr.com/category/research-process/>

Types of Research–1

Application

- **Fundamental Research:** Basic research aims to develop knowledge, theories and predictions
- **Applied Research:** Applied research aims to develop techniques, products and procedures.

Types of Research–2

Objective

- **Exploratory:** Exploratory research aims to explore the main aspects of under-researched problem.
- **Explanatory:** Explanatory research aims to explain the causes and consequences of a well-defined problem.
- **Descriptive:** Descriptive research aims to accurately and systematically describe a population, situation or phenomenon.
- **Correlational:** Correlational research is a type of research method that involves observing two variables in order to establish a statistically corresponding relationship between them.

Types of Research–3

Inquiry Mode

- **Quantitative:** Quantitative research is expressed in numbers and graphs.
- **Qualitative:** Qualitative research is expressed in words.

Fundamental or Basic Research

- Investigation on basic principles and reasons for occurrences of a particular event or process or phenomenon.
- It is also called theoretical research.
- Some natural phenomenon or relating to pure science
- Basic researches sometimes may not lead to immediate use or application.
- It helps to build new frontiers of knowledge
- Outcomes of basic research form the basis for many applied research.

Examples

- An investigation into the secondary symptoms of the Human Papilloma Virus (HPV).
- An investigation into the symptoms of diarrhea.

Applied Research

- Solve certain problems employing well known and accepted theories and principles.
- Most of the experimental research, case studies and inter-disciplinary research essentially applied research.
- Studies individuals or specific cases without the objective to generalize.
- Tries to say how things can be changed
- Tries to correct facts which are problematic.

Examples

- An investigation to determine the healing properties of mushrooms.
- An investigation to determine the side effects of alcohol consumption.

Exploratory Research

- Exploratory research aims to explore the main aspects of under-researched problem.
- Exploratory type of research is usually conducted to have a better understanding of the existing problem, but usually doesn't lead to a conclusive result.

Explanatory Research

- Explanatory research aims to explain the causes and consequences of a well-defined problem.
- Explanatory research is also known as causal research
- Explanatory research is conducted in order to identify the extent and nature of cause-and-effect relationships.
- Causal research can be conducted in order to assess impacts of specific changes on existing norms, various processes etc.
- Causal studies focus on an analysis of a situation or a specific problem to explain the patterns of relationships between variables.

Descriptive Research

- Descriptive research aims to accurately and systematically describe a population, situation or phenomenon. It can answer what, where, when and how questions, but not why questions.
- A descriptive research design can use a wide variety of research methods to investigate one or more variables.

Correlational Research

- Correlational research is a type of research method that involves observing two variables in order to establish a statistically corresponding relationship between them.
- The aim of correlational research is to identify variables that have some sort of relationship to the extent that a change in one creates some change in the other.

Quantitative Research

- Focuses on testing theories and hypotheses
- Analyzed through math and statistical analysis
- Mainly expressed in numbers, graphs and tables
- Requires many respondents
- Closed (multiple choice) questions
- Key terms: testing, measurement, objectivity, replicability

Qualitative Research

- Focuses on exploring ideas and formulating a theory or hypothesis
- Analyzed by summarizing, categorizing and interpreting
- Mainly expressed in words
- Requires few respondents
- Open-ended questions
- Key terms: understanding, context, complexity, subjectivity

Types of Study Design

- **Observational:** A study where a researcher records or observes the observations or measurements without manipulating any variables. These studies show that there may be a relationship but not necessarily a cause and effect relationship.
- **Experimental:** A study that involves some random assignment* of a treatment; researchers can draw cause and effect (or causal) conclusions. An experimental study may also be called a scientific study or an experiment.

(Source: <https://online.stat.psu.edu/stat800/lesson/3/3.4>)

Problem 1: Drinking Tea before Bedtime

A study took random sample of adults and asked them about their bedtime habits. The data showed that people who drank a cup of tea before bedtime were more likely to go to sleep earlier than those who didn't drink tea.

- Observational(Correct Answer)
- Experimental

(Source: <https://www.khanacademy.org/math/ap-statistics/>)

Problem 2: Drinking Tea before Bedtime

Another study took a group of adults and randomly divided them into two groups. One group was told to drink tea every night for a week, while the other group was told not to drink tea that week. Researchers then compared when each group fell asleep.

- Observational
- Experimental(Correct Answer)

(Source: <https://www.khanacademy.org/math/ap-statistics/>)

Cross-sectional Study

- A cross-sectional study is a type of research design in which you collect data from many different individuals at a single point in time. In cross-sectional research, you observe variables without influencing them.
- Researchers in economics, psychology, medicine, epidemiology, and the other social sciences all make use of cross-sectional studies in their work.

When to Use?

- When you want to examine the prevalence of some outcome at a certain moment in time, a cross-sectional study is the best choice.

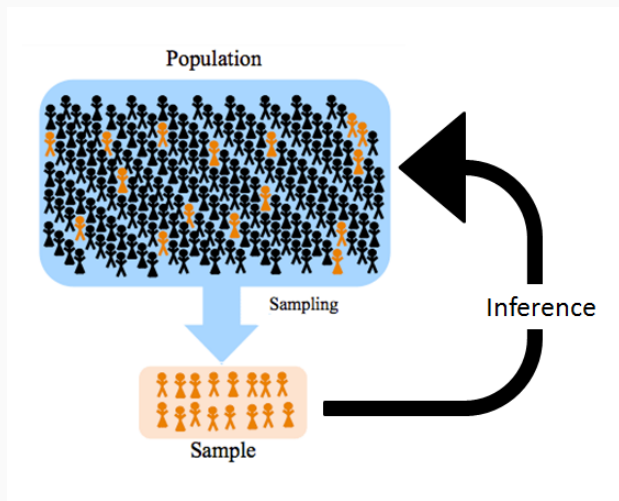
Case Study

- A case study is a detailed study of a specific subject, such as a person, group, place, event, organization, or phenomenon. Case studies are commonly used in social, educational, clinical, and business research.
- A case study research design usually involves qualitative methods, but quantitative methods are sometimes also used.
- Case studies are good for describing, comparing, evaluating and understanding different aspects of a research problem.

When to Use?

- A case study is an appropriate research design when you want to gain concrete, contextual, in-depth knowledge about a specific real-world subject. It allows you to explore the key characteristics, meanings, and implications of the case.

Sampling



Source: <https://online.stat.psu.edu/stat500/>

Types of Sampling Methods

To draw valid conclusions from your results, you have to carefully decide how you will select a sample that is representative of the group as a whole. There are two types of sampling methods

- Probability Sampling
- Non-Probability Sampling

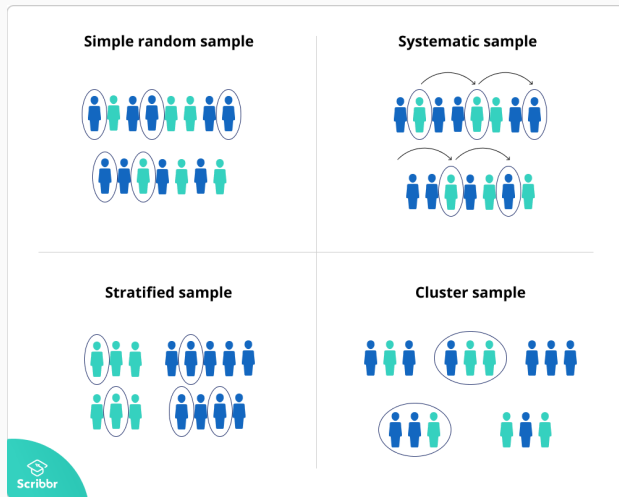
Probability Sampling Methods

- Probability sampling involves random selection, allowing you to make statistical inferences about the whole group.
- Probability sampling means that every member of the population has a chance of being selected.
- It is mainly used in quantitative research.
- If you want to produce results that are representative of the whole population, you need to use a probability sampling technique.

Non-Probability Sampling Methods

- Non-probability sampling involves non-random selection based on convenience or other criteria, allowing you to easily collect initial data.

Types of Probability Sampling Methods



Source: <https://www.scribbr.com/methodology/sampling-methods/>

Simple Random Sampling(SRS)

- In a simple random sample, every member of the population has an equal chance of being selected. Your sampling frame should include the whole population.
- To conduct this type of sampling, you can use tools like random number generators or other techniques that are based entirely on chance.

Systematic Sampling

- Systematic sampling is similar to simple random sampling, but it is usually slightly easier to conduct.
- Every member of the population is listed with a number, but instead of randomly generating numbers, individuals are chosen at regular intervals.

Stratified Sampling

- Stratified sampling involves dividing the population into subpopulations that may differ in important ways. It allows you draw more precise conclusions by ensuring that every subgroup is properly represented in the sample.
- To use this sampling method, you divide the population into subgroups (called strata) based on the relevant characteristic (e.g. gender, age range, income bracket, job role).
- Based on the overall proportions of the population, you calculate how many people should be sampled from each subgroup. Then you use random or systematic sampling to select a sample from each subgroup.

Cluster Sampling

- Cluster sampling also involves dividing the population into subgroups, but each subgroup should have similar characteristics to the whole sample. Instead of sampling individuals from each subgroup, you randomly select entire subgroups.
- If it is practically possible, you might include every individual from each sampled cluster. If the clusters themselves are large, you can also sample individuals from within each cluster using one of the techniques above.
- This method is good for dealing with large and dispersed populations, but there is more risk of error in the sample, as there could be substantial differences between clusters. It's difficult to guarantee that the sampled clusters are really representative of the whole population.

Data Collection Methods

- **Survey:** Survey research allows you to gather large volumes of data that can be analyzed for frequencies, averages and patterns.
- **Observations:** Observations allow you to gather data on behaviours and phenomena without having to rely on the honesty and accuracy of respondents. This method is often used by psychological, social and market researchers to understand how people act in real-life situations.
- **Case Study:** A case study can be used to describe the characteristics of a specific subject (such as a person, group, event or organization). Instead of gathering a large volume of data to identify patterns across time or location, case studies gather detailed data to identify the characteristics of a narrowly defined subject.
- Interviews
- Experiment

Types of Statistical Methods

- **Descriptive Statistics:** Identify important elements in a dataset.
- **Inferential Statistics:** Explain those elements via relationships with other elements.

Descriptive Statistics

Descriptive Statistics divide into 3 categories.

- **Univariate Analysis:** summarize only one variable at a time.
- **Bivariate Analysis:** compare two variables.
- **Multivariate Analysis:** compare more than two variables.

Characteristics of Descriptive Statistical Methods

- Descriptive statistical methods provide an exploratory assessment of the data from a study
- Exploratory data analysis techniques
- Organization and summarization of data
 - Tables
 - Graphs
 - Summary measures

Inferential Statistics

Inferential Statistics divide into 2 categories.

- **Hypothesis Testing:** A hypothesis is a statement that can be tested by scientific research.
- **Model Fitting:** Model fitting is a measure of how well a statistical learning model generalizes to similar data to that on which it was trained.

Characteristics of Inferential Statistical Methods

- Assess the strength of evidence for/against a hypothesis.
- Inferential statistical methods provide a confirmatory data analysis.
- Generalize conclusions from data from part of a group (sample) to the whole group (population)
- Assess the strength of the evidence
- Make comparisons.
- Make predictions.
- Ask more questions; suggest future research.

Steps in a Research Project

(1) Questions → Why? → Literature Review ← Iterative Process

(2) Study Design → Study Subject → Data Collection → Dummy Table/Shell Table

(3) Statistical Analysis → Manuscript → Publication

Planning/ Study Design

- Primary question(s) of interest
 - Quantifying information about a single group?
 - Comparing multiple groups?
- Sample Size
 - How many subjects needed total?
 - How many in each of the groups to be compared?
- Selecting study participants
 - Randomly chosen from “master list?”
 - Selected from a pool of interested persons?
- Sampling Methods
 - Probabilistic methods?
 - Non-probabilistic methods?

Data Collection

First, decide how you will collect data. Your methods depend on what type of data you need to answer your research question.

- Qualitative or Quantitative?
- Primary or Secondary?
- Descriptive or Experimental?

Data Analysis

After collecting data, decide how you will analyze the data.

- For quantitative data, you can use statistical analysis methods to test relationships between variables.
- For qualitative data, you can use methods such as thematic analysis to interpret patterns and meanings in the data.

Interpretation

What do the results mean in terms of practice, the program, the population etc.?

Reporting/Manuscript Writing

- What summary measures will best convey the “main messages” in the data about the primary (and secondary) research questions of interest
- How to convey/ rectify uncertainty in estimates based on the data

Publish Your Research . . .

References

- <https://www.formpl.us/blog/research/home>
- <https://www.scribbr.com/methodology/>
- <https://www.questionpro.com/blog/>
- <https://www.scribbr.com/category/research-process/>
- <https://ocw.jhsph.edu/index.cfm/go/viewCourse/course/FundEpi/coursePage/index/>

Thank You

