

# Fundamentals of Statistical Estimation

M. O.

October 29, 2025

# Outline

- 1 Population and Sample
  - Population
  - Sample
  - Surveys and Characterizing the Population
- 2 Types of Statistical Studies
  - Observational and Experimental Studies
  - Design of Experimental Studies
  - Fisher's Three Principles
  - Design of Observational Studies
- 3 Sample Surveys and Sampling Methods
  - Simple Random Sampling
  - Systematic Sampling
  - Stratified Random Sampling
- 4 aaa

In inferential statistics, distinguishing between the **population** under analysis and the **sample** drawn from it is essential.

The population is defined as the entire set of objects to which the results of the analysis are to be applied.

In the context of statistical estimation, the population typically represents 'that which we wish to know but cannot fully observe,' by definition of the problem setting.

# Sample

The sample, conversely, is originally a subset extracted from the population. By an abuse of terminology, the term often refers to a subset of the characteristic values (e.g., quantities like height or weight) extracted from the population. When the sample is a subset of characteristic values, quantitative analysis and inference become possible, as it can be expressed as  $\{x_1, x_2, \dots, x_n\} \in \mathbb{R}^n$ , for instance.

Some problem settings may permit a 'census,' or a complete survey of the entire population.

Generally, however, populations are too large for a complete survey; therefore, characteristics of the population are inferred by analyzing a sample. This method of investigation is called a 'sample survey,' and the estimations and hypothesis tests based on it constitute 'statistical inference.' (Both 'estimation' and 'testing' are forms of statistical 'inference.')

# Surveys and Characterizing the Population

While knowledge about the distribution of the population is sometimes sought, often the objective is to determine the value of a constant that characterizes that population.

A constant that characterizes the population is called a **parameter**.

Examples of parameters include the population mean, population variance, population standard deviation, and population proportion.

In contrast, a 'function' constructed from the sample is called a **statistic**.

Examples of statistics include the sample mean, sample variance, sample standard deviation, and sample proportion.

Inferential statistics aims to ascertain the value of a parameter based on the value of a statistic calculated from observed data.

# Types of Statistical Studies

The role of statistics is not limited to providing methods for understanding population characteristics by estimating the mean and variance.

Statistics also plays a role in offering methods for drawing a sample correctly from the population.

The 'Simple Random Sampling' method (discussed later) is a well-known example, but many other experimental designs exist.

Research aimed at understanding population characteristics is broadly divided into two types: **experimental studies** and **observational studies**.

An experimental study involves intervention in the subjects of the research.

An observational study is one where intervention in the subjects is not possible.

# Observational and Experimental Studies

Statistics offers methods not only for inferring population characteristics by estimating means and variances but also for ensuring the correct extraction of a sample from the population.

Designing experiments in laboratories or factories, animal experiments for new drug development, and clinical trials are also part of the statistician's role.

Research aimed at understanding population characteristics includes both experimental and observational studies.

**Experimental Studies:** Intervention in the research subjects is possible.

**Observational Studies:** No intervention in the research subjects occurs.

In both cases, the resulting sample is only a part of the population.

The research process must be designed so that the sample represents the population's characteristics without bias.

# Design of Experimental Studies

Despite the name, not all experimental studies take place in a laboratory. An experiment is defined as research where the conditions can be set by the researcher.

Examples include clinical trials to evaluate the efficacy of a new drug or agricultural trials to determine the effect of a fertilizer.

Consider the case of evaluating the efficacy of a new drug. Patients are divided into a group receiving the new drug (the **experimental group**) and a group receiving a control drug (the **control group**), and the difference between the two is measured.

Fisher's Three Principles, described next, are crucial in experimental studies.



# Fisher's Three Principles

Fisher's Three Principles:

**Randomization**

**Replication**

**Local Control**

# Fisher's Three Principles (Randomization)

Randomization is also known as random assignment.

When comparing a new drug (test drug) with an existing one (control drug) in a clinical trial, it is vital to randomly decide which patient receives which drug (known as **random allocation**).

The experimental conditions imposed on the experimental units (in this case, the subjects) when comparing the test drug and the control drug are called the 'treatment.'

It is desirable to keep conditions other than the treatment as uniform as possible. While anticipated biases can be dealt with, if uniformity cannot be achieved for unanticipated biases, random assignment is necessary.

# Fisher's Three Principles (Replication)

Data inherently includes variability.

Even experiments conducted under identical conditions will yield variations in data.

Since the goal is to see if the treatment effect is observable beyond this variability, the magnitude of the variation must be estimated.

Thus, repeating the experiment is necessary.

In clinical trials involving humans, individual differences necessitate collecting data from many subjects, which is also considered replication.

Estimating the required number of repetitions during the planning stage is one of the significant tasks for a statistical analyst.

# Fisher's Three Principles (Local Control)

As mentioned before, to compare the treatment effect with the magnitude of variability, it is desirable to minimize variability other than that caused by the treatment.

Therefore, experiments are divided into several **blocks** with uniform experimental conditions.

This blocking is called local control.

For example, an experiment comparing a new drug administered only to males in their 60s with a control drug administered only to females in their 20s is meaningless in terms of treatment effect.

It is desirable to design the experiment by using age and sex as blocks.

The setting of blocks should aim for maximum uniformity **within** the same block and maximum difference **between** different blocks.

# Design of Observational Studies

While experimental studies are indispensable for demonstrating treatment effects, many problems do not permit experimentation.

Consider a study investigating the effect of smoking on health: subjects cannot be compelled to smoke to study its effects.

In such cases, researchers must rely on observational studies.

Unlike experimental studies, observational studies lack the principle of randomization among Fisher's Three Principles.

The key difference is that subjects self-select the treatment.

Because subjects consciously choose their treatment, there is a potential for bias in treatment selection due to the subjects' characteristics.

Caution is needed when interpreting treatment effects from observational studies.

# Simple Random Sampling

aaa

## Population and Sample

Population

Sample

Surveys and Characterizing  
the Population

## Types of Statistical Studies

Observational and  
Experimental Studies

Design of Experimental  
Studies

Fisher's Three Principles

Design of Observational  
Studies

## Sample Surveys and Sampling Methods

Simple Random Sampling

Systematic Sampling

Stratified Random Sampling

aaa

# Systematic Sampling

aaa

## Population and Sample

Population

Sample

Surveys and Characterizing  
the Population

## Types of Statistical Studies

Observational and  
Experimental Studies

Design of Experimental  
Studies

Fisher's Three Principles

Design of Observational  
Studies

## Sample Surveys and Sampling Methods

Simple Random Sampling

**Systematic Sampling**

Stratified Random Sampling

aaa

# Stratified Random Sampling

aaa

## Population and Sample

Population

Sample

Surveys and Characterizing  
the Population

## Types of Statistical Studies

Observational and  
Experimental Studies

Design of Experimental  
Studies

Fisher's Three Principles

Design of Observational  
Studies

## Sample Surveys and Sampling Methods

Simple Random Sampling

Systematic Sampling

**Stratified Random Sampling**

aaa



aaa

aaa

## Population and Sample

Population

Sample

Surveys and Characterizing  
the Population

## Types of Statistical Studies

Observational and  
Experimental Studies

Design of Experimental  
Studies

Fisher's Three Principles

Design of Observational  
Studies

## Sample Surveys and Sampling Methods

Simple Random Sampling

Systematic Sampling

Stratified Random Sampling

aaa

# References I

 What is a p-value anyway ? - Vickers, Andrew J.

 Statistics is the most powerful discipline (Practical Edition) - Kei Nishiuchi