# Course Introduction

**Course name**: Statistics & Probability (MAS291)

**Textbook**: *Applied Statistics and Probability for Engineers*, $5^{th}$ edition.

**Topics covered**:

- Chapter 1: The Role of Statistics in Engineering
- Chapter 2: Probability
- Chapter 3: Discrete Random Variables and Probability Distribution
- Chapter 4: Continuous Random Variables and Probability Distribution
- Chapter 6: Descriptive Statistics
- Chapter 7: Sampling Distributions and Point Estimation of Paramaters
- Chapter 8: Statistical Intervals for a Single Sample
- Chapter 9: Test of Hypotheses for a Single Sample
- Chapter 10: Statistical Inference for Two Samples
- Chapter 11: Simple Linear Regression and Correlation

Statistics & Probability

# Chapter 1. THE ROLE OF STATISTICS IN ENGINEERING

FPT University

**Department of Mathematics**

*Quy Nhon*, 2023

# Table of Contents


1. The Role of Statistics in Engineering
2. The Engineering Method and Statistical Thinking
3. Collecting Engineering Data
4. Mechanistic and Empirical Models
5. Probability and Probability Models

# Table of Contents


1. The Role of Statistics in Engineering
2. The Engineering Method and Statistical Thinking
3. Collecting Engineering Data
4. Mechanistic and Empirical Models
5. Probability and Probability Models

## Why is Statistics?

- Statistics allows you to understand a subject much more deeply.
- Statistics helps us to make discoveries in science, make decisions based on data, and make predictions.
- Statisticians and statistical methods are important parts of pharmaceutical industry, social scientists, business practice,. . .

# Table of Contents


1. The Role of Statistics in Engineering
2. The Engineering Method and Statistical Thinking
3. Collecting Engineering Data
4. Mechanistic and Empirical Models
5. Probability and Probability Models

# The Engineering Method

**Engineering method**:

1. Develop a clear and concise description of the problem.
2. Identify, at least tentatively, the important factors that affect this problem or that may play a role in its solution.
3. Propose a model for the problem, using scientific or engineering knowledge of the phenomenon being studied. State any limitations or assumptions of the model.
4. Conduct appropriate experiments and collect data to test or validate the tentative model or conclusions made in steps 2 and 3.
5. Refine the model on the basis of the observed data.
6. Manipulate the model to assist in developing a solution to the problem.
7. Conduct an appropriate experiment to confirm that the proposed solution to the problem is both effective and efficient.
8. Draw conclusions or make recommendations based on the problem solution.

# What is Statistics?

Statistics is the science of collecting, organizing, analyzing, and interpreting **DATA** in order to make decisions.

- **Descriptive Statistics**: Involves organizing, summarizing, and displaying data. $\underline{Example}$; Tables, charts,. . .
- **Inferential Statistics**: Involves using sample data to draw conclusions about a population.

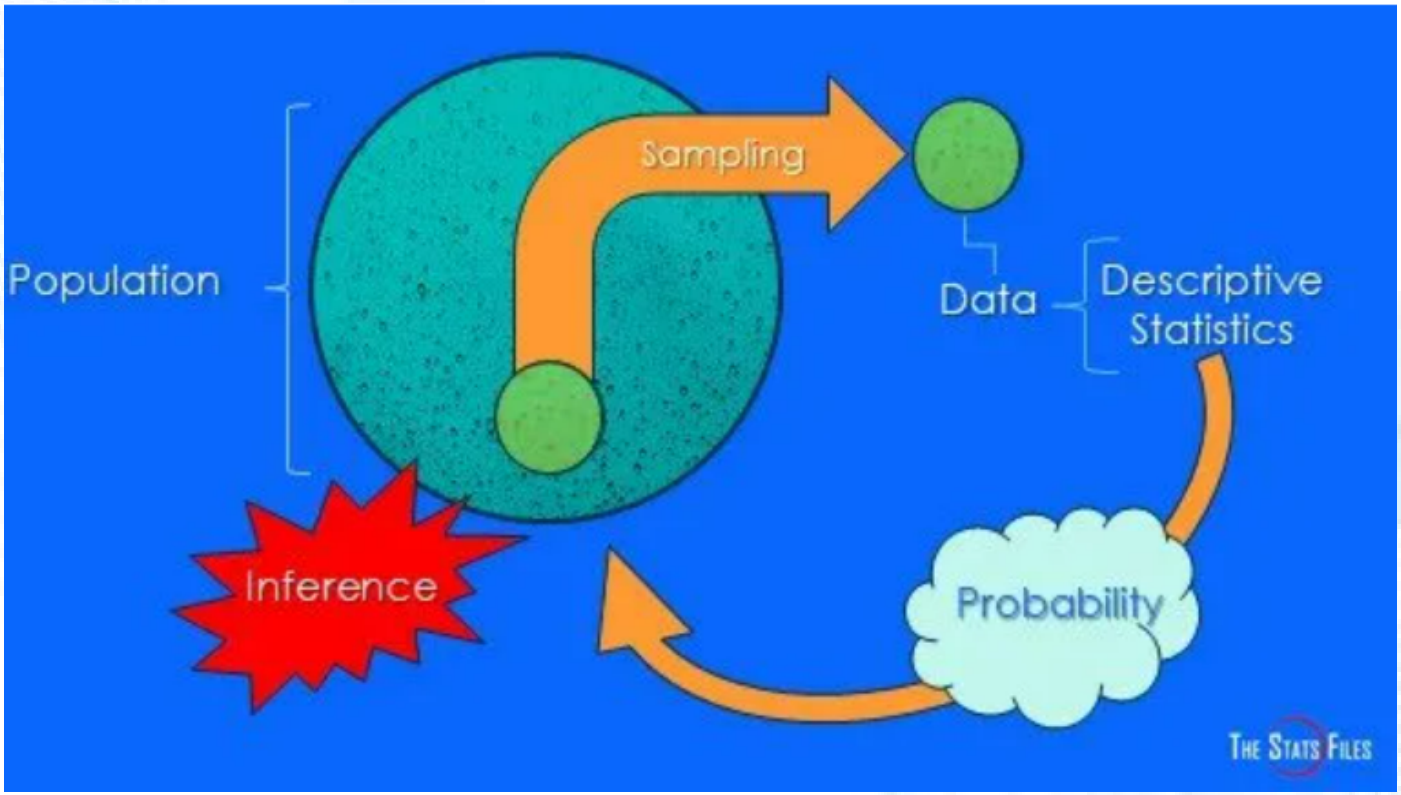


**Figure**. Big picture of Statistics.

# Statistical Thinking

**Statistical thinking** can give us a useful way to incorporate this variability into our decision-making processes.

**Statistics** provides a framework for describing the variability and for learning about which potential sources of variability are the most important or which have the greatest impact on the considering problem.

**Statistics concepts:**

- Population: the complete collection of all individuals to be studied.
- Sample: Sub-collection of members selected from a population.
- Data: consist of information coming from observations, counts, measurements, or responses.
- Parameter: a numerical measurement describing some characteristic of a population.
- Statistic: a numerical measurement describing some characteristic of a sample.
- Random variable: a random variable encompasses all the possible values in a sample space.

**Question 1**. A survey will be given to $100$ students randomly selected from the freshmen class at LQD High School. What is the population?

1. The $100$ selected students.
2. All freshmen at LQD High School.
3. All students at LQD High School

**Question 2**. A survey will be given to $100$ students randomly selected from the freshmen class at LQD High School. What is the sample?

1. The $100$ selected students.
2. All freshmen at LQD High School.
3. All students at LQD High School

**Question 3**. Fifty bottles of water were randomly selected from a large collection of bottles in a company's warehouse. These fifty bottles are referred to as the . . . . . .

1. population.
2. sample.

**Question 4**. Fifty bottles of water were randomly selected from a large collection of bottles in a company's warehouse. The large collection of bottles is referred to as the . . .

1. population.
2. sample.

**Question 5**. A survey of $2000$ American households found that $33\%$ of the respondents own a computer. Is this value a parameter or a statistic?

**Question 6**. The average salary of all automotive workers is $\$42,000$. Is this value a parameter or a statistic?

# Types of Data

1. **Qualitative data**: describes qualities or characteristics.

   $\rightarrow$ This data type is non-numerical in nature.

   **Example**:

   - The cake is orange, blue, and black in color (qualitative).
   - Females have brown, black, blonde, and red hair (qualitative).
   - Major, place of birth,...

2. **Quantitative data**: is any quantifiable information that can be used for mathematical calculation or statistical analysis.

   $\rightarrow$ This data is all about numbers.

   **Example**:

   - There are four cakes and three muffins kept in the basket (quantitative).
   - One glass of fizzy drink has 97.5 calories (quantitative).
   - Age, temperature,...

# Table of Contents


1. The Role of Statistics in Engineering
2. The Engineering Method and Statistical Thinking
3. Collecting Engineering Data
4. Mechanistic and Empirical Models
5. Probability and Probability Models

# Collecting Engineering Data

## Basic Principles

In the engineering environment, the data are almost always a sample that has been selected from the population. There are three basic methods of collecting data including

1. A retrospective study using historical data.
2. An observational study.
3. A designed experiment.

1. **Retrospective study**: A retrospective study would use either all or a sample of the historical process data archived over some period of time.
2. **Observation study**: A researcher observes and measures characteristics of interest of part of a population.
3. **Designed experiments**: A treatment is applied to part of a population and responses are observed.
4. **Observing processes over time**: It is usually very helpful to plot the data versus time in a time series plot.

FUQN

# Table of Contents


1. The Role of Statistics in Engineering
2. The Engineering Method and Statistical Thinking
3. Collecting Engineering Data
4. Mechanistic and Empirical Models
5. Probability and Probability Models

# Mechanistic and Empirical Models

Models play an important role in the analysis of nearly all engineering problems.

1. **Mechanistic model**: built from our underlying knowledge.

   **Example**: Ohm's law: $Current = \dfrac{Voltage}{Resistance}$ or $I = \dfrac{U}{R}$.

2. **Empirical model**: uses our engineering and scientific knowledge of the phenomenon:

$$Response = Deterministic\ function + Random\ error$$

   **Example**: $I = \dfrac{U}{R} + \varepsilon$.

**Remark**

$\varepsilon$ is a term added to the model to account for the fact that the observed values of current flow do not perfectly conform to the mechanistic model.

# Table of Contents


1. The Role of Statistics in Engineering
2. The Engineering Method and Statistical Thinking
3. Collecting Engineering Data
4. Mechanistic and Empirical Models
5. Probability and Probability Models

# Probability and Probability Models

1. **A probability** is a numerical value assigned to a given event.
2. **Probability models** help quantify the risks involved in statistical inference, that is, the risks involved in decisions made every day.

# Thank you!