

# CM 2110 Calculus and Statistical Distributions

Dr. Priyanga D. Talagala

2020-05-29

---

# Contents

|   |               |
|---|---------------|
| <b>Course Syllabus</b>                                | <b>5</b>      |
| Pre-requisites . . . . .                              | 5             |
| Learning Outcomes . . . . .                           | 5             |
| Outline Syllabus . . . . .                            | 5             |
| Method of Assessment . . . . .                        | 6             |
| Recommended Texts . . . . .                           | 6             |
| Lecturer . . . . .                                    | 6             |
| Schedule . . . . .                                    | 6             |
| <br><b>1 Statistical Distributions</b>                | <br><b>1</b>  |
| Recap: CM 1110-Probability . . . . .                  | 1             |
| 1.1 Random Variable . . . . .                         | 3             |
| 1.2 Probability Mass Function . . . . .               | 7             |
| 1.3 Probability density function . . . . .            | 11            |
| 1.4 Cumulative distribution function . . . . .        | 13            |
| 1.5 Descriptive properties of distributions . . . . . | 13            |
| 1.6 Models for discrete distributions . . . . .       | 13            |
| 1.7 Models for continuous distributions . . . . .     | 13            |
| <br><b>2 Estimations</b>                              | <br><b>15</b> |
| 2.1 Point Estimation . . . . .                        | 15            |
| 2.2 Interval Estimation . . . . .                     | 15            |

---

|   |           |
|---|-----------|
| <b>3 Hypothesis Testing</b>   | <b>17</b> |
| 3.1 Null and alternative hypotheses . . . . .                       | 17        |
| 3.2 Errors in testing hypotheses-type I and type II error . . . . . | 17        |
| 3.3 Significance level, size, power of a test . . . . .             | 17        |
| 3.4 Formulation of hypotheses . . . . .                             | 17        |
| 3.5 Methods of testing hypotheses . . . . .                         | 17        |
| <b>4 Design of Experiments</b>                                      | <b>19</b> |
| 4.1 Introduction to experimental design . . . . .                   | 19        |
| 4.2 Basic principles of experimental design . . . . .               | 19        |
| 4.3 Completely randomized design . . . . .                          | 19        |

# Course Syllabus

## Pre-requisites

CM 1110

### **Remark:**

*This course module contains two main sections: (1) mathematics and (2) statistics. This syllabus is designed for the statistics section. Lectures for mathematics section and statistics section are conducted by two lecturers as two separate sub modules (1.5 hour lectures/Week). End Semester Examination is conducted as a single examination.*

## Learning Outcomes

On successful completion of this module, students will be able to plan more carefully the design of experiment in advance which provide evidence for or against theories of cause and effect and make inferences about population characteristics based on sample information and thereby solve data analysis problems in different application domains. (R(<https://cran.r-project.org/>) and RStudio are also freely available to install on your own computer). Get the Open Source Edition of RStudio Desktop. RStudio allows you to run R in a more user-friendly environment.

## Outline Syllabus

- Functions of Several Variables
- Linear Algebra
- Coordinate Systems & Vectors
- Differential Equations
- **Statistical Distributions**
- **Estimation**

- Hypothesis Testing
- Design of Experiments

## Method of Assessment

- Mid-semester examination
- End-semester examination

## Recommended Texts

- Casella, G., & Berger, R. L. (2002). Statistical inference (Vol. 2, pp. 337-472). Pacific Grove, CA: Duxbury.
- Mood, A.M., Graybill, F.A. and Boes, D.C. (2007): Introduction to the Theory of Statistics, 3rd Edn. (Reprint). Tata McGraw-Hill Pub. Co. Ltd.
- Montgomery, D. C. (2017). Design and analysis of experiments. John wiley & sons.

## Lecturer

Dr. Priyanga D. Talagala

## Schedule

Lectures:

- Friday [9.15 am - 10.45 am]

Tutorial:

- Friday [11.00 am - 12.30 pm]

Consultation time:

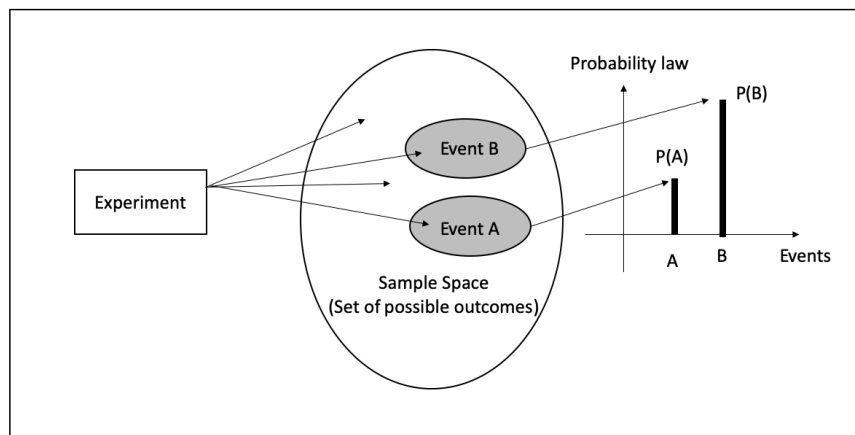
- Friday [8.15 am to 9.00 am]

# Chapter 1

## Statistical Distributions

### Recap: CM 1110-Probability

#### Axioms of probability



- **Probability** of an event quantifies the **uncertainty**, randomness, or the possibility of occurrence the event.
- The probability of event  $E$  is usually denoted by  $P(E)$ .
- Mathematically, the function  $P(\cdot)$  is a set function defined from sample space  $(\Omega)$  to  $[0, 1]$  interval, satisfying the following properties.
- These are called the '**axioms of probability**'.

- **Axiom 1:** For any event  $A$ ,  $P(A) \geq 0$
- **Axiom 2:**  $P(\Omega) = 1$
- **Axiom 3:**
  - (a) If  $A_1, A_2, \dots, A_k$  is a finite collection of mutually exclusive events, then

$$P(A_1 \cup A_2 \cup \dots \cup A_k) = \sum_{i=1}^k P(A_i)$$

- (b) If  $A_1, A_2, \dots$  is an infinite collection of mutually exclusive events, then

$$P(A_1 \cup A_2 \cup \dots) = \sum_{i=1}^{\infty} P(A_i)$$

**NOTE**

- Axioms 1 and 2 imply that for any event  $E$ ,  $0 \leq P(E) \leq 1$ .
- $P(E) = 1 \iff$  the event  $E$  is certain to occur.
- $P(E) = 0 \iff$  the event  $E$  cannot occur.

**Methods for determining Probability**

- There are several ways for determining the probability of events.
- Usually we use the following methods to obtain the probability of events.
  - Classical method
  - Relative frequency method (Empirical approach)
  - Subjective method
  - **Using probability models**

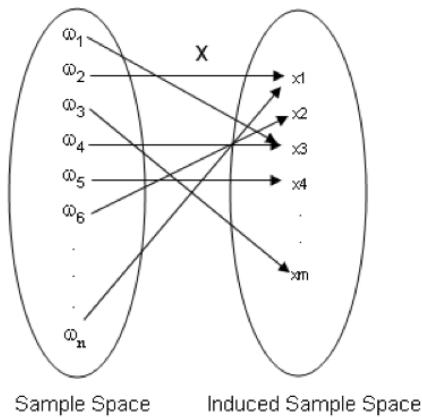


## 1.1 Random Variable

- Some sample spaces contain quantitative (numerical) outcomes, others contain qualitative outcomes.
- Often it is convenient to work with sample spaces containing numerical outcomes.
- A function that maps the original sample space into the real numbers is called a ‘random variable’.
- This is more useful when the original sample space contains qualitative outcomes.

### Definition 1: Random Variable

Let  $\Omega$  be a sample space. Let  $X$  be a function from  $\Omega$  to  $\mathfrak{R}$  (i.e.  $X : \Omega \rightarrow \mathfrak{R}$ ). Then  $X$  is called a random variable.



- A random variable assigns a real number to each outcome of a sample space.
- In other words, to each outcome of an experiment or a sample point  $\omega_i$ , of the sample spaces, there is a unique real number  $x_i$ , known as the value of the random variable  $X$ .
- The range of the random variable is called the *induced sample space*.
- *A note on notation:* Random variables will always denoted with uppercase letters and the realized values of the random variable (or its range) will be denoted by the corresponding lowercase letters. Thus, the random variable  $X$  can take the value  $x$ .
- Each outcome of a sample space occurs with a certain probability. Therefore, each possible value of a random variable is associated with a probability.
- Any events of a sample space can be written in terms of a suitably defined random variable.

### 1.1.1 Types of Random Variables

- A random variable is of two types
  - Discrete Random Variable
  - Continuous Random Variable

#### 1.1.1.1 Discrete Random Variable

- If the induced sample space is discrete, then the random variable is called a **discrete random variable**.

*Example 01* Consider the experiment of tossing a coin. Express the following events using a suitably defined random variable

$H =$  The event of getting a head

$T =$  The event of getting a tail

*Example 02*

Consider the experiment of rolling of a die. Express the following events using a suitably defined random variable

$A =$  *The event that the number faced up is less than 5*

$B =$  *The event that the number faced up is even*

$C =$  *The event that the number faced up is 2 or 5*

*Example 03*

Consider the experiment of tossing a coin 10 times. Then the sample space  $\Omega$  contains  $2^{10} = 1024$  outcomes. Each outcome is a sequence of 10 H's and T's.

Express the following events in terms of a suitably defined random variable.

$D =$  The event that the number of heads is 5

$E =$  The event that the number of tails is less than 4

### 1.1.1.2 Continuous Random Variable

- If the induced sample space is continuous, then the random variable is called a **continuous random variable**.

*Example 04*

Consider the experiment of measuring the lifetime (in hours) of a randomly selected bulb. Express the following events in terms of a suitably defined random variable.

$F$  = The event that the lifetime is less than 300 hours

$G$  = The event that the lifetime is 1000 hours

## 1.2 Probability Mass Function

**Definition 2: Discrete density function of a discrete random variable**

If  $X$  is a discrete random variable with distinct values  $x_1, x_2, \dots, x_n, \dots$ , then the function, denoted by  $f_X(\cdot)$  and defined by

$$f_X(x) = \begin{cases} P(X = x) & \text{if } x = x_j, j = 1, 2, \dots, n, \dots \\ 0 & \text{if } x \neq x_j \end{cases} \quad (1.1)$$

is defined to be the discrete density function of  $X$ .

- The values of a discrete random variable are often called *mass points*.
- $f_X(x)$  denotes the *mass* associated with the *mass point*  $x_j$ .
- **Probability mass function** *discrete frequency function* and *probability function* are other terms used in place of *discrete density function*
- Probability function gives the measure of probability for different values of  $X$ .

### 1.2.1 Properties of a Probability Mass Function

- Let  $X$  be a discrete random variable with probability mass function  $f_X(x)$ . Then,

1. For any  $x \in \mathfrak{R}$ ,  $0 \leq f_X(x) \leq 1$ .
2. Let  $E$  be an event and  $I = \{X(\omega) : \omega \in E\}$ . Then  $P(E) = P(X \in I) = \sum_{x \in I} f_X(x)$ .
3. Let  $R = \{X(\omega) : \omega \in \Omega\}$ . Then  $\sum_{x \in \mathfrak{R}} f_X(x) = 1$ .

### 1.2.2 Representations of Probability Mass Functions

*Example 05*

Consider the experiment of tossing a fair coin. Let

$$X = \begin{cases} 0 & \text{if the outcome is a Tail} \\ 1 & \text{if the outcome is a Head} \end{cases} \quad (1.2)$$

Find the probability mass function of  $X$ . Is  $X$  discrete or continuous?

**1.2.2.1 Using a table**

1.2. PROBABILITY MASS FUNCTION. STATISTICAL DISTRIBUTIONS

**1.2.2.2 Using a function**



### 1.2.2.3 Using a graph

## 1.3 Probability density function

- Let  $X$  be a continuous random variable.
- Then, it is not possible to define a pmf  $f_x$  with properties mentioned in Section 1.2. **Why?**
- Instead, we can find a function  $f_x$  with the some different properties.
- Probability density function (pdf) of a continuous random variable is a function that describes the relative likelihood for this random variable to occur at a given point.

### 1.3.1 Properties of a Probability Density Function

Let  $X$  be a continuous random variable with probability density function  $f_x$ . Then,

1. For any  $x \in \mathfrak{R}$ ,  $f_X(x) \geq 0$ .
2. Let  $E$  be an event and  $I = \{X(\omega) : \omega \in E\}$ . Then  $P(E) = P(X \in I) = \int_I f_X(x)dx$ .
3. Let  $R = \{X(\omega) : \omega \in \Omega\}$ . Then  $\int_{\mathfrak{R}} f_X(x)dx = 1$ .

### 1.3.2 Existence of pdf

- To see the existence of such a function, consider a continuous random variable  $X$ ,
- Suppose that we have a very large number of observations,  $N$ , of  $X$ , measured to high accuracy (large number of decimal places).
- consider the following grouped frequency table and the histogram constructed from those data.
- The height of the bar on a class interval of this histogram is equal to the relative frequency per unit in that class interval.

**1.4 Cumulative distribution function**

**1.5 Descriptive properties of distributions**

**1.6 Models for discrete distributions**

**1.7 Models for continuous distributions**



## Chapter 2

# Estimations

### 2.1 Point Estimation

#### 2.1.1 Methods of finding point estimators

#### 2.1.2 Methods of evaluating point estimators

### 2.2 Interval Estimation

#### 2.2.1 Interpretation of confidence intervals

#### 2.2.2 Methods of finding interval estimators

#### 2.2.3 Methods of evaluating interval estimators



## Chapter 3

# Hypothesis Testing

- 3.1 Null and alternative hypotheses
- 3.2 Errors in testing hypotheses-type I and type II error
- 3.3 Significance level, size, power of a test
- 3.4 Formulation of hypotheses
- 3.5 Methods of testing hypotheses

3.5. METHODS OF TESTING HYPOTHESES 3. HYPOTHESIS TESTING



## Chapter 4

# Design of Experiments

4.1 Introduction to experimental design

4.2 Basic principles of experimental design

4.3 Completely randomized design