COMPUTER SYSTEM MODELING AND SIMULATION
CSC 510

PREPARED BY:
OLUBORODE KAYODE O.

Course Guide

Introduction

CSC 510: Modelling and Simulation is a 3 credit unit course for students studying towards acquiring a Bachelor of Science in Computer Science and other related disciplines. The course is divided into 5 modules and 20 study units. It will first take a brief review of the concepts of Modelling and Simulation. This course will then go ahead to deal with the different stages involved developing good and functional Modelling and Simulation methods. The course went further to deal with different ways of Simulation implementations. This course also introduce such other knowledge that will enable the reader have proper understanding of Simulation such as Statistics.

The course guide therefore gives you an overview of what the course; CSC 510 is all about, the textbooks and other materials to be referenced, what you expect to know in each unit, andhow to work through the course material.

Course Competencies

The overall aim of this course, CSC 510 is to introduce you to basic concepts of Modelling and Simulation in order to enable you to understand the basic elements of Simulation use in day to day activities especially, businesses, Sciences and Technology industries. This course highlights methodology and approaches in the conduct of simulation. In this course of your studies, you will be put through the definitions of common terms in relation to modelling andsimulation, the methodology, theories, experiments and languages use in conducting simulations.

Course Objectives

It is important to note that each unit has specific objectives. Students should study them carefully before proceeding to subsequent units. Therefore, it may be useful to refer to these objectives in the course of your study of the unit to assess your progress. You should always look at the unit objectives after completing a unit. In this way, you can be sure that you have done what is required of you by the end of the unit.

However, below are overall objectives of this course. On completing this course, you should be able to:

- Define a model and modelling.
- Explain when to and why we use models
- Describe the modelling process
- Describe different types of Models.
- Describe how to generate pseudorandom numbers,
- Use QBasic RND function and describe how to simulate randomness,
- Use different Random number generators,
- Explain properties of good random number generator.
- Explain the use of Congruential method for generating Random numbers;
- Choose appropriate parameters for congruential method;
- Translate the method to computer programs;
- Use other very similar random number generating methods such as:
- Describe Monte Carlo method
- Trace the origin of Monte Carlo method
- Give examples of the application of Monte Carlo method
- Define Statistics

- Explain Statistical Distributions
- Compute measures of Central Tendency and Variations
- Explain the Components of Statistical Distributions
- Explain the role of probability distribution functions in simulations
- Describe Probability theory
- Explain the fundamental concepts of Probability theory
- Explain Random Variable
- Explain Limiting theorems
- Describe Probability distributions in simulations
- List common Probability distributions.
- Explain Simulation
- State why we need simulation
- Describe how simulations are done
- Describe various types of Simulations
- Give examples of Simulation
- Show areas of applications of Simulation
- Define Modelling
- Describe some basic modelling concepts
- Differentiate between Visual and Conceptual models
- Explain the Characteristics of Visual, models
- Define Finite Element Method (FEM)

Study Units

This course material contains four modules and sixteen study units as follows:

MODULE 1

UNIT 1 BASICS OF MODELLING AND SIMULATION
UNIT 2 RANDOM NUMBERS
UNIT 3 RANDOM NUMBER GENERATION
UNIT 4 MONTE CARLO METHOD
UNIT 5 STATISTICAL DISTRIBUTION FUNCTIONS
UNIT 6 COMMON PROBABILITY DISTRIBUTIONS

MODULE 2

UNIT 1 SIMULATION AND MODELLING
UNIT 2 MODELLING METHODS
UNIT 3 PHYSICS-BASED FINITE ELEMENT MODEL
UNIT 4 STATISTICS FOR MODELLING AND SIMULATION

MODULE 3

UNIT 1 SIMPLE THEORIES OF QUEUES
UNIT 2 BASIC PROBABILITY THEORIES IN QUEUING
UNIT 3 QUEUING MODELS
UNIT 4: QUEUING EXPERIMENTS

MODULE 4

UNIT 1 SIMULATION LANGUAGES
UNIT 2 THE SIMNET II LANGUAGE

MODULE 5

UNIT 1	STOCHASTIC PROCESSES
UNIT 2	RANDOM WALKS
UNIT 3	DATA COLLECTION
UNIT 4	CODING AND SCREENING

Module 1: MODELLING AND SIMULATION CONCEPTS

Module Introduction

This module is divided into six (6) units

Unit 1: Basics of Modelling and Simulation

Contents

1.0 Introduction

The ability of man to define what may happen in the future and to choose among alternatives lie at the heart of contemporary societies. Our knowledge of the way things work, in society or nature are trailed with clouds of imprecision, and vast harms have followed a belief in certainty. To reduce the level of disparity between outcome and reality, we require a decision analysis and support tool to enable us to evaluate, compare and optimize alternative. Such a tool should be able to provide explanations to various stakeholders and defend the decisions. One such tool that has been successfully employed is simulation which we use to vary the parameter of a model and observe the outcome.

Simulation has been particularly valuable:

a. When there is significant uncertainty regarding the outcome or consequences of a particular alternative under consideration. It allows you to deal with uncertainty and imprecision in a quantifiable way.
b. When the system under consideration involves complex interactions and requires input from multiple

disciplines. In this case, it is difficult for any one person to easily understand the system. A simulation of the model can in such situations act as the framework to integrate the various components in order to better understand their interactions. As such, it becomes a management tool that keeps you focused on the "big picture" without getting lost in unimportant details.

c. when the consequences of a proposed action, plan or design cannot be directly and immediately observed (i.e., the consequences are delayed in time and/or dispersed in space) and/or it is simply impractical or prohibitively expensive to test the alternativesdirectly.

2.0 Intended Learning Outcomes (ILOs)

After studying this unit, you should be able to:

- Define a model and modelling.
- Explain when to and why we use models
- Describe the modelling process
- Describe different types of Models.

3.0 Main Content

Modelling and Simulation Concepts

Modern science would be inconceivable without computers to gather data and run model simulations. Whether it involves bringing back pictures of the surface of the planet Mars ordetailed images to guide brain surgeons, computers have greatly extended our knowledge of the world around us and our ability to turn ideas into engineering reality. Thus modelling and computer simulation are important interdisciplinary tools.

3.1 Definitions

a. **Modelling** is the process of generating abstract, conceptual, graphical and/or mathematical models. Science offers a growing collection of methods, techniquesand theory about all kinds of specialized scientific modelling.

Modelling also means to find relations between systems and models. Stated otherwise, models are abstractions of real or imaginary worlds we create to understand their behaviour, play with them by performing "what if" experiments,make projections, animate or simply have fun.

Modelling: is the process of representing a model which includes its construction and working. This model is similar to a real system, which helps the analyst predict the effect of changes to the system. In other words, modelling is creating a model which represents a system including their properties. It is an act of building a model.

b. A **model** in general is a pattern, plan, representation (especially in miniature), or description designed to show the main object or workings of an object, system, orconcept.

c. A **model** (physical or hypothetical) is a representation of real-world phenomenon or elements (objects, concepts or events). Stated otherwise a model is an attempt toexpress a *possible structure of physical causality*.

Models in science are often theoretical constructs that represent any particular thingwith a set of variables and a set of logical and or quantitative relationships between them. Models in this sense are constructed to enable reasoning within an idealized logical framework about these processes and are an important component of scientific theories.

d. **Simulation -**is the manipulation of a model in such a way that it operates on timeor space to compress it, thus enabling one to perceive the interactions that would not otherwise be apparent because of their separation in time or space.

Simulation of a system is the operation of a model in terms of time or space, which helps analyze the performance of an existing or a proposed system. In other words, simulation is the process of using a model to study the performance of a system. It is an act of using a model for simulation.

e. **Modelling and Simulation** is a discipline for developing a level of understanding ofthe interaction of the parts of a system, and of the system as a whole. The level of understanding which may be developed via this discipline is seldom achievable viaany other discipline.

f. A **computer model** is a simulation or model of a situation in the real world or animaginary world which has parameters that the user can alter.

For example Newton considers movement (of planets and of masses) and writes equations, among which $f = ma$ (where f is force, m mass and a acceleration), that make the dynamics intelligible. Newton by this expression makes a formidable proposition, that force causes acceleration, with mass as proportionality coefficient. Another example, a model airplane is a physical representation of the real airplane; model of airplanes are useful in predicting the behaviour of the real airplane when subjected to different conditions; weather, speed, load, etc. Models help us frame our thinking about objects in the real world. It should be noted that more often than not we model dynamic (changing) systems.

3.2 History of Simulation

The historical perspective of simulation is as enumerated in a chronological order.

- **1940** − A method named 'Monte Carlo' was developed by researchers (John von Neumann, Stanislaw Ulan, Edward Teller, Herman Kahn) and physicists working on a Manhattan project to study neutron scattering.
- **1960** − The first special-purpose simulation languages were developed, such as SIMSCRIPT by Harry Markowitz at the RAND Corporation.
- **1970** − During this period, research was initiated on mathematical foundations of simulation.
- **1980** − During this period, PC-based simulation software, graphical user interfaces and object-oriented programming were developed.
- **1990** − During this period, web-based simulation, fancy animated graphics, simulation-based optimization, Markov-chain Monte Carlo methods were developed.

3.3 Developing Simulation Models

Simulation models consist of the following components: system entities, input variables, performance measures, and functional relationships. Following are the steps to develop a simulation model.

- **Step 1** − Identify the problem with an existing system or set requirements of a proposed system.
- **Step 2** − Design the problem while taking care of the existing system factors and limitations.
- **Step 3** − Collect and start processing the system data, observing its performance and result.
- **Step 4** − Develop the model using network diagrams and verify it using various verifications techniques.

- **Step 5** − Validate the model by comparing its performance under various conditions with the real system.
- **Step 6** − Create a document of the model for future use, which includes objectives, assumptions, input variables and performance in detail.
- **Step 7** − Select an appropriate experimental design as per requirement.
- **Step 8** − Induce experimental conditions on the model and observe the result.

3.3.1 Performing Simulation Analysis

Following are the steps to perform simulation analysis.

- **Step 1** − Prepare a problem statement.
- **Step 2** − Choose input variables and create entities for the simulation process. There are two types of variables - decision variables and uncontrollable variables. Decision variables are controlled by the programmer, whereas uncontrollable variables are the random variables.
- **Step 3** − Create constraints on the decision variables by assigning it to the simulation process.
- **Step 4** − Determine the output variables.
- **Step 5** − Collect data from the real-life system to input into the simulation.
- **Step 6** − Develop a flowchart showing the progress of the simulation process.
- **Step 7** − Choose an appropriate simulation software to run the model.
- **Step 8** − Verify the simulation model by comparing its result with the real-time system.
- **Step 9** − Perform an experiment on the model by changing the variable values to find the best solution.
- **Step 10** − Finally, apply these results into the real-time system.

3.3.2 Modelling & Simulation ─ Advantages

Following are the advantages of using Modelling and Simulation −

- **Easy to understand** − Allows to understand how the system really operates without working on real-time systems.
- **Easy to test** − Allows to make changes into the system and their effect on the output without working on real-time systems.
- **Easy to upgrade** − Allows to determine the system requirements by applying different configurations.
- **Easy to identifying constraints** − Allows to perform bottleneck analysis that causes delay in the work process, information, etc.
- **Easy to diagnose problems** − Certain systems are so complex that it is not easy to understand their interaction at a time. However, Modelling & Simulation allows to understand all the interactions and analyze their effect. Additionally, new policies, operations, and procedures can be explored without affecting the real system.

3.3.2 Modelling & Simulation ─ Disadvantages

Following are the disadvantages of using Modelling and Simulation −

- Designing a model is an art which requires domain knowledge, training and experience.
- Operations are performed on the system using random number, hence difficult to predict the result.
- Simulation requires manpower and it is a time-consuming process.
- Simulation results are difficult to translate. It requires experts to understand.
- Simulation process is expensive.

3.4 What is Modelling and Simulation?

Modelling is a discipline for developing a level of understanding of the interaction of the parts of a system, and of the system as a whole. The level of understanding which may be developed via this discipline is seldom achievable via any other discipline.

A simulation is a technique (not a method) for representing a dynamic real world systemby a model and experimenting with the model in order to gain information about thesystem and therefore take appropriate decision. Simulation can be done by hand or by a computer. Simulations are generally iterative in their development. One develops a model, simulates it, learns from the result, revises the model, and continues the iterations until an adequate level of understanding is attained.

Modelling and Simulation is a discipline, it is also very much an art form. One can learnabout riding a bicycle from reading a book. To really learn to ride a bicycle one must become actively engaged with a bicycle. Modelling and Simulation follows much the samereality. You can learn much about modelling and simulation from reading books and talking with other people. Skill and talent in developing models and performing simulations is only developed through the building of models and simulating them. It is very much ―learn as you go‖ process. From the interaction of the developer and the models emerges an understanding of what makes sense and what doesn't.

3.5 Type of Models

There are many types of models and different ways of classifying/grouping them. For simplicity, Models may be grouped into the following – Physical, Mathematical, Analogue, Simulation, Heuristic, Stochastic and Deterministic models.

a. <u>Physical Models</u>

These are call iconic models. Good examples of physical models are model cars, model railway, model airplane, scale models, etc. A model railway can be used to study the behaviour of a real railway, also scale models can be used to study a plant layout design. Insimulation studies, iconic models are rarely used.

b. <u>Mathematical Models</u>

These are models used for predictive (projecting) purposes. They are abstract and take the form of mathematical expressions of relationships. For examples:

1. $x^2 + y^2 = 1$ (mathematical model of a circle of radius 1)
2. Interest =Principal x Rate x Time/100
3. Linear programming models and so on.

Mathematical models can be as simple as interest earnings on a savings account or as complex as the operation of an entire factory or landing astronauts on the moon. The development of mathematical models requires great deal of skill and knowledge.

c. <u>Analogue Models</u>

These are similar to iconic models. But here some other entities are used to represent directly the entities of the real world. An example is the analogue computer where the magnitudes of the electrical currents flowing in a circuit can be used to represent quantitiesof materials or people moving around in a system. Other examples are; the gauge used to check the pressure in a tyre. The movement of the dial represent the air pressure in the tyre.In medical examination, the marks of electrical current on paper, is the analogue representation of the working of muscles or organs.

d. Simulation Models

Here, instead of entities being represented physically, they are represented by sequences ofrandom numbers subject to the assumptions of the model. These models represent (emulate) the behaviour of a real system. They are used where there are no suitable mathematical models or where the mathematical model is too complex or where it is not possible to experiment upon a working system without causing serious disruption.

e. Heuristic Models

These models use intuitive (or futuristic) rules with the hope that it will produce workable solutions, which can be improved upon. For example, the Arthur C Clerk‘s heuristic modelwas the forerunner of the communications satellite and today‘s international television broadcast.

f. Deterministic Models

These are models that contain certain known and fixed constants throughout their formulation e.g., Economic Order Quantity (EOQ) for inventory control under uncertainty.

g. Stochastic models

probabilities.

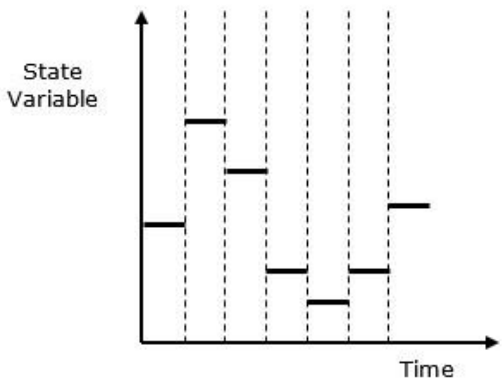
3.6 Advantages of Using Models

- They are safer.
- They are less expensive. For example, Practical Simulators are used to train pilots.
- They are easier to control than the real world counterparts.

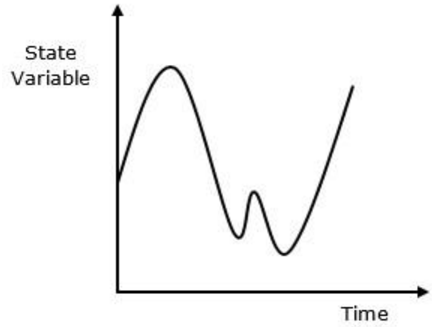
3.6.1 Classification of Models

A system can be classified into the following categories.

- **Discrete-Event Simulation Model** − In this model, the state variable values change only at some discrete points in time where the events occur. Events will only occur at the defined activity time and delays.
- **Stochastic vs. Deterministic Systems** − Stochastic systems are not affected by randomness and their output is not a random variable, whereas deterministic systems are affected by randomness and their output is a random variable.
- **Static vs. Dynamic Simulation** − Static simulation include models which are not affected with time. For example: Monte Carlo Model. Dynamic Simulation include models which are affected with time.
- **Discrete vs. Continuous Systems** − Discrete system is affected by the state variable changes at a discrete point of time. Its behavior is depicted in the following graphical representation.

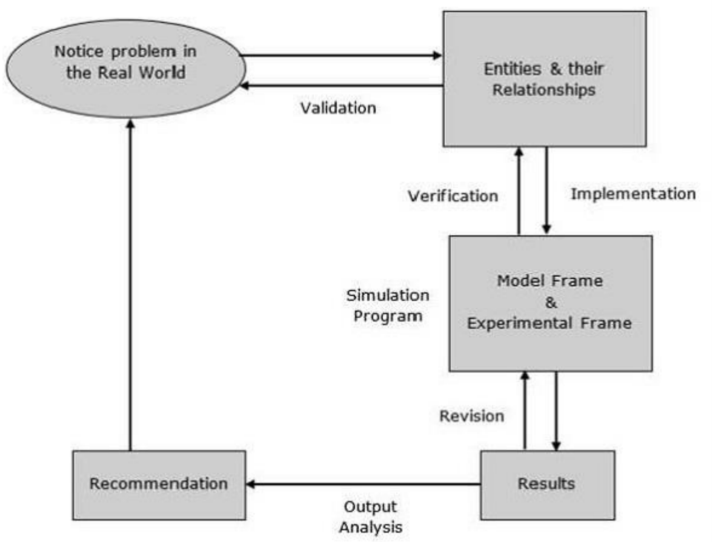


Continuous system is affected by the state variable, which changes continuously as a function with time. Its behavior is depicted in the following graphical representation.



Modelling Process

Modelling process includes the following steps.



Step 1 − Examine the problem. In this stage, we must understand the problem and choose its classification accordingly, such as deterministic or stochastic.

Step 2 − Design a model. In this stage, we have to perform the following simple tasks which help us design a model −

- Collect data as per the system behavior and future requirements.
- Analyze the system features, its assumptions and necessary actions to be taken to make the model successful.
- Determine the variable names, functions, its units, relationships, and their applications used in the model.
- Solve the model using a suitable technique and verify the result using verification methods. Next, validate the result.
- Prepare a report which includes results, interpretations, conclusion, and suggestions.

Step 3 − Provide recommendations after completing the entire process related to the model. It includes investment, resources, algorithms, techniques, etc.

3.7 Applications

One application of scientific modelling is the field of "Modelling and Simulation", generally referred to as "M&S". M&S has a spectrum of applications which range from concept development and analysis, through experimentation, measurement and verification, to disposal analysis. Projects and programs may use hundreds of different simulations, simulators and model analysis tools

3.8 Modelling Procedure

In modelling we construct a suitable representation of an identified real world problem,obtain solution(s) for that representation and interpret each solution in terms of the real situation. The steps involved in modelling are as follows:

1. Examine the real world situation.
2. Extract the essential features from the real world situation.
3. Construct a model of the real (object or system) using just the essential featuresidentified.
4. Solve and experiment with the model.
5. Draw conclusions about the model.
6. If a further refinement necessary, then re-examine the model and readjust parameters and continue at 4, otherwise continue at 7.
7. Proceed with implementation.

Explanation of the Steps

Begin with the real world situation, which is to be investigated with a view to solving someproblem or improving that situation.

The first important step is to extract from the real world situation the essential features to be included in the model. Include only factors that make the model a meaningful representation of reality, while not creating a model, which is difficult by including many variables that do not have much effect. Factors to be considered include ones over which management has control and external factors beyond management control. For the factors included, assumptions have to be made about their behaviour.

Run (simulate) the model and measure what happens. For example, if we have simulation ofa queuing situation where two servers are employed, we can run this for hundreds ofcustomers passing through the system and obtain results such as the average length of the queue and the average waiting time per customer. We can then run it with three servers, say, and see what new values are obtained for these parameters. Many such runs can be carried out making different changes to the structure and assumptions of the model.

In the case of a mathematical model we have to solve a set of equations of some sort, e.g. linear programming problem where we have to solve a set of constraints as simultaneous equations, or in stock control – where we have to use previously accumulated data to predict the future value of a particular variable.

When we have solved our mathematical model or evaluated some simulation runs, we can now draw some conclusions about the model. For example, if we have the average queue length and the average waiting time for a queuing situation varied in some ways, we can use this in conjunction with information on such matters as the wage-rates for servers and value of time lost in the queue to arrive at decisions on the best way to service the queue.

Finally, we use our conclusions about the model to draw some conclusions about the original real world situation. The validity of the conclusions will depend on how well our model actually represented the real world situation.

Usually the first attempt at modelling the situation will almost certainly lead to results at variance with reality. We have to look back at the assumptions in the model and adjust them. The model must be rebuilt and new results obtained. Usually, a large number of iterations of this form will be required before acceptable model is obtained. When an acceptable modelhas been obtained, it is necessary to test the sensitivity of that model to possible changes in condition.

The modelling process can then be considered for implementation when it is decided that themodel is presenting the real world (object or system) sufficiently well for conclusions drawn from it to be a useful guide to action.

The model can be solved by hand, especially if it is simple. It could take time to arrive at an acceptable model. For complex models or models which involve tremendous amount of data, the computer is very useful.

4.0 Self-Assessment Exercise(s)

Answer the following questions:

1. Differentiate between Model, Modelling, Simulation and Computer model.
2. What are the steps followed in modelling?
3. State why we use models

5.0 Conclusion

In this unit we took a look at an overview of major concepts that underlie models to prepare us for the work in this course simulation and modelling.

6.0 Summary

In introducing this unit, it was stated that simulation is a decision support tool which enable us to evaluate, compare and optimize alternative ways of solving a problem and the following were discussed:

- Modelling was defined
- The concepts of modelling were outlined
- Why we use models
- The application of models especially for simulations
- The types of models which include: Physical, Mathematical, Analogue,Simulation, Heuristic, Stochastic and Deterministic models were highlighted.

7.0 Further Readings

- Devore, J. L. (2018). *Probability and statistics for engineering and the sciences*. Toronto,Ontario: Nelson.
- Georgii, H. (2013). *Stochastics: Introduction to probability and statistics*. Berlin: DeGruyter.
- Giri, N. C. (2019). *Introduction to probability and statistics*. London: Routledge.
- Johnson, R. A., Miller, I., & Freund, J. E. (2019). *Miller & Freunds probability andstatistics for engineers*. Boston: Pearson Education.
- Laha, R. G., & Rohatgi, V. K. (2020). *Probability theory*. Mineola, NY: DoverPublications.

Unit 2: Random Numbers

Contents

1.0 Introduction

The use of Random numbers lies at the foundation of modelling and simulations. Computer applications such as simulations, games, graphics, etc., often need the ability to generate random numbers for such application.

The quality of a random number generator is proportional to its **period**, or the number of random numbers it can produce before a repeating pattern sets in. In large-scale simulations, different algorithms (called shift-register and lagged-Fibonacci) can be used, although these also have some drawbacks, combining two different types of generators may produce the best results.

2.0 Intended Learning Outcomes (ILOs)

By the end this unit, you should be able to:

- Describe how to generate pseudorandom numbers,
- Use QBasic RND function and describe how to simulate randomness,
- Use different Random number generators,
- Explain properties of good random number generator.

3.0 Main Content

Random Number can be defined as numbers that show no consistent pattern, with each number in a series and are neither affected in any way by the preceding number, norpredictable from it.

One way to get random digits is to simply start with an arbitrary number with a specified number of digits, for example 4 digits. The first number is called the **seed**. The seed is multiplied by a **constant** number of the same number of digits (length), and the desired number of digits is taken off the right end of the product. The result becomes the new seed. Itis again multiplied by the original constant to generate a new product, and the process is repeated as often as desired. The result is a series of digits that appear randomly distributedas though generated by throwing a die or spinning a wheel. This type of algorithm is called a **congruential generator**.

Generating a random number series from a single seed works fine with most simulations that rely upon generating random events under the control of probabilities (Monte Carlo simulations). However, although the sequence of numbers generated from a given seed is randomly distributed, it is always the same series of numbers for the same seed. Thus, a computer poker game that simply used a given seed would always generate the same hands for each player.

What is needed is a large collection of potential seeds from which one can be more or less randomly chosen. If there are enough possible seeds, the odds of ever getting the same series of numbers become diminishingly small.

One way to do this is to read the time (and perhaps date) from the computer‘s system clock and generate a seed based on that value. Since the clock value is in milliseconds, there are millions of possible values to choose from. Another common technique is to use the interval between the user‘s keystrokes (in milliseconds). Although they are not perfect, these techniques are quite adequate for games.

The so-called true random number generators extract random numbers from physical phenomena such as a radioactive source or even atmospheric noise as detected by a radio receiver.

3.1 Pseudorandom Number Generation

In this section we look at how random numbers may be generated by human beings for use insimulating a system or by computer for use while simulating an event.

What we usually do is to take for instance ten pieces of papers and number them 0,1,2,3,4,5,6,7,8, and 9 , fold and place them in a box. Shake the box and thoroughly mix the slips of paper. Select a slip; then record the number that is on it. Replace the slip and repeat this procedure over and over. The resultant record of digits is a realized sequence of random numbers. Assuming you thoroughly mix the slips before every draw, the nth digit of the sequence has an equal or uniform chance of being any of the digits 0, 1, 2,3,4,5,6,7,8, 9 irrespective of all the preceding digits in the recorded sequence.

In some simulations, we use random numbers that are between 0 and 1. For example, if you need such numbers with four decimal digits, then you can take four at a time from the recorded sequence of random digits, and place a decimal point in front of each group of four. To illustrate, if the sequence of digits is 358083429261… then the four decimal placed random numbers are .3580, .8342, and .9261.

3.2 Random Numbers in Computer

How does computer generate a sequence of random numbers?

One way is to perform the above ―slip-in-a-box‖ experiment and then store the recorded sequence in a computer-backing store.

The RAND Corporation using specially designed electronic equipment, to perform the experiment, actually did generate a table of a million random digits. The table can beobtained on tape, so that blocks of the numbers can be read into the memory of a high- speed computer, as they are needed. Their approach is disadvantageous since considerable computer time was expended in the delays of reading numbers into memory from a tape drive.

Experts in computer science have devised mathematical processes for generating digits that yield sequences satisfying many of the statistical properties of a truly random process. To illustrate, if you examine a long sequence of digits produced by deterministic formulas, each digit will occur with nearly the same frequency, odd numbers will be followed by even numbers about as often as by odd numbers, different pairs of numbers occur with nearly the same frequency, etc. Since such a process is not really random, it is called **pseudo-random number generator**.

The other ways of generating pseudo-random numbers are:

1. Computer simulation languages and indeed some programming languages such as BASIC have built-in pseudo-random number generators. In computer simulationsituations where this facility is not available in the language you are using, you will have to write you own pseudo-random number generator (see how to do this later).
2. The results of experiments such as the one previously describe above are published in books of statistical tables. In hand simulation, it may be appropriate to use a published table of random numbers.
3. The conventional six-sided unbiased die may also be used to generate a sequence of random digits in the set (1, 2, 3, 4, 5, 6) where each digit has a probability 1/6 of occurrence.

Exercise

Suggest one or two experimental set-ups (analogous to the slip-in-a-box approach) forgenerating uniform random digits.

3.3 Using the RND Function in BASIC

The BASIC programming language has a numeric function named RND, which generates random numbers between 0 and 1. Each time RND is executed, a pseudo random number between 0 and 1 is generated. Using RND function at any time will always generate the same sequence of pseudo random numbers unless we vary the random number seed using the BASIC statement:

RANDOMIZE

This way, we can control the sequence of random numbers generated. RANDOMIZE will result to the following prompt on the VDU:

Random Number Seed (-32768 to 32767)?

Suppose your response to the above prompt is 100. Then the computer would use this number, 100, to generate the first random number. This number generated is used to generatethe next random number. Thus by specifying the seed for the first random number, we are ina way controlling all random numbers that will be generated until the seed is reset. A control such as this can be very useful in validating a simulation program or other computer programs that use random numbers.

Consider the following BASIC program:

```
FOR K% = 1 TO 5
PRINT RND
NEXT K%
END
```

If the above program is run, some seven-digit decimal numbers like the following will bedisplayed:
.6291626, .1948297, .6305799, .8625749, .736353.
The particulardigits displayed depend on the system time.

Every time you run the above program, different sequence of numbers will be displayed.

Now add a RANDOMIZE statement to the program:

```
RANDOMIZE TIMER FOR K% = 1 TO 5
PRINT RND
NEXT K%
END
```

If you run this program with 300 as a response to the prompt for the random number seed,the following may be displayed:

.1851404, .9877729, .806621, .8573399, .6208935

Exercise

Find out whether the same set of random numbers will be displayed each time the aboveprogram is run with seed 300.

3.4 Simulating Randomness

Suppose we want to simulate the throwing of a fair die. A random number between 0 and 1 will not always satisfy our needs. If the die is fair, throwing it several times will yield a seriesof uniformly distributed integers 1,2,3,4,5 and 6. Consequently we need to be able to generate a random integer with values in the range 1 and 6 inclusive.

Now the function RND generates a random number between 0 and 1. Specifically, the random variable X is in the range: $0 \leq X < 1$

The expression X = RND *6

Will generate a number in the range: $0 \leq X < 6$

We must convert these numbers to integers as follows: X% = INT (RND*6)

The expression produces an integer in the range: $0 \leq X < 5$

But we need the range: $0 \leq X < 6$

Therefore if we need to add 1 to the above expression in simulating the tossing of a die. Thus,

X% = INT (RND*6) + 1

In general, to generate a random integer between P and N we use the expression:

INT(RND*N+1-P) + P;

where N>P

While for integer number between 0 and N – 1 we use the expression INT (RND *N).

Example 1

A simple QBASIC program that will stimulate the tossing of two dice and display the value obtained after each toss, and the total value of the dice is shown below.

```
CLS
REM DI and D2 represent the individual dice.
RANDOMIZE
DO
```

```
D1% = INT(RND*6) + 1
D2% = INT(RND*6) + 1
TOTAL% = D1% + D2%
PRINT –Die 1:‖;
D1%; –Die 2:‖;
D2%
PRINT: PRINT
INPUT –Toss Again (Y/N)?‖, Y$ LOOP UNTIL UCASE$(Y$) = –N‖

END
```

Exercise

Run the program of example 1several times using different random number seeds to determine if the integers generated for the individual die are uniformly distributed between 1and 6 inclusive.

If we want the computer to be generating the random number seed automatically, we useRANDOMIZE TIMER

In place of RANDOMIZE.

Example 2

Another QBASIC program to simulate the tossing of a fair coin 10 times. The programdisplays a H when a head appears and a T when a tail appears.

```
CLS
REM Program to simulate the tossing of a coin 10 timesREM and print the outcome
RANDOMIZE TIMER
FOR K% = 1 TO 10
RANDNO = RND
IF RANDNO <= 0.5
PRINT –H‖
IF RANDNO > 0.5
PRINT –T‖
NEXT K%
END
```

Example 3

Suppose the output of the program of example 3 is: HHTHHTTTHH and that there are two players X and Y involved in the tossing of the coin. Given that playerX wins N50.00 from player Y if a head appears and loses it to player Y if a tail appears.
Determine who won the game and by how much.

Solution

From the output there are 6 heads and 4 tails.

Player X wins N50.00 x 6 = N300.00 from player Y. He loses N50.00 x 4 = N200.00 toplayer Y.

Thus, player X won the game with N300.00 – N200.00 = N100.00.

3.5 Properties of a Good Random Number Generator

The random numbers generated should;

a. have as nearly as possible a uniform distribution.
b. should be fast
c. not require large amounts of memory.
d. have a long period.
e. be able to generate a different set of random numbers or a series of numbers.
f. not degenerate.

Self-Assessment Exercise(s)

Write a QBASIC program to generate thirty random integer numbers distributed between 20 and 50. Your program should ensure that no number is repeated.

Write a QBASIC program to accept a set of characters from the keyboard and then move thecharacters randomly across the screen. The movement of the characters should stop once a key is pressed on the keyboard. The set of characters should also change colors randomly at the point of the movement.

What is a seed and explain how you can generate random numbers using a seed.Define a period and state how to improve a period.

5.0 Conclusion

In this unit, you have been introduced to Random Numbers generation. You have also learntthe how to manipulate the RND function of QBASIC.

6.0 Summary

What you have learnt in this unit concern:

- The different ways of generating pseudorandom numbers,
- The properties of good random number generator.
- The use of QBasic RND function to simulate randomness,
- The other Random number generating methods,

7.0 Further Readings

- Devore, J. L. (2018). *Probability and statistics for engineering and the sciences*. Toronto,Ontario: Nelson.
- Georgii, H. (2013). *Stochastics: Introduction to probability and statistics*. Berlin: DeGruyter.
- Giri, N. C. (2019). *Introduction to probability and statistics*. London: Routledge.
- Johnson, R. A., Miller, I., & Freund, J. E. (2019). *Miller & Freunds probability andstatistics for engineers*. Boston: Pearson Education.
- Laha, R. G., & Rohatgi, V. K. (2020). *Probability theory*. Mineola, NY: DoverPublications.
- Mathai, A. M., & Haubold, H. J. (2018). *Probability and statistics: A course forphysicists and engineers*. Boston: De Gruyter.
- Pishro-Nik, H. (2014). *Introduction to probability, statistics, and random processes*. BlueBell, PA: Kappa Research, LLC.
- Spiegel, M. R., Schiller, J. J., & Srinivasan, R. A. (2013). *Schaums outline of probabilityand statistics*. New York: McGraw-Hill.

Unit 3: Congruential Random Number Generator

Contents

1.0 Introduction

As has been stated earlier, if you want to write a simulation program and you neither havea simulation language at your disposal nor a programming language with a random numbergenerating function, you must design and write a random number generating function and call it whenever you need it.

Classical uniform random number generators have some major defects, such as, short period length and lack of higher dimension uniformity. However, nowadays there are a class of rather complex random number generators which are as efficient as the classical generators which enjoy the property of a much longer period and of higher dimension uniformity.

2.0 Intended Learning Outcomes (ILOs)

By the end of this unit, the reader should be able to:

- Explain the use Congruential method for generating Random numbers;
- Choose appropriate parameters for congruential method;
- Translate the method to computer programs;
- Use other very similar random number generating methods such as: Mid square, Mid product, Fibonacci

3.0 Main Content

3.1 The Congruential Method

The Congruential Method is widely used. The method is based on modulus arithmetic,which we now discuss briefly.

We say that two numbers x and y are congruent modulo m if (x-y) is an integral multipleof m. Thus we can write: x = y(modulo m)

For example, let m = 10, then we can write:

$13 \equiv 3$ (modulo 10)

$84 \equiv 4$ (modulo 10)

The ***congruential method*** generates random numbers by computing the next random number from the last random number obtained, given an initial random number say, X_0,called the seed.

The method uses the formula:
$X_{n+1} = (aX_n + c)(\text{modulo } m)$
where X_0 = Seed and; a, c < m,
Where a, c and m are carefully chosen positive integer constants of which a and c must be less than m, X_0 is the seed or the last random number generated in the sequence. Stated in the computer language, the above formula becomes:

$X_{(N+1)} = (A^* X_{(N)} + C)$ MOD M (FORTRAN)
or
R = (A* SEED + C) MOD M
(QBASIC)

From the above formula, it follows that the random number generated must be between 0 and (m-1) since MOD (modulo) produces remainder after division. Hence the aboveformula will produce the remainder after dividing ($aX_n + C$) by m. So to generate a random number between p and m we use:

$X_{n+1} = (aX_n + C)(\text{modulo } m + 1\text{-}p) + p$, for $m > p$.

If the value of c is zero, the congruential method is termed Multiplicative Congruential Method. If the value of c is not zero, the method is called Mixed Congruential Method.

The **multiplicative congruential** method is very handy. It is obtained using the general formula:

$r_n = ar_{n-1}$ (modulo m)

where the parameters a, m and the seed r_0 are specified to give desirable statistical properties of the resultant sequence. By virtue of modulo arithmetic, each r_n must be one ofthe numbers 0,1,2,3… m-1. Clearly, you must be careful about the choice of a and r_0. The values of ‗a' and r_0 should be chosen to yield the largest cycle or period, that is to give thelargest value for n at which $r_n = r_0$ for the first time.

Example 4
To illustrate the technique, suppose you want to generate ten decimal place numbers $u_1, u_2, u_3, \ldots$. It can be shown that if you use $u_n = r_n \times 10^{-1}$
where $r_n = 100003 r_{n-1}$ (modulo 10^{10}), and r_0 = any odd number not divisible by 5,
then the period of the sequence will be 5×10^8, that is $r_n = r_0$ for the first time at $n = 5 \times 10^8$ and the cycle subsequently repeats itself.

As an example, using our mixed congruential formula$X_{n+1} = (aX_n + c)$ (modulo m),
And suppose m = 8, a = 5, c = 7 and X_0 (seed) = 4 we can generate a random sequence ofinteger numbers thus:

n	$X_{n+1} = (5X_n+7) \mod 8$
0	X1 = (5*X0+7)mod 8 = (5*X4+7)mod 8 = 27 mod 8 = 3
1	X2 = (5*X1+7)mod 8 = (5*X3+7)mod 8 = 22 mod 8 = 6
2	X3 = (5*X2+7)mod 8 = (5*X6+7)mod 8 = 37 mod 8 = 5
3	X4 = (5*X3+7)mod 8 = (5*X5+7)mod 8 = 32 mod 8 = 0
4	X5 = (5*X4+7)mod 8 = (5*X0+7)mod 8 = 7 mod 8 = 7
5	X6 = (5*X5+7)mod 8 = (5*X7+7)mod 8 = 42 mod 8 = 2
6	X7 = (5*X6+7)mod 8 = (5*X2+7)mod 8 = 17 mod 8 = 1
7	X8 = (5*X7+7)mod 8 = (5*X1+7)mod 8 = 12 mod 8 = 4

Note that the value of X_8 is 4, which is the value of the seed X_0. So if we compute X_9, X_{10},etc the same random numbers 3,6,5,0,7,2,1,4 will be generated once more.

Note also that if we divide the random integer values by 8, we obtain random numbers inthe range $0 \leq X_{n+1} < 1$ which is similar to using the RND function of BASIC.

3.1 Choice of a, c and m

The method of this random number generation by *linear congruential method,* works by computing each successive random number from the previous. Starting with a seed, Xo,the linear congruential method uses the following formula:

$X_{i+1} = (A*Xi + C) \mod M$

In his book, *The Art of Computer Programming,* Donald Knuth presents several rules for maximizing the length of time before the random number generator comes up with thesame value as the seed. This is desirable because once the random number generator comesup with the initial seed, it will start to repeat the same sequence of random numbers (which will not be so random since the second time around we can predict what they will be).According to Knuth's rules, if M is prime, we can let C be 0.
The LCM defined above has full period if and only if the following conditions aresatisfied:

a) m and c are relatively prime
b) If q is a prime number that divides m , then q divides a-1
c) If 4 divides m, then 4 divides a-1

Therefore, the values for a, c and m are not generated randomly, rather they are carefully chosen based on certain considerations. For a binary computer with a word length of r bits, the normal choice for m is $m = 2^{r-1}$. With this choice of **m**, **a** can assume any of the values 1, 5, 9, 13, and **c** can assume any of the values 1, 3, 5, 7… However, experience shows that the congruential method works out very well if the value of a is an odd integer not divisibleby either 3 or 5 and c chosen such that c mod 8 = 5 (for a binary computer) or c mod 200 = 21 (for a decimal computer).

Example 5

Develop a function procedure called RAND in QBASIC which generates a random numberbetween 0 and 1 using the mixed congruential method. Assume a 16-bit computer.

Solution

```
FUNCTION RAND (SEED)
CONST M = 32767, A = 2743, C = 5923
```

```
IF SEED < 0 THEN
SEED = SEED + M
SEED = (A* SEED + C) MOD M
RAND = SEED/M
END FUNCTION
```

Note that in the main program that references the above function in (a), the TIMER function can be used to generate the SEED to be passed to the function RAND as illustrated in example 2.

Example 5

Write a program that can generate that can generate 20 random integer number distributed between 1 and 64 inclusive using mixed congruential method.

Solution QBASIC

```
DECLARE RAND (X)
CLS: REM Mixed Congruential MethodDIM SHARED SEED
SEED = TIMER FOR K% = 1 TO 20

SEED = RAND (SEED)                                   _Call of function RAND
PRINT SEED: SPC (2)
NEXT K%
END _End of main program

FUNCTION RAND (SEED) _Beginning of function subprogramCONST M = 64 A = 27, C = 13
IF SEED = THEN SEED = SEED + MSEED = (a* SEED + C) MOD M + 1 RAND = SEED
END FUNCTION _End of the function program RAND
```

(b) Using FORTRAN

```
PROGRAM RANDNUM COMMON SEED
CLS _Clear screen
DO 50 K = 1, 25
WRITE (*, 5)
FORMAT(/)
50 CONTINUE
WRITE(*,*) _Enter the seed‘READ(*,*) SEED
DO 30 J = 1, 20
SEED = RAND
WRITE (*, 6) SEED
6       FORMAT (I4)
30      CONTINUEEND

FUNCTION RANDCOMMON SEED
PARAMETER (M = 64, A = 27, C =13)
IF (SEED.LT.0) SEED = SEED + M
HOLD = (A*SEED + C)
```

```
SEED = AMOD (HOLD,M) + 1
RAND = SEED
RETURN END
```

3.3 RANECU Random Number Generator

A FORTRAN code for generating uniform random numbers on [0,1]. RANECU is multiplicative linear congruential generator suitable for a 16-bit platform. It combines three simple generators, and has a period exceeding 81012.

It is constructed for more efficient use by providing for a sequence of such numbers (Length), to be returned in a single call. A set of three non-zero integer seeds can be supplied, failing which a default set is employed. If supplied, these three seeds, in order, should lie inthe ranges [1, 32362], [1, 31726] and [1, 31656] respectively. The program is given below.

```
SUBROUTINE RANECU (RVEC,LEN)
C Portable random number generator for 16 bit computer.
C Generates a sequence of LEN pseudo-random numbers, returnedC in RVEC.
DIMENSION RVEC(*)
SAVE ISEED1,ISEED2, ISEED3
DATA ISEED1,ISEED2,ISEED3/1234, 5678, 9876/
C Default values, used if none is suppliednvia an ENTRY
C call at RECUIN
        DO 100 I = 1,LEN
        K=ISEED1/206
ISEED1 = 157 * (ISEED1 - K * 206) - K * 21
IF(ISEED1.LT.0) ISEED1=ISEED1+32363 K=ISEED2/217
ISEED2 = 146 * (ISEED2 - K*217) - K* 45
IF(ISEED2.LT.O) ISEED2=ISEED2+31727K=ISEED3/222
ISEED3 = 142 * (ISEED3 - K *222) - K * 133
IF(ISEED3.LT.0) ISEED3=ISEED3+31657 IZ=ISEED1-ISEED2
IF(IZ.GT.706)IZ = Z - 32362 IZ = 1Z+ISEED3
IF(IZ.LT.1)IZ = 1Z + 32362 RVEC(I)=REAL(IZ) * 3.0899E – 5
100     CONTINUE
RETURN
ENTRY RECUIN(IS1, IS2, IS3)
ISEED1=IS1
ISEED2=IS2
ISEED3=IS3
RETURN
ENTRY RECUUT(IS1,IS2,IS3)
IS1=ISEED1
IS2=ISEED2
IS3=ISEED3
RETURN
END
```

3.4 Other Methods of Generating Random Numbers

Some other methods of generating random numbers are Quadratic Congruential Method, Mid-square method, Mid-product Method and Fibonnachi Method.

3.4.1 The Quadratic congruential method

This method uses the formula:

$Xn=1 = (dX^{2} + cX + a)$ modulo m

Where d is chosen in the same way as c and m should be a power of 2 for the method to yieldsatisfactory results.

3.4.2 The Mid-square method

The first random number is generated from the seed by squaring the seed and discarding all the digits except the middle four digits. This number is subsequently used as the new seed to generate the next random number in the same manner and so on.

The formula is: $X_{n+1} = X_n^{2}$

The mid-square method is rarely used these days as it has the tendency to degenerate rapidly. Also, if the number zero is ever generated, then all subsequent numbers generated will be zero. Furthermore, the method is slow when simulated in the computer since many multiplications and divisions are required to access the middle four digits.

3.4.3 The mid-product method

This method is similar to the mid-square method, except that a successive random number isobtained by multiplying the current number by a constant c, and taking the middle digits.

The formula is: $X_{n+1} = cX_n$

The mid-product method has a longer period and it is more uniformly distributed than themid-square method.

3.4.4 The Fibonacci method

Fibonacci method uses the formula: $X_{n+1} = (X_n + X_{n-1})$ modulo m

In this method, two initial seeds need to be provided. However, experience has shown that the random numbers generated by using Fibonacci method fail to pass tests for randomness.Therefore, the method does not give satisfactory results.

From the foregoing discussions, it is obvious that the last three methods – mid-square, mid-product and Fibonacci are of historical significance and have detrimental and limiting characteristics.

4.0 Self-Assessment Exercise(s)

4.0

1. Write a QBASIC program using Quadratic congruential method to generate 15random integer numbers between 1 and 50.
2. Produce a table of random numbers using multiplicative congruential method, using a
=5, m =8 and $X_0 = 4$. Draw an inference from your solution.
3. Write a QBASIC function that can be referenced as computer random numberbetween 30 and 100 using mixed congruential method.
4. Use the mixed congruential method to generate the following sequences of randomnumbers:

a. A sequence of 10 one-digit random numbers given that

   $X_{n+1} \equiv (X_n + 3)$(modulo 10) and $X_0 = 2$

b. A sequence of eight random numbers between 0 and 7 given that

   $X_{n+1} \equiv (5X_n + 1)$(modulo 8) and $X_0 = 4$

c. A sequence of two-digit random numbers such that

$X_{n+1} \equiv (61X_n + 27)$(modulo 100) and $X_0 = 40$

d. A sequence of five-digit random numbers such that

$X_{n+1} \equiv (21X_n + 53)$(modulo 100) and $X_0 = 33$

5. Define a methods period and state how to improve a period. Show two examples ofsuch improvement.
6. Consider the multiplicative congruential method for generating random digits. Ineach part below, assume modulo 10 arithmetic and determine the length of the cycle:

a. Let a = 2 and r_0 = 1, 3 and 5

b. Let a = 3 and r_0 = 1,2 and 5

5.0 Conclusion

In this unit, you have been introduced to Random Numbers generation. You have alsolearnt how to design random number generator.

6.0 Summary

What you have learnt in this unit concern:

- The Congruential methods of generating random numbers,
- The use of QBasic RND function to simulate randomness,
- The other Random number generating methods,
- The properties of good random number generator.

7.0 Further Readings

- Devore, J. L. (2018). *Probability and statistics for engineering and the sciences*. Toronto,Ontario: Nelson.
- Georgii, H. (2013). *Stochastics: Introduction to probability and statistics*. Berlin: DeGruyter.
- Giri, N. C. (2019). *Introduction to probability and statistics*. London: Routledge.
- Johnson, R. A., Miller, I., & Freund, J. E. (2019). *Miller & Freunds probability andstatistics for engineers*. Boston: Pearson Education.
- Laha, R. G., & Rohatgi, V. K. (2020). *Probability theory*. Mineola, NY: DoverPublications.
- Mathai, A. M., & Haubold, H. J. (2018). *Probability and statistics: A course forphysicists and engineers*. Boston: De Gruyter.
- Pishro-Nik, H. (2014). *Introduction to probability, statistics, and random processes*. BlueBell, PA: Kappa Research, LLC.
- Spiegel, M. R., Schiller, J. J., & Srinivasan, R. A. (2013). *Schaums outline of probabilityand statistics*. New York: McGraw-Hill.

Unit 4: Monte Carlo Methods

Contents

1.0 Introduction

Monte Carlo methods (or Monte Carlo experiments) are a class of computational algorithms that rely on repeated random sampling to compute their results. Monte Carlo methods are often used in simulating physical and mathematical systems. The methods are especially useful in studying systems with a large number of coupled (interacting) degrees of freedom, such as fluids, disordered materials, strongly coupled solids, and cellular structures. More broadly, Monte Carlo methods are useful for modelling phenomena with significant uncertainty in inputs, such as the calculation of risk in business. These methods are also widely used in mathematics: a classic use is for the evaluation of definite integrals, particularly multidimensional integrals with complicated boundary conditions. It is a widely successful method in risk analysis when compared with alternative methods or human intuition. When Monte Carlo simulations have been applied in space exploration and oil exploration, actual observations of failures, cost overruns and schedule overruns are routinely better predicted by the simulations than by human intuition or alternative "soft" methods.

2.0 Intended Learning Outcomes (ILOs)

By the end this unit, you should be able to:

- Describe Monte Carlo method
- Trace the origin of Monte Carlo method
- Give examples of the application of Monte Carlo method

3.0 Main Content

3.1 Overview of Monte Carlo Method

There is no single Monte Carlo method; instead, the term describes a large and widely- use class of approaches.

A **Monte Carlo algorithm** is an algorithm for computers. It is used to simulate the behavior of other systems. It is not an exact method, but a heuristic one. Usually it uses randomness and statistics to get a result.

It is a computation process that uses random numbers to produce an outcome(s). Instead of having fixed inputs, probability distributions are assigned to some or all of the inputs. This will generate a probability distribution for the output after the simulation is run

However, these methods tend to follow the algorithm below:
1. Define a domain of possible inputs.
2. Generate inputs randomly from the domain using a certain specified probabilitydistribution.
3. Perform a deterministic computation using the inputs.
4. Aggregate the results of the individual computations into the final result.

For example, to approximate the value of π using a Monte Carlo method:
1. Draw a square on the ground, then inscribe a circle within it. From plane geometry, the ratio of the area of an inscribed circle to that of the surrounding square is π / 4.
2. Uniformly scatter some objects of uniform size throughout the square. For example,grains of rice or sand.
3. Since the two areas are in the ratio π / 4, the objects should fall in the areas in approximately the same ratio. Thus, counting the number of objects in the circle and dividing by the total number of objects in the square will yield an approximation for π / 4.
4. Multiplying the result by 4 will then yield an approximation for π itself.

Notice how the π approximation follows the general pattern of Monte Carlo algorithms. First,we define an input domain: in this case, it's the square which circumscribes our circle. Next, we generate inputs randomly (scatter individual grains within the square), then perform a computation on each input (test whether it falls within the circle). At the end, we aggregate the results into our final result, the approximation of π.
Note, also, two other common properties of Monte Carlo methods: the computation's relianceon good random numbers, and its slow convergence to a better approximation as more data points are sampled. If grains are purposefully dropped into only, for example, the center of the circle, they will not be uniformly distributed, and so our approximation will be poor. An approximation will also be poor if only a few grains are randomly dropped into the whole square. Thus, the approximation of π will become more accurate both as the grains are dropped more uniformly and as more are dropped.

To understand the Monte Carlo method theoretically, it is useful to think of it as a general technique of numerical integration. It can be shown, at least in a trivial sense, that every application of the Monte Carlo method can be represented as a definite integral. Suppose we need to evaluate a multi-dimensional definite integral of the form:

$$\psi = \int_0^1\int_0^1\cdots\int_0^1 f(u_1,u_2,\ldots,u_n)du_1du_2\cdots du_n = \int_{(0,1)^n} f(\boldsymbol{u})d\boldsymbol{u} \quad \textbf{........6}$$

Most integrals can be converted to this form with a suitable change of variables, so we canconsider this to be a general application suitable for the Monte Carlo method.

The integral represents a non-random problem, but the Monte Carlo method approximates a solution by introducing a random vector ***U*** that is uniformly distributed on the region of integration. Applying the function *f* to ***U***, we obtain a random variable *f*(***U***). This hasexpectation:

$$E[f(\boldsymbol{U})] = \int_{(0,1)^n} f(\boldsymbol{u})\ \phi(\boldsymbol{u})\ d\boldsymbol{u} \quad \text{..........(7)}$$

where is the probability density function ϕf ***U***. Because ϕ equals 1 on the region of integration, [7] becomes:

$$E[f(\boldsymbol{U})] = \int_{(0,1)^n} f(\boldsymbol{u})\ d\boldsymbol{u} \quad \text{..........(8)}$$

Comparing [6] and [8], we obtain a probabilistic expression for the integral Ψ:

$$\Psi = E[f(\boldsymbol{U})] \quad \text{……………..(9)}$$

so random variable $f(\boldsymbol{U})$ has mean Ψ and some standard deviation ζ. We define

$$H = f(\boldsymbol{U}^{[1]}) \quad \text{…………(10)}$$

as an unbiased estimator for Ψ with standard error ζ. This is a little unconventional, since [10] is an estimator that depends upon a sample $\{\boldsymbol{U}^{[1]}\}$ of size one, but it is a valid estimator nonetheless. To estimate Ψ with a standard error lower than ζ, let's generalize our estimator to accommodate a larger sample $\{\boldsymbol{U}^{[1]}, \boldsymbol{U}^{[2]}, \dots, \boldsymbol{U}^{[m]}\}$. Applying the function *f* to each of these yields *m* independent and identically distributed (IID) random variables $f(\boldsymbol{U}^{[1]})$, $f(\boldsymbol{U}^{[2]})$, … , $f(\boldsymbol{U}^{[m]})$, each with expectation Ψ and standard deviation ζ. The generalization of [10]

$$H = \frac{1}{m}\sum_{k=1}^{m} f(\boldsymbol{U}^{[k]}) \quad \text{…………(11)}$$

is an unbiased estimator for Ψ with standard error

$$\frac{\sigma}{\sqrt{m}} \quad \text{…………..(12)}$$

If we have a realization $\{\boldsymbol{u}^{[1]}, \boldsymbol{u}^{[2]}, \dots, \boldsymbol{u}^{[m]}\}$ for our sample, we may estimate Ψ as:

$$h = \frac{1}{m}\sum_{k=1}^{m} f(\boldsymbol{u}^{[k]}) \quad \text{…………….(13)}$$

We call [11] the **crude Monte Carlo estimator**. Formula [12] for its standard error is important for two reasons. First, it tells us that the standard error of a Monte Carlo analysis decreases with the square root of the sample size. If we quadruple the number of realizations used, we will half the standard error. Second, standard error does not depend upon the dimensionality of the integral [6]. Most techniques of numerical integration such as the trapezoidal rule or Simpson's method suffer from the **curse of dimensionality**. When generalized to multiple dimensions, the number of computations required to apply them, increases exponentially with the dimensionality of the integral. For this reason, such methods cannot be applied to integrals of more than a few dimensions. The Monte Carlo method does not suffer from the curse of dimensionality. It is as applicable to a 1000- dimensional integral as it is to a one-dimensional integral.

While increasing the sample size is one technique for reducing the standard error of a Monte Carlo analysis, doing so can be computationally expensive. A better solution is to employ some technique of variance reduction. These techniques incorporate additional information about the analysis directly into the estimator. This allows them to make the Monte Carlo estimator more deterministic, and hence have a lower standard error.

Due to high mathematics required and burden of understanding at this level, we have to stop this discussion here.

3.2 History of Monte Carlo Method

Physicists at Los Alamos Scientific Laboratory were investigating radiation shielding and the distance that neutrons would likely travel through various materials. Despite having most of the necessary data, such as the average distance a neutron would travel in a substance before it collided with an atomic nucleus or how much energy the neutron was likely to give off following a collision, the problem could not be solved with analytical calculations. John von Neumann and Stanislaw Ulam suggested that the problem be solvedby modelling the experiment on a computer using chance. Being secret, their work required a code name. Von Neumann chose the name "Monte Carlo". The name is a reference to the Monte Carlo Casino in Monaco where Ulam's uncle would borrow money to gamble.

Random methods of computation and experimentation (generally considered forms of stochastic simulation) can be arguably traced back to the earliest pioneers of probability theory but are more specifically traced to the pre-electronic computing era. The general difference usually described about a Monte Carlo form of simulation is that it systematically "inverts" the typical mode of simulation, treating deterministic problems by *first* finding a probabilistic analogy. Previous methods of simulation and statistical sampling generally did the opposite: using simulation to test a previously understood deterministic problem. Though examples of an "inverted" approach do exist historically, they were not considered a general method until the popularity of the Monte Carlo methodspread.

It was only after electronic computers were first built (from 1945 on) that Monte Carlo methods began to be studied in depth. In the 1950s they were used at Los Alamos for early work relating to the development of the hydrogen bomb, and became popularized in the fields of physics, physical chemistry, and operations research. The Rand Corporation and the U.S. Air Force were two of the major organizations responsible for funding anddisseminating information on Monte Carlo methods during this time, and they began to find a wide application in many different fields.
Uses of Monte Carlo methods require large amounts of random numbers, and it was their use that spurred the development of pseudorandom number generators, which were far quicker to use than the tables of random numbers which had been previously used for statistical sampling.

3.4 Applications of Monte Carlo Methods

As stated above, Monte Carlo simulation methods are especially useful for modelling phenomena with significant uncertainty in inputs and in studying systems with a large number of coupled degrees of freedom. Areas of application include:

a. Physical sciences:

Monte Carlo methods are very important in computational physics, physical chemistry, andrelated applied fields, and have diverse applications from complicated quantum calculations to designing heat shields and aerodynamic forms. The Monte Carlo method is widely used in statistical physics, particularly Monte Carlo molecular modelling as an alternative for computational molecular dynamics as well as to compute statistical field theories of simple particle and polymer models. In experimental particle physics, these methods are used for designing detectors, understanding their behaviour and comparing experimental data to theory, or on vastly large scale of the galaxy modelling.

Monte Carlo methods are also used in the models that form the basis of modern weatherforecasting operations.

b. Engineering

Monte Carlo methods are widely used in engineering for sensitivity analysis and quantitative probabilistic analysis in process design. The need arises from the interactive, co-linear and non-linear behaviour of typical process simulations. For example,

- in **microelectronics** engineering, Monte Carlo methods are applied to analyze correlated and uncorrelated variations in analog and digital integrated circuits. Thisenables designers to estimate realistic 3 sigma corners and effectively optimise circuit yields.

- in **geostatistics and geometallurgy**, Monte Carlo methods strengthen the design ofmineral processing flow sheets and contribute to quantitative risk analysis.

c. Visual Designs

Monte Carlo methods have also proven efficient in solving coupled integral differential equations of radiation fields and energy transport, and thus these methods have been usedin global illumination computations which produce photorealistic images of virtual 3D models, with applications in video games, architecture, design and computer generated films.

d. Finance and business

Monte Carlo methods in finance are often used to calculate the value of companies, to evaluate investments in projects at a business unit or corporate level, or to evaluatefinancial derivatives. Monte Carlo methods used in these cases allow the construction of stochastic or probabilistic financial models as opposed to the traditional static and deterministic models, thereby enhancing the treatment of uncertainty in the calculation.

e. Telecommunications

When planning a wireless network, design must be proved to work for a wide variety of scenarios that depend mainly on the number of users, their locations and the services they want to use. Monte Carlo methods are typically used to generate these users and theirstates. The network performance is then evaluated and, if results are not satisfactory, the network design goes through an optimization process.

f. Games

Monte Carlo methods have recently been applied in game playing related artificial intelligence theory. Most notably the game of Battleship have seen remarkably successful Monte Carlo algorithm based computer players. One of the main problems that this approach has in game playing is that it sometimes misses an isolated, very good move. These approaches are often strong strategically but weak tactically, as tactical decisions tend to rely on a small number of crucial moves which are easily missed by the randomly searching Monte Carlo algorithm.

Monte Carlo simulation versus "what if" scenarios

The opposite of Monte Carlo simulation might be considered deterministic modelling using single-point estimates. Each uncertain variable within a model is assigned a ―best guess‖ estimate. Various combinations of each input variable are manually chosen (such as best case, worst case, and most likely case), and the results recorded for each so-called
―what if‖ scenario.

By contrast, Monte Carlo simulation considers random sampling of probability distributionfunctions as model inputs to produce hundreds or thousands of possible outcomes instead of a few discrete scenarios. The results provide probabilities of different outcomes occurring.

For example, a comparison of a spreadsheet cost construction model run using traditional
―what if‖ scenarios, and then run again with Monte Carlo simulation and Triangular probability distributions shows that the Monte Carlo analysis has a narrower range thanthe ―what if‖ analysis. This is because the ―what if‖ analysis gives equal weight to all scenarios.

A **randomized algorithm** or **probabilistic algorithm** is an algorithm which employs a degree of randomness as part of its logic. The algorithm typically uses uniformly distributed random bits as an auxiliary input to guide its behaviour, in the hope of achieving good performance in the "average case" over all possible choices of random bits.Formally, the algorithm's performance will be a random variable determined by the random bits; thus either the running time, or the output (or both) are random variables

In common practice, randomized algorithms are approximated using a pseudorandomnumber generator in place of a true source of random bits; such an implementation may deviate from the expected theoretical behaviour.

g. Uses in mathematics

In general, Monte Carlo methods are used in mathematics to solve various problems by generating suitable random numbers and observing that fraction of the numbers which obeys some property or properties. The method is useful for obtaining numerical solutions to problems which are too complicated to solve analytically. The most common application of the Monte Carlo method in mathematics are:

i. Integration

Deterministic methods of numerical integration usually operate by taking a number of evenly spaced samples from a function. In general, this works very well for functions of one variable. However, for functions of vectors, deterministic quadrature methods can be very inefficient. To numerically integrate a function of a two-dimensional vector, equally spaced grid points over a two-dimensional surface are required. For instance a 10x10 grid requires 100 points. If the vector has 100 dimensions, the same spacing on the grid would require 10^{100} points which is far too many to be computed. But 100 dimensions is by no means unusual, since in many physical problems, a "dimension" is equivalent to a degree of freedom.

Monte Carlo methods provide a way out of this *exponential time-increase*. As long as the function in question is reasonably well-behaved, it can be estimated by randomly selecting points in 100-dimensional space, and taking some kind of average of the function values at these points. By the law of large numbers, this method will display convergence (i.e. quadrupling the number of sampled points will halve the error, regardless of the number of dimensions).

ii. Optimization

Most Monte Carlo optimization methods are based on **random walks**. The program will move around a marker in multi-dimensional space, tending to move in directions which lead to a lower function, but sometimes moving against the gradient.

Another popular application for random numbers in numerical simulation is in numerical optimization (choosing the best element from some set of available alternatives). These problems use functions of some often large-dimensional vector that are to be minimized(or maximized). Many problems can be phrased in this way: for example a computer chess program could be seen as trying to find the optimal set of, say, 10 moves which produces the best evaluation function at the end. The travelling salesman problem is another optimization problem. There are also applications to engineering design, such as design optimization.

iii. Inverse problems

Probabilistic formulation of inverse problems leads to the definition of a probability distribution in the space models. This probability distribution combines ***a priori*** (prior knowledge about a population, rather than that estimated by recent observation) information with new information obtained by measuring some observable parameters (data). As, in the general case, the theory linking data with model parameters is nonlinear, the **aposteriori** probability in the model space may not be easy to describe (it may be multimodal, some moments may not be defined, etc.).

When analyzing an inverse problem, obtaining a maximum likelihood model is usually not sufficient, as we normally also wish to have information on the resolution power of the data. In the general case we may have a large number of model parameters, and an inspection of the marginal probability densities of interest may be impractical, or even useless. But it is possible to pseudorandomly generate a large collection of models according to the posterior probability distribution and to analyze and display the models insuch a way that information on the relative likelihoods of model properties is conveyed to the spectator. This can be accomplished by means of an efficient Monte Carlo method, even in cases where no explicit formula for the a priori distribution is available.

h. Computational mathematics

Monte Carlo methods are useful in many areas of computational mathematics, where a *lucky choice* can find the correct result. A classic example is Rabin's algorithm for primality testing (algorithm which determines whether a given number is prime). It states that for any *n* which is not prime, a random *x* has at least a 75% chance of proving that *n* is not prime. Hence, if *n* is not prime, but *x* says that it might be, we have observed at most a 1-in-4 event. If 10 different random *x* say that "*n* is probably prime" when it is not, we have observed a one-in-a-million event. In general a Monte Carlo algorithm of this kind produces one correct answer with a guarantee that ***n* is composite, and *x* proves it so**, but another one without, but with a guarantee of not getting this answer when it is wrong too often; in this case at most 25% of the time.

Remark:

In physics, two systems are **coupled** if they are interacting with each other. Of special interest is the **coupling** of two (or more) vibratory systems (e.g. pendula or resonant circuits) by means of springs or magnetic fields, etc. Characteristic for a coupled oscillation is the effect of beat.

3.5 Monte Carlo Simulation

Monte Carlo simulation is a computerized mathematical technique to generate random sample data based on some known distribution for numerical experiments. This method is applied to risk quantitative analysis and decision making problems. This method is used by the professionals of various profiles such as finance, project management, energy, manufacturing, engineering, research & development, insurance, oil & gas, transportation, etc.

This method was first used by scientists working on the atom bomb in 1940. This method can be used in those situations where we need to make an estimate and uncertain decisions such as weather forecast predictions.

Monte Carlo Simulation ─ Important Characteristics

Following are the three important characteristics of Monte-Carlo method −

- Its output must generate random samples.
- Its input distribution must be known.
- Its result must be known while performing an experiment.

Monte Carlo Simulation ─ Advantages

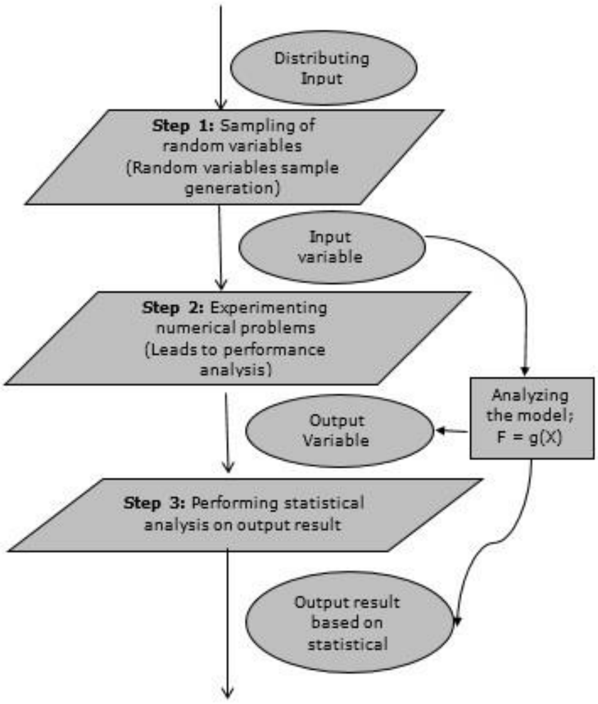
- Easy to implement.
- Provides statistical sampling for numerical experiments using the computer.
- Provides approximate solution to mathematical problems.
- Can be used for both stochastic and deterministic problems.

Monte Carlo Simulation ─ Disadvantages

- Time consuming as there is a need to generate large number of sampling to get the desired output.
- The results of this method are only the approximation of true values, not the exact.

Monte Carlo Simulation Method ─ Flow Diagram

The following illustration shows a generalized flowchart of Monte Carlo simulation.



4.0 Self-Assessment Exercise(s)

Answer the following questions:

1. How is Monte Carlo method different in approach from the typical mode ofsimulation, in deterministic problems?
2. How is Monte Carlo method used in Engineering and Mathematics?

5.0 Conclusion

Monte Carlo methods, relies on repeated computation of random or pseudo-random numbers. These methods are most suited to computations by a computer and tend to be used when it is unfeasible or impossible to compute an exact result with a deterministic algorithm (i.e. an algorithm whose behaviour can be completely predicted from the input)

6.0 Summary

In this unit we discussed the following:

- The algorithm of Monte Carlo method
- The history of Monte Carlo method which spurred the development of pseudorandomnumber

generator

- The application of Monte Carlo methods in areas such as physical sciences, Engineering, Finance and Business, telecommunications, Games, Mathematics, etc.

7.0 Further Readings

- Devore, J. L. (2018). *Probability and statistics for engineering and the sciences*. Toronto,Ontario: Nelson.
- Georgii, H. (2013). *Stochastics: Introduction to probability and statistics*. Berlin: De Gruyter.
- Giri, N. C. (2019). *Introduction to probability and statistics*. London: Routledge.
- Johnson, R. A., Miller, I., & Freund, J. E. (2019). *Miller & Freunds probability and statistics for engineers*. Boston: Pearson Education.
- Laha, R. G., & Rohatgi, V. K. (2020). *Probability theory*. Mineola, NY: Dover Publications.
- Mathai, A. M., & Haubold, H. J. (2018). *Probability and statistics: A course for physicists and engineers*. Boston: De Gruyter.
- Pishro-Nik, H. (2014). *Introduction to probability, statistics, and random processes*. BlueBell, PA: Kappa Research, LLC.
- Spiegel, M. R., Schiller, J. J., & Srinivasan, R. A. (2013). *Schaums outline of probabilityand statistics*. New York: McGraw-Hill.

Unit 5: Statistical Distribution Functions

Contents

1.0 Introduction

Although simulation can be a valuable tool for better understanding the underlying mechanisms that control the behaviour of a system, using simulation to make ***predictions***of the future behaviour of a system can be difficult. This is because, for most real-world systems, at least some of the controlling parameters, processes and events are often stochastic, uncertain and/or poorly understood. The objective of many simulations is to identify and quantify the risks associated with a particular option, plan or design. Simulating a system in the face of such uncertainty and computing such risks requires that the uncertainties be quantitatively included in the calculations. To do this we collect data about the system parameters and subject them to statistical analysis.

2.0 Intended Learning Outcomes (ILOs)

After studying this unit the reader should be able to

- Define Statistics
- Explain Statistical Distributions
- Compute measures of Central Tendency and Variations
- Explain the Components of Statistical Distributions
  - Normal Distributions,
  - z-score
  - percentile,
  - Skewed Distributions

- Ways to transform data to Graphs

3.0 Main Content

3.1 What is Statistics?

The field of statistics is concerned with the collection, description, and interpretation ofdata (data are numbers obtained through measurement). In the field of statistics, the term –statistic‖ denotes a measurement taken on a sample (as opposed to a population). In general conversation, –statistics‖ also refers to facts and figures.

3.2 What is a Statistical Distribution?

A statistical distribution describes the numbers of times each possible outcome occurs in a sample. If you have 10 test scores with 5 possible outcomes of A, B, C, D, or F, a statistical distribution describes the relative number of times an A,B,C,D or F occurs. For example, 2 A‘s, 4 B‘s, 4 C‘s, 0 D‘s, 0 F‘s.

3.3 Measures of Central Tendency

Suppose we have a sample with the following 4 observations: 4, 1, 4, 3.

Mean - the sum of a set of numbers divided by the number of observations.

$$\text{Mean} = \frac{4+1+4+3}{4} = \frac{12}{4} = 3$$

Median - the middle point of a set of numbers (for odd numberedsamples). the mean of the middle two points (for even samples).

$$\text{Median}=1,\underline{3},\underline{4},4 \text{ or } \frac{3+4}{2}=\frac{7}{2}=3.5$$

Mode - the most frequently occurring number.

Mode = 4 (4 occurs most).

The mean, median and mode are called measures of central tendency.

3.4 Measures of Variation

Range - the maximum value minus the minimum value in a set ofnumbers. Range = 4-1 = 3.

Standard Deviation - the average distance a data point is away from the mean.

$$\text{standard deviation} = \frac{|4-3|+|1-3|+|4-3|+|3-3|}{4} = \frac{1+2+1+0}{4} = \frac{4}{4} = 1$$

Standard deviation computes the difference between each data point and the mean. Take the absolute value of each difference. Sum the absolute values. Divide this sum by the number of data points. Median: first arrange data points in increasing order.

Mean, Median, Mode, Range, and Standard Deviations are measurements in a sample (statistics) and can also be used to make inferences on a population.

3.5 Showing Data Distribution in Graphs

- **Bar graphs** use bars to compare frequencies of possible data values (see Fig a).
- **Double bar graphs** use two sets of bars to compare frequencies of data valuesbetween two levels of data (e.g. boys and girls) (see fig b).

- **Histograms** use bars to show how frequently data occur within equal spaces withinan interval (see fig c & d).
- **Pie Charts** use portion of a circle to show contributions of data values (see fig c & d).

3.6 The Difference between a Continuous and a Discrete Distribution

Continuous distributions describe an infinite number of possible data values (as shownby the curve). For example someone‘s height could be 1.7m, 1.705m, 1.71m, ...

Discrete distributions describe a finite number of possible values. (shown by the bars)

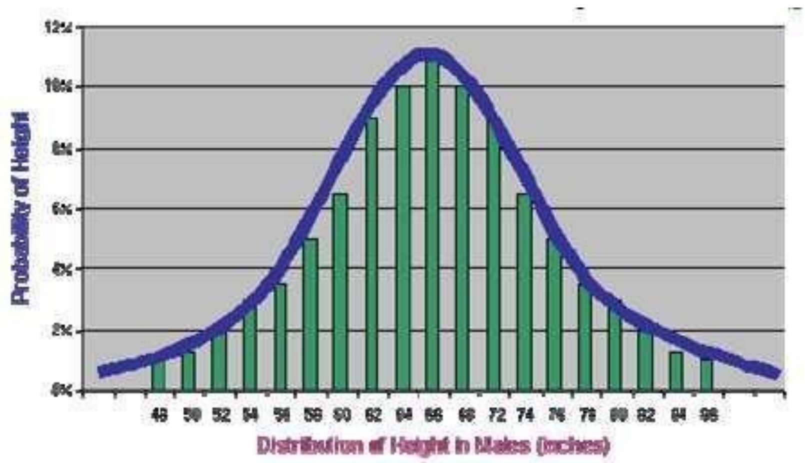


Fig 2: Distribution of Height in Males

3.7 Normal Distribution

A normal distribution is a continuous distribution that is ―bell-shaped‖. Data are often assumed to be normal. Normal distributions can estimate probabilities over a continuous interval of data values.
The **normal distribution** refers to a family of continuous probability distributions described by the normal equation.

In a normal distribution, data are most likely to be at the mean. Data are less likely to be farther away from the mean.
The normal distribution is defined by the following equation:

$Y = [\ 1/\zeta * \text{sqrt}(2\pi)\ ] * e^{-(x\ -\ \mu)2/2\zeta 2}$

where *X* is a normal random variable, μ is the mean, ζ is the standard deviation, π is approximately 3.14159, and *e* is approximately 2.71828.
The random variable *X* in the normal equation is called the **normal random variable**. The normal equation is the probability density function for the normal distribution.

The graph of the normal distribution depends on two factors - the mean and the standard deviation. The mean of the distribution determines the location of the center of the graph, and the standard deviation determines the height and width of the graph. When the standard deviation is large, the curve is short and wide; when the standard deviation is small, the curve is tall and narrow. All normal distributions look like a symmetric, bell- shaped curve, as shown in figure 3a and 3b.

(a) (b)

Fig. 3: Graph of Normal Distribution Based on Size of Mean and Standard Deviation

The curve on the left is shorter and wider than the curve on the right, because the curve onthe left has a bigger standard deviation.

3.7.1 Standard Normal Distribution

The **standard normal distribution** is a special case of the normal distribution. It is the distribution that occurs when a normal random variable has a mean of zero and a standarddeviation of one.

The normal random variable of a standard normal distribution is called a **standard score**or a **z-score**. Every normal random variable X can be transformed into a z score via the following equation:

$$z = (X - \mu) / \zeta$$

where X is a normal random variable, μ is the mean of X, and ζ is the standard deviation of X.

3.7.2 The Normal Distribution as a Model for Measurements

Often, phenomena in the real world follow a normal (or near-normal) distribution. This allows researchers to use the normal distribution as a model for assessing probabilities associated with real-world phenomena. Typically, the analysis involves two steps.

- Transform raw data. Usually, the raw data are not in the form of z-scores. They need to be transformed into z-scores, using the transformation equation presentedearlier: $z = (X - \mu) / \zeta$.
- Find the probability. Once the data have been transformed into z-scores, you canuse standard normal distribution tables, online calculators (e.g., Stat Trek's free normal distribution calculator) to find probabilities associated with the z-scores.

  The problem in the next section demonstrates the use of the normal distribution as a modelfor measurement.

  Example 1 - Ada earned a score of 940 on a national achievement test. The mean test score was 850 with a standard deviation of 100. What proportion of students had a higherscore than Ada? (Assume that test scores are normally distributed.)

  Solution - As part of the solution to this problem, we assume that test scores are normallydistributed. In this way, we use the normal distribution as a model for measurement.

  Given an assumption of normality, the solution involves three steps.

- First, we transform Ada's test score into a z-score, using the z-scoretransformationequation.

  $z = (X - \mu) / \zeta = (940 - 850) / 100 = 0.90$
- Then, using a standard normal distribution table, we find the cumulative probability associated with the z-score. In this case, we find $P(Z < 0.90) = 0.8159$.

- Therefore, the P(Z > 0.90) = 1 - P(Z < 0.90) = 1 - 0.8159 = 0.1841.
  Thus, we estimate that 18.41 percent of the students tested had a higher score than Ada.

  Example 2 - An average light bulb manufactured by the Acme Corporation lasts 300 days with a standard deviation of 50 days. Assuming that bulb life is normally distributed, what is the probability that an Acme light bulb will last at most 365 days?

Solution: Given a mean score of 300 days and a standard deviation of 50 days, we want tofind the cumulative probability that bulb life is less than or equal to 365 days. Thus, we know the following:

- The value of the normal random variable is 365 days.
- The mean is equal to 300 days.
- The standard deviation is equal to 50 days.

We enter these values into the formula and compute the cumulative probability. The answer is: $P(X \leq 365) = 0.90$. Hence, there is a 90% chance that a light bulb will burn outwithin 365 days.

3.7.3 Conversion to a Standard Normal Distribution

The values for points in a standard normal distribution are **z-scores**. We can use a standardnormal table to find the probability of getting at or below a z-score. (a percentile).

- Subtract the mean from each observation in your normal distribution, thenew mean=0.
- Divide each observation by the standard deviation, the new standarddeviation=1.

3.7.4 Skewed Distributions µ

Skewness is the degree of asymmetry or departure from symmetry, of a distribution. Skewed distributions are not symmetric. If the frequency curve of a distribution has a longer tail to the right of the right of the central maximum than to the left, the distributionis said to be skewed to the right, or have a positive skewness. If the reverse is the case, it is said to be skewed to the left or negative skewness.

For skewed distributions, the mean tend to lie on the same side of the mode as the longer tail. Thus a measure of the asymmetry is supplied by the difference:

Mean – mode. This can be made dimensionless if we divide it by a measure of dispersion, such as the standard deviation, leading to the definition:

$$\text{Skewness} = \frac{mean - \text{mod}\, e}{SD} = \frac{\mu - \text{mod}\, e}{s} \quad (1)$$

To avoid using mode, we can use the empirical formula:

$$\text{Skewness} = \frac{3(mean - \text{median})}{SD} = \frac{3(\mu - \text{median})}{s} \quad (2)$$

Equations (1) and (2) are called; Pearson‘s first and second coefficients of skewness.

3.8 What is a Percentile?

A percentile (or **cumulative probability**) is the proportion of data in a distribution less than or equal to a data point. If you scored a 90 on a math test and 80% of the class had scores of 90 or lower; your percentile is 80. In the figure 4, b=90 and P(Z<b)=80.

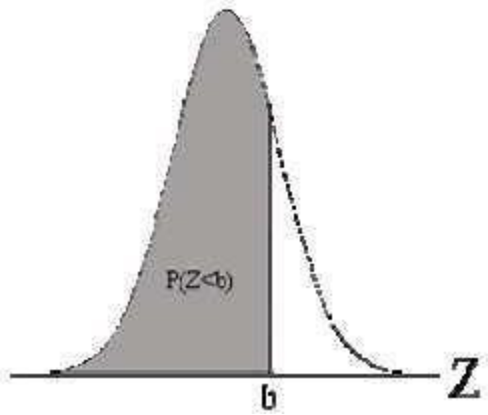


Fig. 4: illustration of Percentiles

3.9 Probabilities in Discrete Distributions

Suppose for your 10 tests you received 5 As, 2 Bs, 2 Cs, 1 D and want to find the probability of receiving an A or a B. Sum the frequencies for A and B and divide by the sample size. The probability of receiving an A or a B is (5+2)/10 = .7 (a 70% chance).

3.10 Probability and the Normal Curve

The normal distribution is a continuous probability distribution. This has severalimplications for probability.

- The total area under the normal curve is equal to 1.
- The probability that a normal random variable *X* equals any particular value is 0.
- The probability that *X* is greater than b equals the area under the normal curve bounded by *b* and plus infinity (as indicated by the *non-shaded* area in the figure 4).
- The probability that *X* is less than *a* equals the area under the normal curve boundedby b and minus infinity (as indicated by the *shaded* area in the figure below).

4.0 Self-Assessment Exercise(s)

Answer the following questions:

4.0

1. Why Convert to a Standard Normal Distribution?
2. What is the difference between a Continuous and a Discrete Distribution?
3. Given the following: mean=279.76, median=279.06, mode=277.5 and SD=15.6, findthe first and second coefficients of skewness
4. Find the mode, median and mean deviation of the following sets of data: (a) 3, 7, 9, 5and (b) 8, 10, 9, 12, 4, 8, 2.

5.0 Conclusion

We use Statistical distributions to: investigate how a change in one variable relates to a change in a second variable, represent situations with numbers, tables, graphs, and verbal descriptions, understand measurable attributes of objects and their units, systems, andprocesses of measurement, identify relationships among attributes of entities or systemsand their association.

6.0 Summary

In this unit:

- We defined Statistics as field of study that is concerned with the collection,description, and interpretation of data.
- We saw that Statistical Distributions describe the numbers of times each possibleoutcome occurs in a sample.
- We computed various measures of Central Tendency and Variations which can beused to make inferences.
- And explained the following components of Statistical Distributions:
  - Normal Distributions,
  - z-score
  - percentile,
  - Skewed Distributions
  - Ways to transform data to Graphs

7.0 Further Readings

- Devore, J. L. (2018). *Probability and statistics for engineering and the sciences*. Toronto, Ontario: Nelson.
- Georgii, H. (2013). *Stochastics: Introduction to probability and statistics*. Berlin: DeGruyter.
- Giri, N. C. (2019). *Introduction to probability and statistics*. London: Routledge.
- Johnson, R. A., Miller, I., & Freund, J. E. (2019). *Miller & Freunds probability andstatistics for engineers*. Boston: Pearson Education.
- Laha, R. G., & Rohatgi, V. K. (2020). *Probability theory*. Mineola, NY: DoverPublications.
- Mathai, A. M., & Haubold, H. J. (2018). *Probability and statistics: A course forphysicists and engineers*. Boston: De Gruyter.
- Pishro-Nik, H. (2014). *Introduction to probability, statistics, and random processes*. BlueBell, PA: Kappa Research, LLC.
- Spiegel, M. R., Schiller, J. J., & Srinivasan, R. A. (2013). *Schaums outline of probabilityand statistics*. New York: McGraw-Hill.

Unit 6: Common Probability Distributions

Contents

1.0 Introduction

In this section we look at the branch of statistics that deals with analysis of random events. Probability is the numerical assessment of likelihood on a scale from 0 (impossibility) to 1 (absolute certainty). Probability is usually expressed as the ratio between the number of waysan event can happen and the total number of things that can happen (e.g., there are 13 waysof picking a diamond from a deck of 52 cards, so the probability of picking a diamond is 13/52, or ¼). Probability theory grew out of attempts to understand card games and gambling. As science became more rigorous, analogies between certain biological, physical, and social phenomena and games of chance became more evident (e.g., the sexes of newborninfants follow sequences similar to those of coin tosses). As a result, probability became a fundamental tool of modern genetics and many other disciplines.

2.0 Intended Learning Outcomes (ILOs)

By the end of this unit, the reader should be able to:

- Explain the role of probability distribution functions in simulations
- Describe Probability theory
- Explain the fundamental concepts of Probability theory
- Explain Random Variable
- Explain Limiting theorems
- Describe Probability distributions in simulations

- List common Probability distributions.

3.0 Main Content

3.1 Distribution Functions and Simulation

Many simulation tools and approaches are ***deterministic***. In a deterministic simulation, the input parameters for a model are represented using single values (which typically are described either as "the best guess" or "worst case" values). Unfortunately, this kind of simulation, while it may provide some insight into the underlying mechanisms, is not well- suited to making predictions to support decision-making, as it cannot quantitatively address the risks and uncertainties that are inherently present.

However, it is possible to quantitatively represent uncertainties in simulations. ***Probabilistic simulation*** is the process of explicitly representing these uncertainties by specifying inputs as probability distributions. If the inputs describing a system are uncertain, the prediction of future performance is necessarily uncertain. That is, the result of any analysis based on inputs represented by probability distributions is itself a probability distribution. Hence, whereas the result of a deterministic simulation of an uncertain system is a *qualified statement* ("if we build the dam, the salmon population could go extinct"), the result of a probabilistic simulation of such a system is a *quantified probability* ("if we build the dam, there is a 20% chance that the salmon population will go extinct"). Such a result (in this case, quantifying the risk of extinction) is typically much more useful to decision-makerswho might utilize the simulation results.

3.2 Probability Definitions

The word *probability* does not have a consistent direct definition. In fact, there are two broadcategories of **probability interpretations**, whose adherents possess different (and sometimesconflicting) views about the fundamental nature of probability:

1. Frequentists talk about probabilities only when dealing with experiments that are random and well-defined. The probability of a random event denotes the *relative frequency of occurrence* of an experiment's outcome, when repeating the experiment.Frequentists consider probability to be the relative frequency "in the long run" of outcomes.
2. Bayesians, however, assign probabilities to any statement whatsoever, even when norandom process is involved. Probability, for a Bayesian, is a way to represent an individual's *degree of belief* in a statement, or an objective degree of rational belief, given the evidence

   The scientific study of probability is a modern development. Gambling shows thatthere has been an interest in quantifying the ideas of probability for millennia, butexact mathematical descriptions of use in those problems only arose much later.

Probability Distribution

A probability distribution gathers together all possible outcomes of a random variable (i.e. any quantity for which more than one value is possible), and summarizes these outcomes by indicating the probability of each of them. While a probability distribution is often associatedwith the bell-shaped curve, recognize that such a curve is only indicative of one specific type of probability, the so-called normal probability distribution. However, in real life, a probability distribution can take any shape, size and form.

Example: Probability Distribution

For example, if we wanted to choose a day at random in the future to schedule an event, and we wanted to know the probability that this day would fall on a Sunday, as we will need to avoid scheduling it on a

Sunday. With seven days in a week, the probability that a random day would happen to be a Sunday would be given by one-seventh or about 14.29%. Of course, the same 14.29% probability would be true for any of the other six days.

In this case, we would have a uniform probability distribution: the chances that our random day would fall on any particular day are the same, and the graph of our probability distribution would be a straight line.

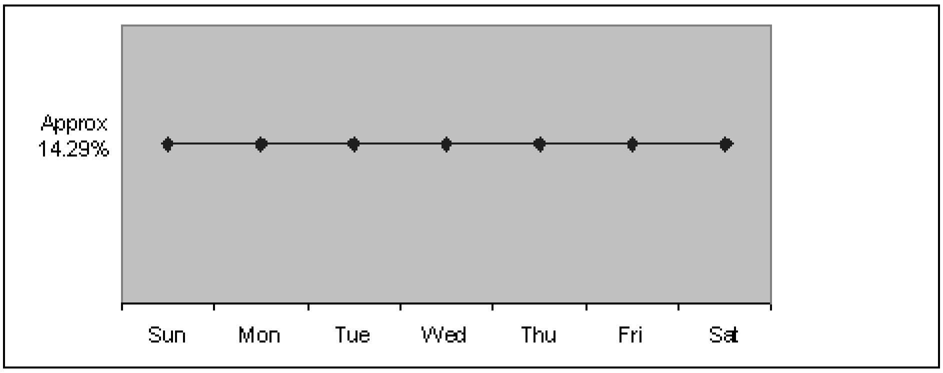


fig.1: Uniform Probability Distribution

Probability distributions can be simple to understand as in this example, or they can bevery complex and require sophisticated techniques (e.g., option pricing models, Monte Carlo simulations) to help describe all possible outcomes.

3.3 Random Variables

Random variable is **discrete** random variables if it can take on a finite or countable numberof possible outcomes. The previous example asking for a day of the week is an example of a discrete variable, since it can only take seven possible values. Monetary variables expressed in dollars and cents are always discrete, since money is rounded to the nearest
$0.01.

A random variable is **continuous** random variable if it has has infinite possible outcomes.A rate of return (e.g. growth rate) is continuous:

- a stock can grow by 9% next year or by 10%, and in between this range it couldgrow by 9.3%, 9.4%, 9.5%
- Clearly there is no end to how precise the outcomes could be broken down; thus it‘sdescribed as a continuous variable.

Outcomes in Discrete vs. Continuous Variables

The rule of thumb is that a discrete variable can have all possibilities listed out, while a continuous variable must be expressed in terms of its upper and lower limits, and greater- than or less-than indicators. Of course, listing out a large set of possible outcomes (whichis usually the case for money variables) is usually impractical – thus money variables will usually have outcomes expressed as if they were continuous.

Examples:

- Rates of return can theoretically range from –100% to positive infinity.
- Time is bound on the lower side by 0.

- Market price of a security will also have a lower limit of $0, while its upper limit willdepend on the security – stocks have no upper limit (thus a stock price‘s outcome ≥ $0),
- Bond prices are more complicated, bound by factors such as time-to-maturity andembedded call options. If a face value of a bond is $1,000, there‘s an upper limit (somewhere above $1,000) above which the price of the bond will not go, but pinpointing the upper value of that set is imprecise.

3.4 Probability Function

A probability function gives the probabilities that a random variable will take on a given list of specific values. For a discrete variable, if (x_1, x_2, x_3, x_4 …) are the complete set of possible outcomes, p(x) indicates the chances that X will be equal to x. Each x in the listfor a discrete variable will have a p(x). For a continuous variable, a probability function is expressed as f(x).

The two key properties of a probability function, p(x) (or f(x) for continuous), are thefollowing:

1. $0 \leq p(x) \leq 1$, since probability must always be between 0 and 1.
2. Add up all probabilities of all distinct possible outcomes of a random variable, and thesum must equal 1.

Determining whether a function satisfies the first property should be easy to spot since we know that probabilities always lie between 0 and 1. In other words, p(x) could never be 1.4 or –0.2. To illustrate the second property, say we are given a set of three possibilities for X: (1, 2, 3) and a set of three for Y: (6, 7, 8), and given the probability functions f(x) and g(y).

x	f(x)	y	g(y)
1	0.31	6	0.32
2	0.43	7	0.40
3	0.26	8	0.23

For all possibilities of f(x), the sum is 0.31+0.43+0.26=1, so we know it is a validprobability function. For all possibilities of g(y), the sum is 0.32+0.40+0.23 = 0.95, which violates our second principle. Either the given probabilities for g(y) are wrong, or there is afourth possibility for y where g(y) = 0.05. Either way it needs to sum to 1.

Probability Density Function

A probability density function (or pdf) describes a probability function in the case of a continuous random variable. Also known as simply the ―density‖, a probability density function is denoted by ―f(x)‖. Since a pdf refers to a continuous random variable, its probabilities would be expressed as ranges of variables rather than probabilities assigned toindividual values as is done for a discrete variable. For example, if a stock has a 20% chance of a negative return, the pdf in its simplest terms could be expressed as:

x	f(x)
< 0	0.2
≥ 0	0.8

3.5 Mathematical Treatment of Probability

In mathematics, a probability of an event *A* is represented by a real number in the range from 0 to 1 and written as P(*A*), p(*A*) or Pr(*A*). An impossible event has a probability of 0, and a certain event has a probability of 1. However, the converses are not always true: probability 0 events are not always impossible, nor probability 1 events certain. The rather subtle distinction between "certain" and "probability 1" is treated at greater length in the article on "almost surely".

The *opposite* or *complement* of an event *A* is the event [not *A*] (that is, the event of *A* notoccurring); its probability is given by P(not *A*) = 1 - P(*A*). As an example, the chance of not rolling a six on a six-sided die is 1 – (chance of rolling a six) $1 - \frac{1}{6} = \frac{5}{6}$.

Joint Probabilty

If both the events *A* and *B* occur on a single performance of an experiment this is called the ***intersection or joint probability*** of *A* and *B*, denoted as and $P(A \cap B)$.

If two events, *A* and *B* are ***independent*** then the joint probability is:

$$P(A \text{ and } B) = P(A \cap B) = P(A)P(B),$$

for example, if two coins are flipped the chance of both being heads is: $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$.Mutually Exclusive Events

If either event *A* or event *B* or both events occur on a single performance of an experiment this is called the union of the events *A* and *B* denoted as $P(A \cup B)$.

If two events are ***mutually exclusive*** then the probability of either occurring is:

$$P(A \text{ or } B) = P(A \cup B) = P(A) + P(B) - P(A \cap B) = P(A) + P(B) - 0 = P(A) + P(B)$$

.

For example, the chance of rolling a 1 or 2 on a six-sided die is

$$P(1 \text{ or } 2) = P(1) + P(2) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}.$$

If the events are not mutually exclusive then

$$P(A \text{ or } B) = P(A \cup B) = P(A) + P(B) - P(A \text{ and } B).$$

For example, when drawing a single card at random from a regular deck of cards, the chance of getting a heart or a face card (J,Q,K) (or one that is both) is $\frac{13}{52} + \frac{12}{52} - \frac{3}{52} = \frac{11}{26}$, because of the 52 cards of a deck 13 are hearts, 12 are face cards, and 3 are both: here the possibilitiesincluded in the "3 that are both" are included in each of the "13 hearts" and the "12 face cards" but should only be counted once.

Conditional Probability

This is the probability of some event *A*, given the occurrence of some other event *B*. Conditional probability is written *P*(*A*|*B*), and is read "the probability of *A*, given *B*". It is defined by:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}.$$

If *P*(*B*) = 0 then is undefined.

$P(A \mid B)$

Summary of probabilities

Event	Probability
A	P(A) □ [0,1]
Not A	P(A) = 1-P(A)
A or B	P(A U B) = P(A) + P(B)-P(A∩B) = P(A)+P(B) If A and B are mutually exclusive
A and B	P(A∩B) = P(A|B)P(B) = P(A)P(B) If A and B are independent
A given B	P(A|B) = P(A∩B)/P(B)

Two or more events are mutually exclusive if the occurrence of any one of them excludes the occurrence of the others.

3.6 Probability theory

Like other theories, the theory of probability is a representation of probabilistic concepts informal terms—that is, in terms that can be considered separately from their meaning. These formal terms are manipulated by the rules of mathematics and logic, and any results are then interpreted or translated back into the problem domain.

There have been at least two successful attempts to formalize probability, namely the Kolmogorov formulation and the Cox formulation. In Kolmogorov's formulation sets are interpreted as events and probability itself as a measure on a class of sets. In Cox's theorem, probability is taken as a primitive (that is, not further analyzed) and the emphasis is on constructing a consistent assignment of probability values to propositions. In both cases, the laws of probability are the same, except for technical details.

Probability theory is a mathematical science that permits one to find, using the probabilities of some random events, the probabilities of other random events connected in some way with the first.
The assertion that a certain event occurs with a probability equal, for example, to 1/2, isstill not, in itself, of ultimate value, because we are striving for definite knowledge. Of definitive, cognitive value are those results of probability theory that allow us to state that the probability of occurrence of some event *A* is very close to 1 or (which is the same thing) that the probability of the non-occurrence of event *A* is very small. According to the principle of ―disregarding sufficiently small probabilities,‖ such an event is considered practically reliable. Such conclusions, which are of scientific and practical interest, are usually based on the assumption that the occurrence or non-occurrence of event *A* dependson a large number of factors that are slightly connected with each other.

Consequently, it can also be said that **probability theory** is a mathematical science that clarifies the regularities that arise in the interaction of a large number of random factors.
To describe the regular connection between certain conditions *S* and event *A*, whoseoccurrence or non-occurrence under given conditions can be accurately established, naturalscience usually uses one of the following schemes:

(a) For each realization of conditions *S*, event *A* occurs. All the laws of classical mechanics have such a form, stating that for specified initial conditions and forces acting on an object or system of objects, the motion will proceed in an unambiguously definite manner.

(b) Under conditions S, event A has a definite probability $P(A/S)$ equal to p.
Thus, for example, the laws of radioactive emission assert that for eachradioactive substance there exists the specific probability that, for a given amount of a substance, a certain number of atoms N will decompose within a given time interval.

Let us call the frequency of event A in a given set of n trials (that is, of n repeated realizations of conditions S) the ratio $h = m/n$ of the number m of those trials in which A occurs to the total number of trials n. The existence of a specific probability equal to p for an event A under conditions S is manifested in the fact that in almost every sufficientlylong series of trials, the frequency of event A is approximately equal to p.

Statistical laws, that is, laws described by a scheme of type (b), were first discovered in games of chance similar to dice. The statistical rules of birth and death (for example, the probability of the birth of a boy is 0.515) have also been known for a long time. A great number of statistical laws in physics, chemistry, biology, and other sciences were discovered at the end of the 19th and in the first half of the 20th century.

The possibility of applying the methods of probability theory to the investigation ofstatistical laws, which pertain to a very wide range of scientific fields, is based on the fact that the probabilities of events always satisfy certain simple relationships, which will be discussed in the next section. The investigation of the properties of probabilities of events on the basis of these simple relationships is also a topic of probability theory.

3.6.1 Fundamental concepts of Probability theory.

The fundamental concepts of probability theory as a mathematical discipline are most simply defined in the framework of so-called elementary probability theory. Each trial T considered in elementary probability theory is such that it is ended by one and only one of the events $E_1, E_2, \dots, E_s$ (by one or another, depending on the case). These events are called outcomes of the trial. Each outcome E_k is connected with a positive number p_k, the probability of this outcome. The numbers p_k must add up to 1. Events A, which consist of the fact that ―either E_i, or E_j … , or E_k occurs,‖ are then considered. The outcomes $E_i, \dots, E_k$ are said to be favorable to A, and according to the definition, it is assumed that the probability $P(A)$ of event A is equal to the sum of the probabilities of the outcomes favorable to it:

(1) $\boldsymbol{P(A) = p_i + p_j + \dots + p_k}$

The particular case $p_1 = p_2 = p_s = 1/s$ leads to the formula:

(2) $\boldsymbol{P(A) = r/s}$

Formula (2) expresses the so-called classical definition of probability according to which the probability of some event A is equal to the ratio of the number r of outcomes favorable to A to the number s of all ―equally likely‖ outcomes. The classical definition of probability only reduces the concept of probability to the concept of equal possibility, which remains without a clear definition.

EXAMPLE. In the tossing of two dice, each of the 36 possible outcomes can be designatedby (i, j), where i is the number of pips that comes up on the first dice and j, the number on the second. The outcomes are assumed to be equally likely. To the event A, ―the sum of the pips is 4,‖ three outcomes are favorable: (1,3); (2,2); (3,1). Consequently, $P(A) = 3/36 = 1/12$.

Starting from certain given events, it is possible to define two new events: their union (sum) and intersection (product). Event B is called the **union** of events $A_1, A_2, \ldots, A_r$ if it has the form ‒A_1 or $A_2, \ldots,$ or A_r occurs.‖

Event C is called the **intersection** of events $A_1, A_2 \ldots, A_r$ if it has the form ‒A_1, and $A_2, \ldots$
, and A_r occurs.‖
The union of events is designated by the symbol ∪, and the intersection, by ∩. Thus, wewrite:

$B = A_1 \cup A_2 \cup \ldots \cup A_r$ $C = A_1 \cap A_2 \cap \ldots \cap A_r$
Events A and B are called disjoint if their simultaneous occurrence is impossible—that is,if among the outcomes of a trial not one is favourable to A and B simultaneously.
Two of the basic theorems of probability theory are connected with the operations of unionand intersection of events; these are the theorems of addition and multiplication of probabilities.

3.6.2 Theorem of Addition of Probabilities.
If events $A_1, A_2, \ldots, A_r$ are such that each two of them are disjoint, then the probability oftheir union is equal to the sum of their probabilities.

Thus, in the example presented above of tossing two dice, event B, ‒the sum of the pips does not exceed 4,‖ is the union of three disjoint events A_2, A_3, A_4, consisting of the fact thesum of the pips is equal to 2, 3, and 4, respectively. The probabilities of these events are 1/36, 2/36, and 3/36, respectively. According to the theorem of addition of probabilities, probability $P(B)$ is:

1/36 + 2/36 + 3/36 = 6/36 = 1/6

The conditional probability of event B under condition A is determined by the formula:

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

which, as can be proved, corresponds completely with the properties of frequencies.
Events $A_1, A_2, \ldots, A_r$ are said to be independent if the conditional probability of each ofthem, under the condition that some of the remaining events have occurred, is equal to its ‒absolute‖ probability.

3.6.3 Theorem of Multiplication of Probabilities.
The probability of the intersection of events $A_1, A_2, \ldots, A_r$ is equal to the probability ofevent A_1 multiplied by the probability of event A_2 under the condition that A_1 has occurred,
…, multiplied by the probability of A_r under the condition that $A_1, A_2, \ldots, A_{r-1}$ haveoccurred. For independent events, the multiplication theorem reduces to the formula:

$$P(A_1 \cap A_2 \cap \ldots \cap A_r) = P(A_1) \times P(A_2) \times \ldots \times P(A_r) \ldots\ldots\ldots (3)$$

that is, the probability of the intersection of independent events is equal to the product of the probabilities of these events. Formula (3) remains correct, if on both sides some of the events are replaced by their inverses.

EXAMPLE:

Four shots are fired at a target, and the hit probability is 0.2 for each shot. The target hitsby different shots are assumed to be independent events. What is the probability of hitting the target three times?

Each outcome of the trial can be designated by a sequence of four letters [for example, (s,f, f, s) denotes that the first and fourth shots hit the target (success), and the second and third miss (failure)]. There are 2 · 2 · 2 · 2 = 16 outcomes in all. In accordance with the assumption of independence of the results of individual shots, one should use formula (3) and the remarks about it to determine the probabilities of these outcomes. Thus, theprobability of the outcome (s, f, f, f) is set equal to 0.2 x 0.8 x 0.8 x 0.8 = 0.1024; here,

0.8 = 1 - 0.2 is the probability of a miss for a single shot. For the event ―three shots hit the target,‖ the possible outcomes are: (s, s, s, f), (s, s, f, s), (s, f, s, s), and (f, s, s, s) are favorable and the probability of each is the same:

0.2 · 0.2 · 0.2 · 0.8 = · · · = 0.8 · 0.2 · 0.2 · 0.2 = 0.0064

Consequently, the desired probability is 4 x 0.0064 = 0.0256.

Generalizing the discussion of the given example, it is possible to derive one of the fundamental formulas of probability theory: if events A_1, A_2, …, A_n are independent and each has a probability p, then the probability of exactly m such events occurring is:

$$P_n(m) = C_n^m (1-p)^{n-m} \quad (4)$$

Here, C^m_n denotes the number of combinations of n elements taken m at a time. For large n, the calculation using formula (4) becomes difficult. In the preceding example, let the numberof shots equal 100; the problem then becomes one of finding the probability x that the number of hits lies in the range from 8 to 32. The use of formula (4) and the addition theoremgives an accurate, but not a practically useful, expression for the desired probability

$$x = \sum_{m=8}^{32} C_{100}^{m} (0.2)^m (0.8)^{100-m}$$

The approximate value of the probability x can be found by the **Laplace theorem**

$$x \approx \frac{1}{\sqrt{2\pi}} \int_{-3}^{+3} e^{-z^2/2}\, dz = 0.9973$$

with the error not exceeding 0.0009. The obtained result demonstrates that the event $8 \leq m \leq 32$ is practically certain. This is the simplest, but a typical, example of the use of the limittheorems of probability theory.

Another fundamental formula of elementary probability theory is the so-called total probability formula: if events A_1, A_2, …, A_r are disjoint in pairs and their union is a certain event, then the probability of any event B is the sum

$$P(B) = \sum_{k=1}^{r} P(B|A_k)P(A_k)$$

The theorem of multiplication of probabilities turns out to be particularly useful in the consideration of compound trials. Let us say that trial T consists of trials T_1, T_2, …, T_{n-1}, T_n,if each outcome of trial T is the intersection of certain outcomes A_i, B_i, … , x_k, Y_l of the corresponding trials T_1, T_2, …, T_{n-1}, T_n. From one or another consideration, the following probabilities are often known:

$P(A_1)$, $P(B_j/A_i)$, … , $P(Y_l/A_i \cap B_j \cap … \cap X_k)$……………………..(5)

According to the probabilities of (5), probabilities *P(E)* for all the outcomes of *E* of the compound trial and, in addition, the probabilities of all events connected with this trial canbe determined using the multiplication theorem (just as was done in the example above).

Two types of compound trials are the most significant from a practical point of view:

(a) the component trials are independent, that is, the probabilities (5) are equal to the unconditional probabilities $P(A_i)$, $P(B_j)$, …, $P(Y_l)$; and

(b) the results of only the directly preceding trial have any effect on the probabilities ofthe outcomes of any trial—that is, the probabilities (5) are equal, respectively, to $P(A_i)$, $P(B_j/A_i)$, … , $P(Y_l/X_k)$.

In this case, it is said that the trials are connected in a Markov chain. The probabilities ofall the events connected with the compound trial are completely determined here by the initial probabilities $P(A_i)$ and the transition probabilities $P(B_j/A_i)$, … , $P(Y_l/X_k)$.

Often, instead of the complete specification of a probability distribution of a random variable, it is preferable to use a small number of numerical characteristics. The most frequently used are the mathematical expectation and the dispersion.

In addition to mathematical expectations and dispersions of these variables, a joint distribution of several random variables is characterized by correlation coefficients and so forth. The meaning of the listed characteristics is to a large extent explained by the **limit theorems**

3.7 The Limit theorems.

In the formal presentation of probability theory, limit theorems appear in the form of a superstructure over its elementary sections, in which all problems have a finite, purely arithmetic character. However, the cognitive value of probability theory is revealed only bythe limit theorems. Thus, the **Bernoulli theorem** proves that in independent trials, the frequency of appearance of any event, as a rule, deviates little from its probability, and the **Laplace theorem** indicates the probabilities of one or another deviation. Similarly, the meaning of such characteristics of a random variable as its mathematical expectation and dispersion is explained by the law of large numbers and the **central limit theorem**.

Let $X_1, X_2, …, X_n$, be independent random variables that have one and the same probabilitydistribution with $\mathbf{E}X_K = a$, $\mathbf{D}X_K = \zeta^2$ and Y_n be the arithmetic mean of the first *n* variablesof sequence such that:
$Y_n = (X_1 + X_2 + X_2 + \cdots + X_n)/n$

In accordance with the law of large numbers, for any $\varepsilon > 0$, the probability of the inequality
$| Y_n - a | \leq \varepsilon$ has the limit 1 as $n \to \infty$, and thus Y_n, as a rule, differs little from *a*.

The **central limit theorem** makes this result specific by demonstrating that the deviations of Y_n from *a* are approximately subordinate to a normal distribution with mean zero and dispersion $8^2/n$. Thus, to determine the probabilities of one or another deviation of Y_n from *a* for large *n*, there is no need to know all the details about the distribution of the variables X_n; it is sufficient to know only their dispersion.

In the 1920‘s it was discovered that even in the scheme of a sequence of identically distributed and independent random variables, limiting distributions that differ from the normal can arise in a completely

natural manner. Thus, for example, if X_1 is the time until the first reversion of some randomly varying system to the original state, and X_2 is the timebetween the first and second reversions, and so on, then under very general conditions the distribution of the sum $X_1 + \cdot \cdot \cdot + X_n$ (that is, of the time until the Aith reversion), after multiplication by $n^{-1/\alpha}$(α is a constant less than 1), converges to some limiting distribution. Thus, the time until the *nth* reversion increases, roughly speaking, as $n^{1/\alpha}$, that is, more rapidly than n (in the case of applicability of the law of large numbers, it is of the order of n).
The mechanism of the emergence of the majority of limiting regularities can be understoodultimately only in connection with the theory of random processes.

Random processes.
In a number of physical and chemical investigations of recent decades, the need has arisen to consider, in addition to one-dimensional and multidimensional random variables, random processes—that is, processes for which the probability of one or another of their courses is defined. In probability theory, a random process is usually considered as a one- parameter family of random variables $X(t)$. In an overwhelming number of applications, the parameter t represents time, but this parameter can be, for example, a point in space, and then we usually speak of a random function. In the case when the parameter t runs through the integer-valued numbers, the random function is called a **random sequence**. Just as a random variable is characterized by a distribution law, a random process can be characterized by a set of joint distribution laws for $X(t_1)$, $X(t_2)$, …, $X(t_n)$ for all possible moments of t_1, t_2, …, t_n for any $n > 0$.

3.7 Probability Distribution Functions
In probability theory and statistics, a **probability distribution** identifies either the probability of each value of a random variable (when the variable is discrete), or the probability of the value falling within a particular interval (when the variable is continuous). The probability distribution describes the range of possible values that a random variable can attain and the probability that the value of the random variable is within any (measurable) subset of that range.

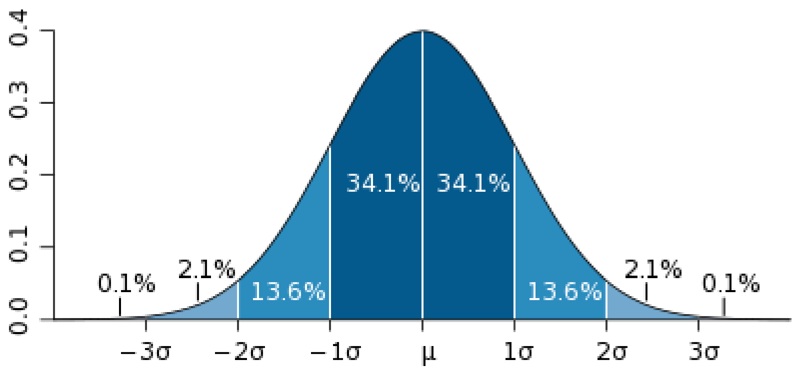


Figure 2: The Normal distribution, often called the "bell curve".

When the random variable takes values in the set of real numbers, the probability distribution is completely described by the *cumulative distribution function*, whose value ateach real x is the probability that the random variable is smaller than or equal to x.

The concept of the probability distribution and the random variables which they describe underlies the mathematical discipline of probability theory, and the science of statistics. There is spread or variability in almost any value that can be measured in a population (e.g.height of people, durability of a metal, sales growth, traffic flow, etc.); almost all measurements are made with some intrinsic error; also in physics many processes are described probabilistically, from the kinetic properties of gases to the quantum mechanical description of fundamental particles. For these and many other reasons, simple numbers are often inadequate for describing a quantity, while probability distributions are often more appropriate.

3.8.1 Probability distributions of real-valued random variables

Because a probability distribution Pr on the real line is determined by the probability of a real-valued random variable *X* being in a half-open interval (-∞, *x*], the probability distribution is completely characterized by its cumulative distribution function given as:

$$F(x) = \Pr[X \leq x] \qquad \forall x \in \mathbb{R}.$$

a. Discrete probability distribution

A probability distribution is called ***discrete*** if its cumulative distribution function only increases in jumps. More precisely, a probability distribution is discrete if there is a finiteor countable set whose probability is 1.

For many familiar discrete distributions, the set of possible values is discrete in the sense that all its points are isolated points. But, there are discrete distributions for which this countable set is dense on the real line.

Discrete distributions are characterized by a probability mass function, *p* such that:

$$\Pr[X = x] = p(x).$$

b. Continuous probability distribution

By one convention, a probability distribution μ is called ***continuous*** if its cumulative distribution function $F(x) = \mu(-\infty, x]$ is continuous and, therefore, the probability measure of singletons $\mu\{x\} = 0$.

Another convention reserves the term ***continuous probability distribution*** for absolutely continuous distributions. These distributions can be characterized by a probability densityfunction: a non-negative Lebesgue integrable function *f* defined on the real numbers suchthat

$$F(x) = \mu(-\infty, x] = \int_{-\infty}^{x} f(t)\,dt.$$

Discrete distributions and some continuous distributions do not admit such a density.

Terminologies

The **support** of a distribution is the smallest closed interval/set whose complement has probability zero. It may be understood as the points or elements that are actual members of the distribution.

A **discrete random variable** is a random variable whose probability distribution is discrete. Similarly, a **continuous random variable** is a random variable whose probabilitydistribution is continuous.

Some properties

- The probability density function of the sum of two independent random variables isthe **convolution** of each of their density functions.
- The probability density function of the difference of two independent randomvariables is the **cross-correlation** of their density functions.
- Probability distributions are not a vector space – they are not closed under linear

combinations, as these do not preserve non-negativity or total integral 1 – but they are closed under convex combination, thus forming a convex subset of the space of functions (or measures).
In mathematics and, in particular, functional analysis, **convolution** is a mathematicaloperation on two functions *f* and *g*, producing a third function that is typically viewed as a modified version of one of the original functions. Convolution is similar to cross- correlation. It has applications that include statistics, computer vision, image and signal processing, electrical engineering, and differential equations

3.8 Summary of Common Probability Distributions

The following is a list of some of the most common probability distributions, grouped by the type of process that they are related to.
Note that all of the univariate distributions below are singly-peaked; that is, it is assumed that the values cluster around a single point. In practice, actually-observed quantities may cluster around multiple values. Such quantities can be modeled using a mixture distribution.

3.9.1 Related to real-valued quantities that grow linearly (e.g. errors, offsets)

- Normal distribution (aka Gaussian distribution), for a single such quantity; the mostcommon continuous distribution;
- Multivariate normal distribution (aka multivariate Gaussian distribution), forvectors of correlated outcomes that are individually Gaussian-distributed;

3.9.2 Related to positive real-valued quantities that grow exponentially (e.g.prices, incomes, populations)

- Log-normal distribution, for a single such quantity whose log is normallydistributed
- Pareto distribution, for a single such quantity whose log is exponentiallydistributed; the prototypical power law distribution.

3.9.3 Related to real-valued quantities that are assumed to be uniformlydistributed over a (possibly unknown) region

- Discrete uniform distribution, for a finite set of values (e.g. the outcome of a fairdie)
- Continuous uniform distribution, for continuously-distributed values.

3.9.4 Related to Bernoulli trials (yes/no events, with a given probability)

- *Bernoulli* distribution, for the outcome of a single Bernoulli trial (e.g.success/failure, yes/no);
- *Binomial* distribution, for the number of "positive occurrences" (e.g. successes, yesvotes, etc.) given a fixed total number of independent occurrences;
- *Negative binomial* distribution, for binomial-type observations but where the quantity of interest is the number of failures before a given number of successes occurs‘
- *Geometric* distribution, for binomial-type observations but where the quantity ofinterest is the number of failures before the first success; a special case of the negative binomial distribution.

3.9.5 Related to sampling schemes over a finite population

- *Binomial* distribution, for the number of "positive occurrences" (e.g. successes, yesvotes, etc.) given a fixed number of total occurrences, using sampling with replacement
- *Hypergeometric* distribution, for the number of "positive occurrences" (e.g. successes, yes votes, etc.) given a fixed number of total occurrences, using samplingwithout replacement

- *Beta-binomial* distribution, for the number of "positive occurrences" (e.g. successes, yes votes, etc.) given a fixed number of total occurrences, sampling using a Polya urnscheme (in some sense, the "opposite" of sampling without replacement)

3.9.6 Related to categorical outcomes (events with *K* possible outcomes, with a givenprobability for each outcome)

- *Categorical* distribution, for a single categorical outcome (e.g. yes/no/maybe in a survey); a generalization of the Bernoulli distribution;
- *Multinomial* distribution, for the number of each type of categorical outcome, given afixed number of total outcomes; a generalization of the binomial distribution;
- *Multivariate* hyper geometric distribution, similar to the multinomial distribution, butusing sampling without replacement; a generalization of the hyper geometric distribution;

3.9.7 Related to events in a Poisson process (events that occur independently with agiven rate)

- *Poisson* distribution, for the number of occurrences of a Poisson-type event in a givenperiod of time
- *Exponential* distribution, for the time before the next Poisson-type event occurs

3.9.8 Useful for hypothesis testing related to normally-distributed outcomes

- *Chi-square* distribution, the distribution of a sum of squared standard normal variables; useful e.g. for inference regarding the sample variance of normally-distributed samples
- *Student's* t distribution, the distribution of the ratio of a standard normal variable andthe square root of a scaled chi squared variable; useful for inference regarding the mean of normally-distributed samples with unknown variance
- *F-distribution*, the distribution of the ratio of two scaled chi squared variables; useful

e.g. for inferences that involve comparing variances or involving R-squared (thesquared correlation coefficient).

Useful as conjugate prior distributions in Bayesian inference

- *Beta distribution*, for a single probability (real number between 0 and 1); conjugate tothe Bernoulli distribution and binomial distribution
- *Gamma distribution*, for a non-negative scaling parameter; conjugate to the rate parameter of a Poisson distribution or exponential distribution, the precision (inversevariance) of a normal distribution, etc.
- *Dirichlet distribution*, for a vector of probabilities that must sum to 1; conjugate tothe categorical distribution and multinomial distribution; generalization of the betadistribution
- *Wishart distribution*, for a symmetric non-negative definite matrix; conjugate to theinverse of the covariance matrix of a multivariate normal distribution; generalzationof the gamma distribution

4.0 Self-Assessment Exercise(s)

Answer the following questions:

1. Define probability distribution
2. What is the relationship between a random variable and probability distribution
3. List the distributions related to:

a. Bernoulli trials
b. Categorical outcomes

c. Hypothesis testing
4. The student is expected to familiarize him/herself with these probability distributionsand their applications.

5.0 Conclusion

The basis of simulation is randomness. Here we have discussed this fundamental basis which offers us the possibility to quantitatively represent uncertainties in simulations. With **Probabilities** in simulation we can explicitly represent uncertainties by specifying inputs as probability distributions.

6.0 Summary

In this unit we discussed the following:

- Defined Probability as
- Discussed the fundamental concepts of probability theory
- The limit theorem
- Random variables and Random processes
- Probability distributions
- Provided a listing of common probability distributions grouped by their relatedprocesses

7.0 Further Readings

- Devore, J. L. (2018). *Probability and statistics for engineering and the sciences*. Toronto,Ontario: Nelson.
- Georgii, H. (2013). *Stochastics: Introduction to probability and statistics*. Berlin: DeGruyter.
- Giri, N. C. (2019). *Introduction to probability and statistics*. London: Routledge.
- Johnson, R. A., Miller, I., & Freund, J. E. (2019). *Miller & Freunds probability andstatistics for engineers*. Boston: Pearson Education.
- Laha, R. G., & Rohatgi, V. K. (2020). *Probability theory*. Mineola, NY: DoverPublications.
- Mathai, A. M., & Haubold, H. J. (2018). *Probability and statistics: A course forphysicists and engineers*. Boston: De Gruyter.
- Pishro-Nik, H. (2014). *Introduction to probability, statistics, and random processes*.Blue Bell, PA: Kappa Research, LLC.
- Spiegel, M. R., Schiller, J. J., & Srinivasan, R. A. (2013). *Schaums outline of probabilityand statistics*. New York: McGraw-Hill.

Module 2: MODELLING AND SIMULATION CONCEPTS

Module Introduction

This module is divided into four (4) units

Unit 1: Simulation and Modelling
Unit 2: Modelling Methods
Unit 3: Physics-Based Finite Element Model
Unit 4: Statistics for Modelling and Simulation

Unit 1: Simulation and Modelling

Contents

1.0 Introduction

Real world phenomenon are very dynamic thus difficult to exactly predict. To make decisions in such circumstances, we need a tool or verifiable procedures that guide decision makers to an informed and provable decision and action. In this unit we will look at one such tool; simulation which has become cornerstone to many probabilistic projects.

2.0 Intended Learning Outcomes (ILOs)

At the end of this unit you should be to:

- Say what simulation is about
- State why we need simulation
- Describe how simulations are done
- Describe various types of Simulations

- Give examples of Simulation
- Show areas of applications of Simulation

3.0 Main Content

3.1 What is Simulation?

The term **simulation** is used in different ways by different people. As used here, simulation is defined as the process of creating a *model* (i.e., an abstract representation or exact copy) of an existing or proposed *system* (e.g., a project, a business, a mine, a forest, the organs in your body, etc.) in order to identify and understand those factors which control the system and/or to predict (forecast) the future behaviour of the system. Almost any system which can be quantitatively described using equations and/or rules can be simulated.

The underlying purpose of simulation is to shed light on the underlying mechanisms that control the behaviour of a system. More practically, simulation can be used to predict (forecast) the future behaviour of a system, and determine what you can do to influence that future behaviour. That is, simulation can be used to predict the way in which the system will evolve and respond to its surroundings, so that you can identify any necessary changes that will help make the system perform the way that you want it to.

For example, a fisheries biologist could dynamically simulate the salmon population in a river in order to predict changes to the population, and quantitatively understand theimpacts on the salmon of possible actions (e.g., fishing, loss of habitat) to ensure that they do not go extinct at some point in the future.
Also flight simulator on PC is also a computer model of some aspect of the flight; it shows on the screen the controls and what the pilot is supposed to see from the ―cockpit‖ (his armchair).

Simulation therefore, is a technique (not a method) for representing a dynamic real world system by a model and experimenting with the model in order to gain information aboutthe system and hence take appropriate decision. Simulation can be done by hand or by a computer.

Simulation is a powerful and important tool because it provides a way in which alternativedesigns, plans and/or policies can be evaluated without having to experiment on a real system, which may be prohibitively costly, time-consuming, or simply impractical to do. That is, it allows you to ask *"What if?" questions* about a system without having to experiment on the actual system itself (and hence incur the costs of field tests, prototypes, etc.).

3.2 When to Use simulation

Simulation is used in systems that change with time, such as a gas station, where cars come and go (called dynamic systems) and involve randomness. In such a system nobody can guess at exactly which time the next car should arrive at the station. Modelling complex dynamic systems theoretically need too many simplifications and the emerging models may not therefore be valid. Simulation does not require many simplifying assumptions, making it the only tool even in absence of randomness.
Simulation is used to observe the dynamic behaviour of a model of real or imaginary system. Indeed, by simulating a complex system we are able to understand the behaviour atlow cost. Otherwise we would have to carry out a complicated theoretical research or to build a device (an electric heater, a building or a plane), and observe how it changes to get hints for improvements in the design.
If you run a shop, an hospital or a bank, then computer simulation may show you bottlenecks, service time, flows, and queues of clients and provide important information on how to improve your business.

Note that often we describe a real world system by:
1. A physical model
2. A mathematical or analytic model
3. An analogue model.

What happens when a system is not amenable to treatment using the above model?Constructing a real physical system could be very expensive and what more testing it with live human beings and observing what happens could be fatal. Training a new pilot using an airplane is suicidal. This is why simulation is designed and utilized.
Thus simulation is the answer to our question. Many operations Research analysts consider simulation to be a method of last resort. This is because it may be useful when otherapproaches cannot be used, for example when a real world situation is complex. Note that nothing prevents you from using simulation approach to analytic problem. Results can at least be compared!
Thus, before designing and implementing a real life system, it is necessary to find out via simulation studies whether the system will work otherwise the whole exercise will be a wild goose chase. Inevitably huge sums of money might have been wasted.

Unlike the situation in mathematical programming, so far there are no clear cut underlying principle guiding the formulation of simulation models. Each application is ad-hoc to a large extent. In general there are three basic objectives of simulation studies:
1. To Describe a Current System – Suppose that a manufacturing company has suddenlyobserved a marked deterioration in meeting due-dates of customers order. It may be necessary to build a simulation model to see how the current procedures for estimating due dates, scheduling production and ordering raw materials are giving rise to observed delays.
2. To explore a Hypothetical System – such as installing a new system, which will cost alot of money, it might be better to build a hypothetical model of the system and learn from its behaviour.
3. To Design an Improved System – for example consider a supermarket that has onepayment counter. Due to increase in patronage, it is considering to increase the number of pay points. A simulation experiment may identify if one, two or moreadditional points are needed or not needed.

3.3 Types of Simulations
Computer models can be classified according to several criteria including:
- Stochastic or deterministic (and as a special case of deterministic, chaotic)
- Steady-state or dynamic
- Continuous or discrete (and as an important special case of discrete, discrete event orDE models)
- Local or distributed.For example:
- Steady-state models use equations defining the relationships between elements of the modelled system and attempt to find a state in which the system is in equilibrium. Such models are often used in simulating physical systems, as a simpler modelling case before dynamic simulation is attempted.
- Dynamic simulations model changes in a system in response to (usually changing) input signals.
- *Stochastic* models use *random number generators* to model the chance or random events; they are also called Monte Carlo simulations.

There are two basic types of simulation for which models are built, and the process of choosing the subset of characteristics or features is different for each. The distinction between the two types is based on how time is represented; either as a **continuous** variable oras a **discrete** variable.

3.3.1 Continuous simulation

Continuous simulations treat time as continuous and express changes in terms of a set of differential equations that reflect the relationships among the set of characteristics. Thus the characteristics or features chosen to model the system must be those whose behaviour is understood mathematically.

Continuous simulation is used in systems where the state changes all the time, not just at the time of some discrete events. For example, the water level in a reservoir due to in and outflow changes all the time. In such cases ***continuous simulation*** is more appropriate, although discrete events simulation can serve as an approximation.

A meteorological modelling falls is another example in this category. The characteristics of weather models are wind components, temperature, water vapour, cloud formation, precipitation, and so on. The interaction of these components over time can be modelled by a set of partial differential equations, which measure the rate of change of each component over some three-dimensional region.

A *continuous dynamic simulation* performs numerical solution of differential-algebraic equations or differential equations (either partial or ordinary). Periodically, the simulation program solves all the equations, and uses the numbers to change the state and output of the simulation. Applications include flight simulators, racing-car games, chemical process modelling, and simulations of electrical circuits. Originally, these kinds of simulations were actually implemented on analog computers, where the differential equations could be represented directly by various electrical components such as operational amplifiers. By the late 1980s, however, most "analogue" simulations were run on conventional digital computers that emulate the behaviour of an analog computer.

A typical Continuous (stochastic) system has a large number of control parameters that can have a significant impact on the performance of the system. To establish a basic knowledgeof the behaviour of a system under variation of input parameters, sensitivity analysis is usually performed, which applies small changes from one state to the nominal values of inputparameters. For such simulation, variations of the input parameter cannot be made infinitely small. The sensitivity of the performance measure with respect to an input parameter is therefore defined as (partial) derivative.

Sensitivity analysis is concerned with evaluating sensitivities (gradient) of performancemeasures with respect to parameter of interest. It provides guidance for design and operational decisions and plays a pivotal role in identifying the most significant system parameters, as well as bottleneck of subsystems.

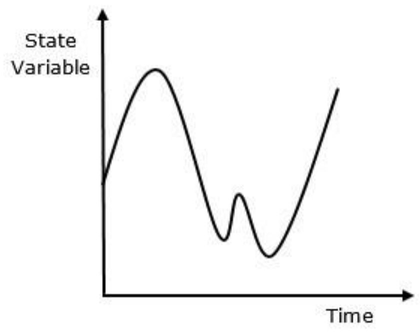
In designing, analysing and operating such complex systems, one is interested not only in performance evaluation but also in sensitivity analysis and optimisation.

3.3.2 Modelling & Simulation - Continuous

A continuous system is one in which important activities of the system completes smoothly without any delay, i.e. no queue of events, no sorting of time simulation, etc. When a continuous system is modeled mathematically, its variables representing the attributes are controlled by continuous functions.

3.3.2.1 What is Continuous Simulation?

Continuous simulation is a type of simulation in which state variables change continuously with respect to time. Following is the graphical representation of its behavior.



3.3.2.2 Why Use Continuous Simulation?

We have to use continuous simulation as it depends on differential equation of various parameters associated with the system and their estimated results known to us.

3.3.2.3 Application Areas

Continuous simulation is used in the following sectors. In civil engineering for the construction of dam embankment and tunnel constructions. In military applications for simulation of missile trajectory, simulation of fighter aircraft training, and designing & testing of intelligent controller for underwater vehicles.

In logistics for designing of toll plaza, passenger flow analysis at the airport terminal, and proactive flight schedule evaluation. In business development for product development planning, staff management planning, and market study analysis.

3.3.2.4 Modelling & Simulation ─ Application Areas

Modelling & Simulation can be applied to the following areas − Military applications, training & support, designing semiconductors, telecommunications, civil engineering designs & presentations, and E-business models.

Additionally, it is used to study the internal structure of a complex system such as the biological system. It is used while optimizing the system design such as routing algorithm, assembly line, etc. It is used to test new designs and policies. It is used to verify analytic solutions.

3.3.2.5 Various Concepts & Classification of modelling

Models & Events

Following are the basic concepts of Modelling & Simulation.

- **Object** is an entity which exists in the real world to study the behavior of a model.
- **Base Model** is a hypothetical explanation of object properties and its behavior, which is valid across the model.
- **System** is the articulate object under definite conditions, which exists in the real world.
- **Experimental Frame** is used to study a system in the real world, such as experimental conditions, aspects, objectives, etc. Basic Experimental Frame consists of two sets of variables − the Frame Input Variables & the Frame Output Variables, which matches the system or model terminals. The

Frame input variable is responsible for matching the inputs applied to the system or a model. The Frame output variable is responsible for matching the output values to the system or a model.

- **Lumped Model** is an exact explanation of a system which follows the specified conditions of a given Experimental Frame.
- **Verification** is the process of comparing two or more items to ensure their accuracy. In Modelling & Simulation, verification can be done by comparing the consistency of a simulation program and the lumped model to ensure their performance. There are various ways to perform validation process, which we will cover in a separate chapter.
- **Validation** is the process of comparing two results. In Modelling & Simulation, validation is performed by comparing experiment measurements with the simulation results within the context of an Experimental Frame. The model is invalid, if the results mismatch. There are various ways to perform validation process, which we will cover in separate chapter.

3.3.2.6 System State Variables

The system state variables are a set of data, required to define the internal process within the system at a given point of time.

- In a **discrete-event model**, the system state variables remain constant over intervals of time and the values change at defined points called event times.
- In **continuous-event model**, the system state variables are defined by differential equation results whose value changes continuously over time.

Following are some of the system state variables −

- **Entities & Attributes** − An entity represents an object whose value can be static or dynamic, depending upon the process with other entities. Attributes are the local values used by the entity.
- **Resources** − A resource is an entity that provides service to one or more dynamic entities at a time. The dynamic entity can request one or more units of a resource; if accepted then the entity can use the resource and release when completed. If rejected, the entity can join a queue.
- **Lists** − Lists are used to represent the queues used by the entities and resources. There are various possibilities of queues such as LIFO, FIFO, etc. depending upon the process.
- **Delay** − It is an indefinite duration that is caused by some combination of system conditions.

3.3.4 Discrete-event simulation,

Discrete event models are made up of *entitie*s, *attributes*, and *events*. An entity represents some object in the real system that must be explicitly defined. That is, the characteristic or feature of the system or an object. For example, if we were modelling a manufacturing plant, the different machines, and the product being created, would be entities. An attribute is some characteristic of a particular entity. The identification number, the purchase date, and the maintenance history would be attributes of a particular machine. An event is an interaction between entities. For example, the sending of the output from one machine as input to the next machine would be an event.

Suppose we are interested in a gas station. We may describe the behaviour of this system graphically by plotting the number of cars in the station; the state of the system. Every time a car arrives the graph increases by one unit while a departing car causes the graph to drop by one unit. This graph (called sample path), could be obtained from observation of real station, but could also be artificially constructed. Such *artificial construction and analysis of the resulting sample path (or more sample paths in more complex cases) consists of the simulation*.

The path consists of only horizontal and vertical lines, as cars arrivals and departures occurred, at distinct points in time, what we refer to as events. Between two consecutive events, nothing happens – the graph is horizontal. When the number of events are finite, we call the simulation ***discrete event***.

Discrete event systems (DES) are dynamic systems, which evolve in time by the occurrence of events at possible irregular time intervals. DES abound in real-world applications. Examples include traffic systems, flexible manufacturing systems, computer communication systems, production lines, flow networks etc. Most of these systems can be modelled in termsof discrete events whose occurrence causes the system to change from one state to another.
Simulations may be performed manually. Most often, however, the system model is writteneither as a computer program or as some kind of input into simulator software.

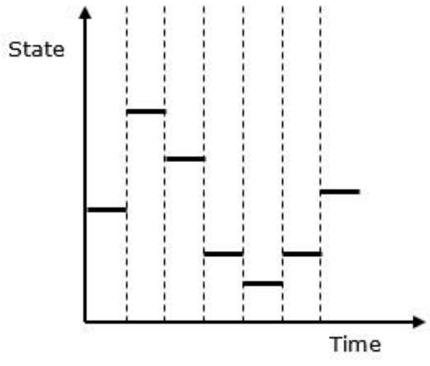
A *discrete event simulation* (DE) manages events in time. Most computer, logic-test and fault-tree simulations are of this type. In this type of simulation, the simulator maintains a queue of events sorted by the simulated time they should occur. The simulator reads the queue and triggers new events as each event is processed. It is not important to execute the simulation in real time. It's often more important to be able to access the data produced by the simulation, to discover logic defects in the design, or the sequence of events.

A special type of discrete simulation which does not rely on a model with an underlying equation, but can nonetheless be represented formally, is *agent-based simulation*. In agent- based simulation, the individual entities (such as molecules, cells, trees or consumers) in the model are represented directly (rather than by their density or concentration) and possess an internal *state* and set of behaviours or *rules* which determine how the agent's state is updated from one time-step to the next.

Discrete System Simulation

In discrete systems, the changes in the system state are discontinuous and each change in the state of the system is called an **event**. The model used in a discrete system simulation has a set of numbers to represent the state of the system, called as a **state descriptor**. In this chapter, we will also learn about queuing simulation, which is a very important aspect in discrete event simulation along with simulation of time-sharing system.

Following is the graphical representation of the behavior of a discrete system simulation.



Discrete Event Simulation ─ Key Features

Discrete event simulation is generally carried out by a software designed in high level programming languages such as Pascal, C++, or any specialized simulation language. Following are the five key features –

- **Entities** − These are the representation of real elements like the parts of machines.
- **Relationships** − It means to link entities together.
- **Simulation Executive** − It is responsible for controlling the advance time and executing discrete events.
- **Random Number Generator** − It helps to simulate different data coming into the simulation model.
- **Results & Statistics** − It validates the model and provides its performance measures.

Time Graph Representation

Every system depends on a time parameter. In a graphical representation it is referred to as clock time or time counter and initially it is set to zero. Time is updated based on the following two factors −

- **Time Slicing** − It is the time defined by a model for each event until the absence of any event.
- **Next Event** − It is the event defined by the model for the next event to be executed instead of a time interval. It is more efficient than Time Slicing.

Simulation of a Queuing System

A queue is the combination of all entities in the system being served and those waiting for their turn.

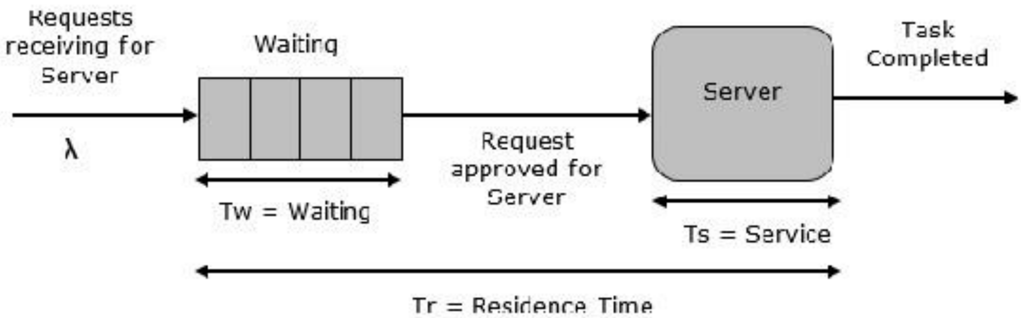
Parameters

Following is the list of parameters used in the Queuing System.

Symbol	Description
λ	Denotes the arrival rate which is the number of arrivals per second
Ts	Denotes the mean service time for each arrival excluding the waiting time in the queue
σTs	Denotes the standard deviation of service time
ρ	Denotes the server time utilization, both when it was idle and busy
u	Denotes traffic intensity
r	Denotes the mean of items in the system
R	Denotes the total number of items in the system
Tr	Denotes the mean time of an item in the system
TR	Denotes the total time of an item in the system
σr	Denotes the standard deviation of r
σTr	Denotes the standard deviation of Tr
w	Denotes the mean number of items waiting in the queue
σw	Denotes the standard deviation of w
Tw	Denotes the mean waiting time of all items
Td	Denotes the mean waiting time of the items waiting in the queue
N	Denotes the number of servers in a system
mx(y)	Denotes the y^{th} percentile which means the value of y below which x occurs y percent of the time

Single Server Queue

This is the simplest queuing system as represented in the following figure. The central element of the system is a server, which provides service to the connected devices or items. Items request to the system to be served, if the server is idle. Then, it is served immediately, else it joins a waiting queue. After the task is completed by the server, the item departs.

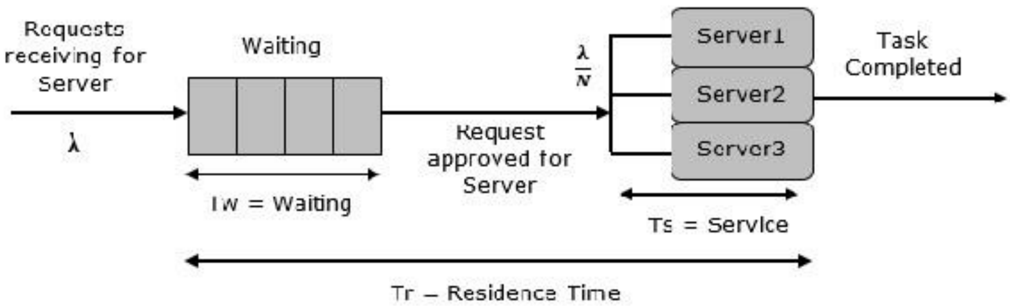


Multi Server Queue

As the name suggests, the system consists of multiple servers and a common queue for all items. When any item requests for the server, it is allocated if at-least one server is available. Else the queue begins to start until the server is free. In this system, we assume that all servers are identical, i.e. there is no difference which server is chosen for which item.

There is an exception of utilization. Let **N** be the identical servers, then **ρ** is the utilization of each server. Consider **Nρ** to be the utilization of the entire system; then the maximum utilization is **N*100%**, and the maximum input rate is −

$\lambda max = \frac{\text{N}}{\text{T}s}$



Queuing Relationships

The following table shows some basic queuing relationships.

General Terms	Single Server	Multi server
r = λTr Little's formula	ρ = λTs	ρ = λTs/N
w = λTw Little's formula	r = w + ρ	u = λTs = ρN
Tr = Tw + Ts		r = w + Nρ

Simulation of Time-Sharing System

Time-sharing system is designed in such a manner that each user uses a small portion of time shared on a system, which results in multiple users sharing the system simultaneously. The switching of each user is so rapid that each user feels like using their own system. It is based on the concept of CPU scheduling and multi-programming where multiple resources can be utilized effectively by executing multiple jobs simultaneously on a system.

Example − SimOS Simulation System.

It is designed by Stanford University to study the complex computer hardware designs, to analyze application performance, and to study the operating systems. SimOS contains software simulation of all the hardware components of the modern computer systems, i.e. processors, Memory Management Units (MMU), caches, etc.

3.4 Steps In Constructing A Simulation Model.

1. Formulate the model (see modelling)
2. Design the Experiment – Workout details of experimental procedures before runningthe model subsystems, parameters, relationships, data structures, etc.
3. Develop the Computer Programs – Each historical evolution of the model, includinggeneration of random events and generation of objects, will take place within the computer. if a model has a simple structure, you can use **BASIC, FORTRAN, PASCAL or C** and so on to develop the computerized version. However, it is betterto use a simulation language such as **SIMULATIONSCRIPT, GPSS, SIMULATIONULA (SIMULA), SIMULATIONNET (SIMNET) II, QMS**, etc.

3.4.3 To Extract the terms in Simulation

Let us consider building a simulation of gas station with a single pump served by a singleservice man.
Assumptions: arrival of cars as well as their times are random.
At first identify the:
State: number of cars waiting for service and number of cars served at any moment.
Event: arrival of cars, start of service, end of service.
Entities: these are the cars.
Queue: the queue of cars in front of the pump waiting for service.
Random realization**:** interval times, service times.
Distribution: we shall assume exponential distributions for the interval time and service time.Next,

specify what to do at each event. The above example would look like this:

At event of entity arrival: Create next arrival. If the server is free, send entity for start of service. Otherwise it joins the queue. At event of service start: Server becomes occupied. Schedule end of service for this entity. At event of service end: Server becomes free. If any entity is waiting in the queue: remove the first entity from the queue; send it for start of service.

Some initiation is still required, for example, the creation of the first arrival. Lastly, the above is translated into code. This is easy with appropriate library function, which has subroutine for creation, scheduling, proper timing of events, queue manipulations, random variate generation and statistics collection.

3.4.4 Simulation terminologies:

State – A variable characterizing an attribute in the system such as level of stock in inventoryor number of jobs in waiting for processing.
Event: **-** An occurrence at a point in time which may change the state of the system, such asarrival of a customer or start of work on a job.
Entity: An object that passes through the system, such as cars in an intersection or orders ina factory. Often an event (e.g., arrival) is associated with an entity (e.g., customer).

Queue: A queue is not only a physical queue of people, or cars, etc it can also be a task list, abuffer of finished goods waiting for transportation or any place where entities are waiting for something to happen for any reason.
Creating: Is causing an arrival of new entity into the system at some point in time.
Scheduling: is the act of assigning a new future event to an existing entity.
Random variable: is a quantity that is uncertain, such as interval time between two incoming flightsor number of defectives parts in a shipment.
Random Variate: is an artificially generated random variable.
Distribution: is the mathematical law, which governs the probabilistic features of a randomvariable.

3.5 Applications of Computer Simulation

Computer simulation has become a useful part of modelling many natural systems in physics, chemistry and biology, and human systems in economics and social science (the computational sociology) as well as in engineering to gain insight into the operation of those systems. A good example of the usefulness of using computers to simulate can be found in the field of network traffic simulation. In such simulations the model behaviour will change each simulation according to the set of initial parameters assumed for the environment. Computer simulations are often considered to be *human out of the loop* simulations.

Computer graphics can be used to display the results of a computer simulation. Animations can be used to experience a simulation in real-time e.g. in training simulations. In some casesanimations may also be useful in faster than real-time or even slower than real-time modes. For example, faster than real-time animations can be useful in visualizing the build up of queues in the simulation of humans evacuating a building.

There are many different types of computer simulation; the common feature they all share is the attempt to generate a sample of representative scenarios for a model in which a complete enumeration of all possible states of the model would be prohibitive or impossible. Several software packages also exist for running computer-based simulation modelling that makes the modelling almost effortless and simple.

a. Simulation in computer science

In computer science, simulation has an even more specialized meaning: Alan Turing uses the term "simulation" to refer to what happens when a digital computer runs a state transition table (runs a program) that describes the state transitions, inputs and outputs of a subject discrete-state machine. The computer simulates the subject machine.

In computer programming, a simulator is often used to execute a program that has to run on some inconvenient type of computer, or in a tightly controlled testing environment For example, simulators are usually used to debug a micro program or sometimes commercial application programs. Since the operation of the computer is simulated, all of the informationabout the computer's operation is directly available to the programmer, and the speed and execution of the simulation can be varied at will.

Simulators may also be used to interpret fault trees, or test very large scale integration (VLSI) logic designs before they are constructed. In theoretical computer science the term *simulation* represents a relation between state transition systems.

b. Simulation in training

Simulation is often used in the training of civilian and military personnel. This usually occurswhen it is prohibitively expensive or simply too dangerous to allow trainees to use the real equipment in the real world. In such situations they will spend time learning valuable lessonsin a "safe" virtual environment. Often the convenience is to permit mistakes during training for a safety-critical system.

Training simulations typically come in one of three categories:

- **live** simulation (where real people use simulated (or "dummy") equipment in the realworld);
- **virtual** simulation (where real people use simulated equipment in a simulated world(or "virtual environment")), or
- **constructive** simulation (where simulated people use simulated equipment in a simulated environment). Constructive simulation is often referred to as "wargaming"since it bears some resemblance to table-top war games in which players command armies of soldiers and equipment which move around a board.

c. Simulation in Education

Simulations in education are somewhat like training simulations. They focus on specific tasks. In the past, video has been used for teachers and students to observe, problem solveand role play; however, a more recent use of simulation in education include animated narrative vignettes (ANV). ANVs are cartoon-like video narratives of hypothetical and reality-based stories involving classroom teaching and learning. ANVs have been used to assess knowledge, problem solving skills and dispositions of children, and pre-service and in-service teachers.

Another form of simulation has been finding favour in business education in recent years. Business simulations that incorporate a dynamic model enable experimentation with businessstrategies in a risk free environment and provide a useful extension to case study discussions.

d. Medical Simulators

Medical simulators are increasingly being developed and deployed to teach therapeutic and diagnostic procedures as well as medical concepts and decision making to personnel in the health professions. Simulators have been developed for training procedures ranging from the basics such as blood draw, to laparoscopic surgery and trauma care.

Many medical simulators involve a computer connected to a plastic simulation of the relevant anatomy. Sophisticated simulators of this type employ a life size mannequin which responds to injected drugs and can be programmed to create simulations of life- threatening emergencies. In others simulations, visual components of the procedure are reproduced by computer graphics techniques, while touch-based components are reproduced by haptic feedback devices combined with physical simulation routines computed in response to the user's actions. Medical simulations of this sort will often use 3D CT or MRI scans of patient data to enhance realism.

Another important medical application of a simulator -- although, perhaps, denoting a slightly different meaning of *simulator* -- is the use of a *placebo* drug, a formulation which simulates the active drug in trials of drug efficacy.

e. City Simulators / Urban Simulation

A City Simulator is a tool used by urban planners to understand how cities are likely to evolve in response to various policy decisions. UrbanSim. The City Simulator developed at the University of Washington and ILUTE developed at the University of Toronto are examples of modern, large-scale urban simulators designed for use by urban planners. City simulators are generally agent-based simulations with explicit representations for land useand transportation.

f. Flight simulators

A flight simulator is used to train pilots on the ground. It permits a pilot to crash his simulated "aircraft" without being hurt. Flight simulators are often used to train pilots to operate aircraft in extremely hazardous situations, such as landings with no engines, or complete electrical or hydraulic failures. The most advanced simulators have high-fidelity visual systems and hydraulic motion systems. The simulator is normally cheaper to operate than a real trainer aircraft.

g. Marine simulators

This bears resemblance to flight simulators. The marine simulators are used to train ship personnel. Simulators like these are mostly used to simulate large or complex vessels, such as cruise ships and dredging ships. They often consist of a replication of a ships' bridge, with operating desk(s), and a number of screens on which the virtual surroundings are projected.

Simulation in Engineering (Technology) Process

Simulation is an important feature in engineering systems or any system that involves many processes. For example in electrical engineering, delay lines may be used to simulate propagation delay and phase shift caused by an actual transmission line. Similarly, dummy loads may be used to simulate impedance without simulating propagation, and is used in situations where propagation is unwanted. A simulator may imitate only a few of the operations and functions of the unit it simulates. *Contrast with*: emulate.

Most engineering simulations entail mathematical modelling and computer assistedinvestigation. There are many cases, however, where mathematical modelling is not reliable. Simulation of fluid dynamics problems often requires both mathematical and physical simulations. In these cases the physical models require dynamic similitude. Physical and chemical simulations have also direct realistic uses, rather than research uses; in chemical engineering, for example, process simulations are used to give the process parameters immediately used for operating chemical plants, such as oil refineries.

Discrete Event Simulation is often used in industrial engineering, operations management and operational research to model many systems (commerce, health, defence, manufacturing,logistics, etc.); for example, the value-adding transformation processes in businesses, to optimize business performance. Imagine a business, where each person could do 30 tasks, where thousands of products or services involved dozens of tasks in a sequence, where customer demand varied seasonally and forecasting was inaccurate- this is the domain where such simulation helps with business decisions across all functions.

h. Simulation and games

Strategy games - both traditional and modern - may be viewed as simulations of abstracted decision-making for the purpose of training military and political leaders. In a narrower sense, many video games are also simulators, implemented inexpensively. These are sometimes called "sim games". Such games can simulate various aspects of reality, from economics to piloting vehicles, such as flight simulators (described above).

i. The "classroom of the future"

The "classroom of the future" will probably contain several kinds of simulators, in additionto textual and visual learning tools. This will allow students to enter school better prepared, and with a higher skill level. The advanced student or postgraduate will have a more concise and comprehensive method of retraining -- or of incorporating new academic contents into their skill set -- and regulatory bodies and institution managers will find it easier to assess theproficiency and competence of individuals.

In classrooms of the future, the simulator will be more than a "living" textbook; it will become an integral a part of the practice of Education and training. The simulator environment will also provide a standard platform for curriculum development in educationalinstitutions.

3.6 Model Evaluation

An important part of the modelling process is the evaluation of an acquired model. *How do we know if a mathematical model describes the system well?* This is not an easy question to answer. Usually the engineer has a set of measurements from the system which are used in creating the model. Then, if the model was built well, the model will adequately show the relations between system variables for the measurements at hand. The question then becomes: How do we know that the measurement data are a representative set of possible values? Does the model describe well the properties of the system between the measurement data (interpolation)? Does the model describe well events outside the measurement data (extrapolation)?

A common approach is to split the measured data into two parts; training data and verification data. The training data are used to *train* the model, that is, to estimate the model parameters (see above). The verification data are used to evaluate model performance. Assuming that the training data and verification data are not the same, we can assume that if the model describes the verification data well, then the model describes the real system well.

However, this still leaves the *extrapolation question* open. How well does this model describe events outside the measured data? Consider again Newtonian classical mechanics- model. Newton made his measurements without advanced equipment, so he could not measure properties of particles travelling at speeds close to the speed of light. Likewise, he did not measure the movements of molecules and other small particles, but macro particles only. It is then not surprising that his model does not extrapolate well into these domains, even though his model is quite sufficient for ordinary life physics.

The reliability and the trust people put in computer simulations depends on the validity of thesimulation model, therefore verification and validation are of crucial importance in thedevelopment of computer simulations. Another important aspect of computer simulations is that of reproducibility of the results, meaning that a simulation model should not provide a different answer for each execution. Although this might seem obvious, this is a special pointof attention in stochastic simulations, where random numbers should actually be semi-random numbers. An exception to reproducibility are human in the loop simulations such as flight simulations and computer games. Here a human is part of the simulation and thus influences the outcome in a way that is hard if not impossible to reproduce exactly.

Case Studies

Operations study to add a new plane arrival at La Guardia southwest terminal LaGuardia airport planned to add a new flight to the schedule of the southwest terminal. Theairport administration wanted to understand how the introduction of a new flight would influence terminal capacity.

Problem

In order to understand the scale of the problem, the developers conducted a preliminary staticpedestrian flow analysis based on data of how long before the flight passengers arrived at the airport. In the picture, the solid line represented the number of seats in the waiting area, the red stacks represented the number of passengers in the terminal before introducing the new flight, and additional passengers from the new flight were represented by purple areas. The graph showed that if the new plane took off in the afternoon at 5:00 pm, the already crowded waiting area would have to bear an additional burden that could lead to a significant problem.The developers used the AnyLogic Pedestrian Library to create a crowd simulation model of the terminal in order to examine the use of seats under different scenarios. The basic model displayed the operation of all terminal areas before the introduction of the new flight, andthen various assumptions could be checked against this model. The best situation was when people were waiting for departure at their gates, but the consultants wanted to check how far they would have to move away from their gates to wait for their departure.

To set up the crowd simulation model, the developers used tables of passenger preference for waiting areas.

The model showed how far from their gate people would have to wait. The results of modeling the base scenario, without the new flight, showed that some of the peaks were reduced compared with the static analysis. This was due to passengers lining up 30 minutes before their flights. The model also showed where the people would actually wait. From this, it could be verified that there was no overflow and that the situation was stable.

In the afternoon, the waiting area was a lot more heavily utilized. There were a lot of passengers mixing in different areas and waiting for different gates. With the new flight at this peak time, some of these areas would get extremely overloaded. This pedestrian simulation was very useful in showing the operations of this terminal and how adding the new flight would affect the passengers in this area, including how far they would have to move to wait for their flights.

Solution

Designing large transport facilities requires careful consideration and agreement on every detail. That means that such projects must go through a great deal of decision making. The initial task of engineers usually produces alternatives and functional designs. These consider physical requirements and standards, but whether business or operating objectives will be metcan be hard to determine accurately. It is here that AnyLogic based modeling helps by enabling faster decision-making and significantly improving insight into the various tasksthat engineers face when planning large transport facilities.

4.0 Self-Assessment Exercise(s)

Answer the following questions:

- What are the objectives of simulation?
- In one sentence for each distinguish between different types of simulation
- Briefly describe simulation in five application areas.

5.0 Conclusion

Simulation is used to shed light on the underlying mechanisms that control the behaviour of asystem. It can be used to predict (forecast) the future behaviour of a system, and determine what you can do to influence that future behaviour. We simulation when we require information to solve bottlenecks, service time, flows, and queues of clients and provide important information on how to improve your business.

We simulate when a system is not amenable to treatment using any of the physical model, mathematical or analogue models. Other reasons to resort to simulation include when it is very expensive to construct a real physical system and what more testing it with live human beings and observing what happens could be fatal. Training a new pilot using an airplane is suicidal. These are where and when simulations are designed and utilized.

6.0 Summary

In this unit we:

- defined simulation as the process of creating a *model* (i.e., an abstract representation or exact copy) of an existing or proposed *system* (e.g., a project, a business, a mine, aforest, the organs in your body, etc.) in order to identify and understand those factorswhich control the system and/or to predict (forecast) the future behaviour of the system.
- Stated that simulation is required when a system is not amenable to treatment using any existing model or when it is very expensive to construct a real physical system orwhen testing it with live human beings could be fatal.
- classified simulation into:
    - Stochastic or deterministic (and as a special case of deterministic, chaotic)
    - Steady-state or dynamic
    - Continuous or discrete (and as an important special case of discrete, discreteevent or DE models)
    - Local or distributed
- Stated that simulations are done by: Formulating the model, Design the Experimentand Developing the Computer Programs.
- Listed areas of applications of Simulation to include: Computer science, Medicine,Education, City/Urban planning, Training etc.

7.0 Further Readings

- Gordon, S. I., & Guilfoos, B. (2017). *Introduction to Modeling and Simulation with MATLAB® and Python*. Milton: CRC Press.
- Zeigler, B. P., Muzy, A., & Kofman, E. (2019). *Theory of modeling and simulation: Discrete event and iterative system computational foundations*. San Diego (Calif.):Academic Press.
- Kluever, C. A. (2020). *Dynamic systems modeling, simulation, and control*. Hoboken,N.J: John Wiley & Sons.
- Law, A. M. (2015). *Simulation modeling and analysis*. New York: McGraw-Hill.
- Verschuuren, G. M., & Travise, S. (2016). *100 Excel Simulations: Using Excel to ModelRisk, Investments, Genetics, Growth, Gambling and Monte Carlo Analysis*. Holy Macro! Books.
- Grigoryev, I. (2015). *AnyLogic 6 in three days: A quick course in simulation modeling*.Hampton, NJ: AnyLogic North America.
- Dimotikalis, I., Skiadas, C., & Skiadas, C. H. (2011). *Chaos theory: Modeling, simulationand applications: Selected papers from the 3rd Cghaotic Modeling and Simulation International Conference (CHAOS2010), Chania, Crete, Greece, 1-4 June, 2010*. Singapore: World Scientific.
- Velten, K. (2010). *Mathematical modeling and simulation: Introduction for scientists and engineers*. Weinheim: Wiley-VCH.

Unit 2: Modelling Methods

Contents

1.0 Introduction

Modelling is an essential and inseparable part of all scientific activity, and many scientific disciplines have their own ideas about specific types of modelling. There is little general theory about scientific modelling, offered by the philosophy of science, systems theory, and new fields like knowledge visualization.

We create **models** for representation of the objects within a system together with the rules that govern the interactions of the objects. The representation may be concrete as in the case of the spaceship or flight simulators or abstract as in the case of the computer programthat examines the number of checkout stations in service queue.

2.0 Intended Learning Outcomes (ILOs)

At the end of this unit, the student should be able to:

- Define Modelling
- Describe some basic modelling concepts
- Differentiate between Visual and Conceptual models
- Explain the Characteristics of Visual, models

3.0 Main ContentDefinitions of Modelling

Modelling is the process of generating abstract, conceptual, graphical and/or mathematical models. Science offers a growing collection of methods, techniques and theory about all kinds of specialized scientific modelling. A scientific model can provide a way to read elements easily which have been broken down to a simpler form.

Model

A scientific model seeks to represent empirical objects, phenomena, and physical processesin a logical and objective way. All models are simplified reflections of reality, but despite their inherent falsity, they are nevertheless extremely useful. Building and disputing models is fundamental to the scientific enterprise. Complete and true representation maybe impossible but scientific debate often concerns which is the better model for a given task, such as the most accurate climate model for seasonal forecasting.

For the scientist, a **model** is also a way in which the human thought processes can be amplified. For instance, models that are rendered in software allow scientists to leverage computational power to simulate, visualize, manipulate and gain intuition about the entity, phenomenon or process being represented.

3.1 Basic Modelling Concepts

Modelling as a substitute for direct measurement and experimentation

Models are typically used when it is either impossible or impractical to create experimentalconditions in which scientists can directly measure outcomes. Direct measurement of outcomes under controlled conditions will always be more accurate than modelled estimates of outcomes. When predicting outcomes, models use *assumptions*, while measurements do not. As the number of assumptions in a model increases, the accuracyand relevance of the model diminishes.

Modelling language

A *modelling language* is any *artificial language* that can be used to express information or knowledge or systems in a structure that is defined by a consistent set of rules. The rulesare used for interpretation of the meaning of components in the structure.

Simulation

A *simulation* is the implementation of a model. A steady state simulation provides information about the system at an instant in time (usually at equilibrium, if it exists). A dynamic simulation provides information over time. A simulation brings a model to lifeand shows how a particular object or phenomenon will behave. It is useful for testing, analysis or training where real-world systems or concepts can be represented by a model.

Structure

Structure is a fundamental and sometimes intangible notion covering the recognition,observation, nature, and stability of patterns and relationships of entities. From a child's verbal description of a *snow*, to the detailed *scientific analysis* of the properties of magnetic fields, the concept of structure is an essential foundation of nearly every mode of inquiry and discovery in science, philosophy, and art.

Systems

A *system* is a set of interacting or interdependent entities, real or abstract, forming anintegrated whole. In general, a system is a construct or collection of different elements thattogether can produce results not obtainable by the elements alone. The concept of an 'integrated whole' can also be stated in terms of a system embodying a set of relationshipswhich are differentiated from relationships of the set to other elements, and fromrelationships between an element of the set and elements not a part of the relational regime.

There are two types of systems:

1) Discrete, in which the variables change instantaneously at separate points in time and,
2) Continuous, where the state variables change continuously with respect to time.

3.1.1 The process of generating a model

Modelling refers to the process of generating a model as a conceptual representation of some phenomenon. Typically a model will refer only to some aspects of the phenomenonin question, and two models of the same phenomenon may be essentially different, that isin which the difference is more than just a simple renaming. This may be due to differing requirements of the model's end users or to conceptual or aesthetic differences by the modellers and decisions made during the modelling process. *Aesthetic* considerations that may influence the *structure* of a model might be the modeller's preferences regarding probabilistic models vis-a-vis deterministic ones, discrete vs continuous time etc. For this reason users of a model need to understand the model's original purpose and the assumptions of its validity.

3.1.2 Factors in evaluating a model

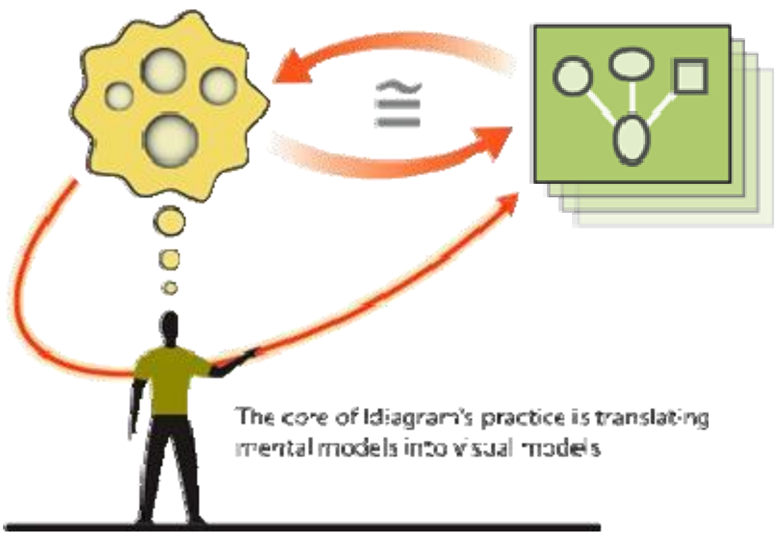
A model is evaluated first and foremost by its consistency to empirical data; any model inconsistent with reproducible observations must be modified or rejected. However, a fit toempirical data alone is not sufficient for a model to be accepted as valid. Other factors important in evaluating a model include:

- Ability to explain past observations
- Ability to predict future observations
- Cost of use, especially in combination with other models
- Refutability, enabling estimation of the degree of confidence in the model
- Simplicity, or even aesthetic appeal

3.2 Visual and Conceptual models

Visualization is any technique for creating images, diagrams, or animations tocommunicate a message. Visualization through visual imagery has been an effective wayto communicate both abstract and concrete ideas since the dawn of man. Examples from history include cave paintings, Egyptian hieroglyphs, Greek geometry, and Leonardo da Vinci's revolutionary methods of technical drawing for engineering and scientific purposes.We should not hold a narrow definition of exactly what a **visual model** should *look like*. We should rather use whatever visual elements or styles such as diagrams, maps, graphs, charts, pictures, cartoons, etc. – that will most effectively represent the problem at hand.

We can however define visual models by what they strive to do, and list some of the important characteristics that distinguish 'visual models' from other kinds of graphic art.



Visual representation of data depends fundamentally on an appropriate **visual** scheme for mapping numbers into graphic patterns (Berlin 1983). One reason for the widespread use of graphical methods for quantitative data is the availability of a natural **visual** mapping: magnitude can be represented by length, as in a bar chart, or by position along a scale, as indot charts and scatter plots. One reason for the relative paucity of graphical methods for categorical data may be that a natural **visual** mapping for frequency data is not so apparent.

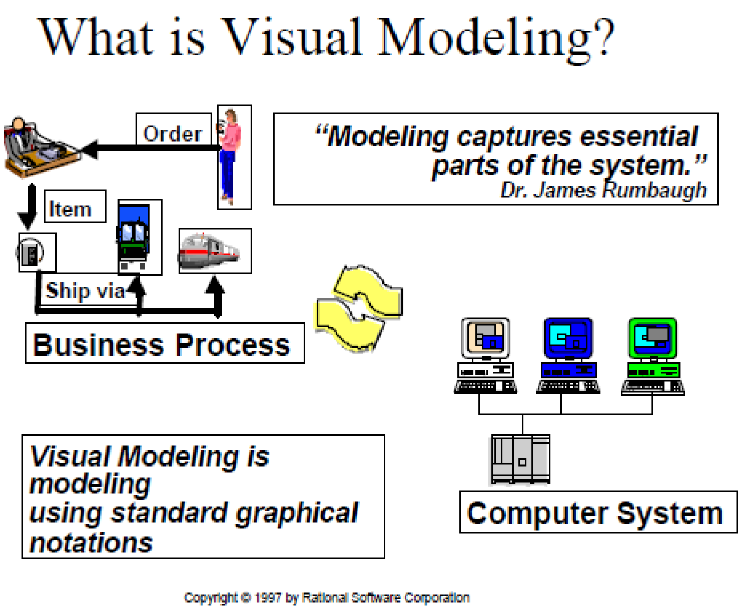
Conceptual model helps you interpret what is shown in a drawing or graph. A good **conceptual** model for a graphical display will have deeper connections with underlying statistical ideas as well.

For quantitative data, position along a scale can be related to mechanical models in which fitting data by least squares or least absolute deviations correspond directly to balancing forces or minimizing potential energy (Farebrother 1987).

The mechanical model for least squares regression, for example, likens each observation toa unit mass connected vertically to a rod by springs of unit modulus. Sall (1991a) shows how this mechanical model neatly describes the effects of sample size on power of a test, the leverage of outlying observations in regression, principal components, and collinearity among others.

Conceptual Modelling

- Is used for abstract (visual) representation of the problem domain
- It serves to enhance understanding of complex problem domains.
- It provides a basis for communication among project team members



3.3 Features of Visual and Conceptual ModelA visual model should:

- Render **conceptual knowledge** as opposed to quantitative data (information visualization) or physical things (technical illustration). We usually express conceptual knowledge with words alone, and yet the meaning behind those wordsis often *inherently visual*. Visual models seek to render directly the image- schematic (meaning that lies behind our words).
- Be **good models** - the images should accurately reflect the situation in the world and embody the characteristics of a useful model.

- **Integrate** the most salient aspects of the problem into a clear and coherent picture.
- **Fit the visual structure to the problem** – and not force the problem into apredefined visual structure.
- Use a **consistent visual gramma**r.
- Should be **visually and cognitively *tractable***. Visual models exist to support robust qualitative thinking: they're software for 'human-simulation' (as opposed to computer-simulation) of the issue at hand. To serve as effective 'simulation software', visual models must be 'readable' and 'run able' by our visio-cognitive 'hardware' and should positively engage our prodigious visual intelligence.
- Tap into the power of **elegant design**. In other words, they shouldn't be ugly

Conceptual Modelling

- A good conceptual model should NOT reflect a solution bias.
- Should model the problem domain, not the solution domain.
- Initial conceptual model may be rough and general.
- May be refined during incremental development.

3.4 Cognitive Affordances of Visual Models

Due to the limited capacity of our working memory, 7 ± 2 ‗chunks' of information, we cannot hold in our minds concepts, arguments, or problems that consist of more than 5 to 9objects or relationships. While this cognitive limitation severely restricts our ability tothink about complex things, we can do what we often do: extend our intellectual abilities with external representations or 'models' of the problem.

The particular affordances diagrams – their ability to simultaneously show many objects and relationships – make them an ideal tool for thinking about conceptually-complex problems. Diagrams provide an external mnemonic aid that enables us to see complicated relationships and easily move between various mind-sized groupings of things.

Self-Assessment Exercise(s)

Answer the following questions:

1. Differentiate between modelling and a model
2. What factors are important in evaluating a model?
3. What are the desirable features of a visual model?

5.0 Conclusion

The essence of constructing a model is to identify a small subset of characteristics or features that are sufficient to describe the behaviour of the system under investigation. Since a model is an abstraction of a real system and not the system itself, there is therefore,a fine line between having too few characteristics to accurately describe the behaviour of the system and more than you need to accurately describe the system. The goal should beto build the simplest model that effectively describes the relevant behaviour of the system.

6.0 Summary

- We defined **modelling** as the process of generating abstract, conceptual, graphical and/or mathematical models. Science offers a growing collection of methods, techniques andtheory about all kinds of specialized scientific modelling.

- We Listed and briefly explained some basic modelling concepts
- Differentiating between Visual and Conceptual models
- we discussed the important factors in evaluating a model to include:
  - Ability to explain past observations
  - Ability to predict future observations
  - Cost of use, especially in combination with other models
  - Refutability, enabling estimation of the degree of confidence in the model
  - Simplicity, or even aesthetic appeal
- We discussed the features of a good visual model which include:
  - Ability to render conceptual knowledge as opposed to quantitative data(information visualization) or physical things (technical illustration),
  - the images should accurately reflect the situation in the world,
  - the model should Integrate the most salient aspects of the problem into a clearand coherent picture,
  - Fit the visual structure to the problem,
  - It should Use a consistent visual grammar,
  - Should be visually and cognitively tractable.
- We also stated the Characteristics of Conceptual models

7.0 Further Readings

- Gordon, S. I., & Guilfoos, B. (2017). *Introduction to Modeling and Simulation withMATLAB® and Python*. Milton: CRC Press.
- Zeigler, B. P., Muzy, A., & Kofman, E. (2019). *Theory of modeling and simulation: Discrete event and iterative system computational foundations*. San Diego (Calif.):Academic Press.
- Kluever, C. A. (2020). *Dynamic systems modeling, simulation, and control*. Hoboken,N.J: John Wiley & Sons.
- Law, A. M. (2015). *Simulation modeling and analysis*. New York: McGraw-Hill.
- Verschuuren, G. M., & Travise, S. (2016). *100 Excel Simulations: Using Excel to ModelRisk, Investments, Genetics, Growth, Gambling and Monte Carlo Analysis*. Holy Macro! Books.
- Grigoryev, I. (2015). *AnyLogic 6 in three days: A quick course in simulation modeling*.Hampton, NJ: AnyLogic North America.
- Dimotikalis, I., Skiadas, C., & Skiadas, C. H. (2011). *Chaos theory: Modeling, simulationand applications: Selected papers from the 3rd Cghaotic Modeling and Simulation International Conference (CHAOS2010), Chania, Crete, Greece, 1-4 June, 2010*. Singapore: World Scientific.
- Velten, K. (2010). *Mathematical modeling and simulation: Introduction for scientists and engineers*. Weinheim: Wiley-VCH.

Unit 3: Finite Element Model and Database Model

Contents

1.0 Introduction

The finite element method is one of the most powerful approaches for approximate solutions to a wide range of problems in mathematical physics. The method has achieved acceptance in nearly every branch of engineering and is the preferred approach in structural mechanics and heat transfer. Its application has extended to soil mechanics, heat transfer, fluid flow, magnetic field calculations, and other areas.

Managing large quantities of **structured** and **unstructured** data is a primary function of **information systems**. Data models describe structure of **data** for **storage** in data management systems such as relational databases. They typically do not describe unstructured data, such as documents, word processing, email messages, pictures, digital audio, and video

2.0 Intended Learning Outcomes (ILOs)

After studying this unit the reader should be able to:

- Define Finite Element Method (FEM)
- Describe the relationship between FEM and Finite element analysis
- State the origin and Applications of FEM
- Describe the Basics of FEM
- Define Data modeling
- Describe the different types and the three perspectives of data models
- Have an overview of database models

3.0 Main Content

3.1 Finite Element Model

Many physical phenomena in engineering and science can be described in terms of partial differential equations (PDE). In general, solving these equations by classical analytical methods for arbitrary shapes is almost impossible. The finite element method (FEM) is a numerical approach by which these PDE can

be solved approximately.
The FEM is a function/basis-based approach to solve PDE. FEs are widely used in diverse fields to solve static and dynamic problems − Solid or fluid mechanics, electromagnetic, biomechanics, etc.

The **finite element method (FEM)** (its practical application often known as **finite elementanalysis** (FEA)) is a numerical technique for finding approximate solutions of partial differential equations (PDE) as well as of integral equations. The solution approach isbased either on eliminating the differential equation completely (steady state problems), or rendering the PDE into an approximating system of ordinary differential equations, which are then numerically integrated using standard techniques such as Euler's method, Runge- Kutta, etc.

In solving partial differential equations, the primary challenge is to create an equation that approximates the equation to be studied, but is numerically stable, meaning that errors in the input and intermediate calculations do not accumulate and cause the resulting output to be meaningless. There are many ways of doing this, all with advantages and disadvantages.The steps may be broken down as follows:

1. Definition of the physical problem:- development of the model.
2. Formulation of the governing equations - Systems of PDE, ODE, algebraic equations,define initial conditions and/or boundary conditions to get a well-posed problem,
3. Discretization of the equations.
4. Solution of the discrete system of equations.
5. Interpretation of the obtained results.
6. Errors analysis.

The Finite Element Method is a good choice for solving partial differential equations over complicated domains (like cars and oil pipelines), when the domain changes (as during a solid state reaction with a moving boundary), when the desired precision varies over the entire domain, or when the solution lacks smoothness. For instance, in a frontal crash simulation it is possible to increase prediction accuracy in "important" areas like the front of the car and reduce it in its rear (thus reducing cost of the simulation); another example would be the simulation of the weather pattern on Earth, where it is more important to haveaccurate predictions over land than over the wide-open sea.

The Finite Element Analysis

Finite Element Analysis is a method to computationally model reality in a mathematical form to better understand a highly complex problem. In the real world *everything* that occurs is as a results of interactions between atoms (and sub-particles of those atoms), billions and billions of them. If we were to simulate the world in a computer, we would have to simulate this interaction based on the simple laws of physics. However, no computer can process the near infinite number of atoms in objects, so instead we model 'finite' groups of them.

For example, we might model a gallon of water by dividing it up into 1000 parts and measuring the interaction of these linked parts. If you divide into too few parts, your simulation will be too inaccurate. If you divide into too many, your computer will sit therefor years calculating the result!

3.1.1 Why use FEA?

Simulation in general is always a good idea, as it lets you test designs and ideas without spending money or effort actually building anything. By using simulation, you can find fault points within your designs, simulate ideas as you think of them, and even quantitize and optimize them. One can even use simulation to verify theories - if the theoretical simulation matches what actually happens, then the theory is proven!

Sometimes you can hand calculate certain designs. But sometimes a design can be toocomplex, making FEA great for non-symmetric problems with ultra-complicated geometries.

3.1.2 History of FEM

The finite element method originated from the need for solving complex elasticity and structural analysis problems in civil and aeronautical engineering. The development can betraced back to the work by Alexander Hrennikoff (1941) and Richard Courant (1942). While the approaches used by these pioneers are dramatically different, they share one essential characteristic: mesh discretization of a continuous domain into a set of discrete sub-domains, usually called elements.

Hrennikoff's work discretizes the domain by using a lattice analogy while Courant's approach divides the domain into finite triangular subregions for solution of second order elliptic partial differential equations (PDEs) that arise from the problem of torsion of a cylinder. Courant's contribution was evolutionary, drawing on a large body of earlier results for PDEs developed by Rayleigh, Ritz, and Galerkin.
Development of the finite element method began in earnest in the middle to late 1950s for airframe and structural analysis and gathered momentum at the University of Stuttgart through the work of John Argyris and at Berkeley through the work of Ray W. Clough in the 1960s for use in civil engineering. By late 1950s, the key concepts of stiffness matrix and element assembly existed essentially in the form used today. NASA issued a request for proposals for the development of the finite element software NASTRAN in 1965. The method was again provided with a rigorous mathematical foundation in 1973 with the publication of Strang and Fix's *An Analysis of The Finite Element Method* has since been generalized into a branch of applied mathematics for numerical modelling of physical systems in a wide variety of engineering disciplines, e.g., electromagnetism, thanks to Peter P. Silvester and fluid dynamics.

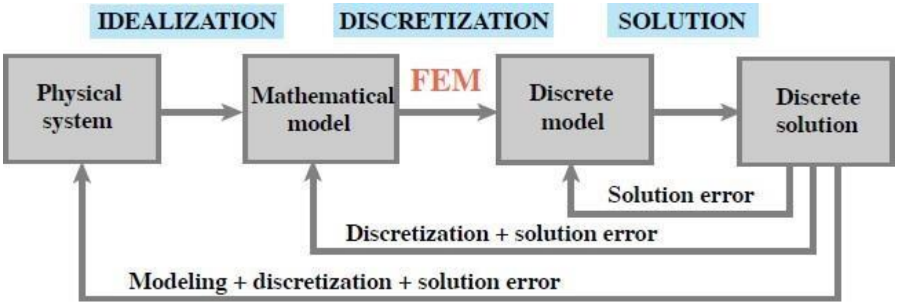
3.1.3 Applications of FEM

A variety of specializations under the umbrella of the mechanical engineering discipline (such as aeronautical, biomechanical, and automotive industries) commonly use integratedFEM in design and development of their products. Several modern FEM packages include specific components such as thermal, electromagnetic, fluid, and structural working environments. In a structural simulation, FEM helps tremendously in producing stiffness and strength visualizations and also in minimizing weight, materials, and costs.

3.2 Overview of Basic FEM

The basic steps of the finite element method are discussed next in more generality. Although attention is focused on structural problems, most of the steps translate to other applications problems as noted above. he role of FEM in numerical simulation is schematized in (**Figure 10**). below. This diagram displays the three key simulation steps: idealization, discretization and solution. It also indicates the fact that each step introduces different of errors. For example the discretization error is the discrepancy obtained when the discrete solution is substitutedin the mathematical model.

Figure10: Steps of the physical simulation process: idealization, discretization and solution.



3.2.1 Idealization

a. Models

The word ―model‖ has the traditional meaning of a scaled copy or representation of an object. And that is precisely how most dictionaries define it. We use here the term in a more modern sense, increasingly common since the advent of computers:

A model is a symbolic device built to simulate and predict aspects of behaviour of asystem.
Note the careful distinction made between ―behaviour‖ and ―aspects of behaviour.‖ To predict everything, in all physical scales, you must deal with the actual system.

A model *abstracts* aspects of interest to the modeler. The term ―symbolic means that a model represents a system in terms of the symbols and language of another science. For example, engineering systems may be (and are) modeled with the symbols of mathematics and/or computer sciences.

b. Mathematical Models

Mathematical modelling, or *idealization*, is a process by which the engineer passes fromthe actual physical system under study, to a *mathematical model* of the system, where the term *model* is understood in the wider sense defined above.

The process is called *idealization* because the mathematical model is necessarily an abstraction of the physical reality. (Note the phrase *aspects of behaviour* in the definition.) The analytical or numerical results obtained for the mathematical model are re-interpreted in physical terms only for those aspects.

Why is the mathematical model an abstraction of reality?

Engineering systems such as structures tend to be highly complex. To simulate its behaviour it is necessary to reduce that complexity to manageable proportions. Mathematical modelling is an abstraction tool by which complexity can be brought under control. This is achieved by ―filtering out‖ physical details that are not relevant to the analysis process. For example, a continuum material model necessarily filters out the aggregate, crystal, molecular and atomic levels of matter. If you are designing a bridge or building such levels are irrelevant. Consequently, choosing a mathematical model is equivalent to choosing an information filter.

3.2.2 Implicit vs. Explicit Modelling

Suppose that you have to analyze a structure and at your disposal is a ―black box‖ general- purpose finite element program. This is also known in the trade as a ―canned program.‖ Those programs usually offer a *catalog* of element types; for example, bars, beams, plates, shells, axisymmetric solids, general 3D solids, and so on. The moment you choose specific elements from the catalog you automatically accept

the mathematical models on which the elements are based. This is *implicit modelling*, a process depicted in Figure 2.
Ideally you should be fully aware of the implications of your choice. Providing such
–finite element literacy‖ is one of the objectives of this course.

Unfortunately many users of commercial programs are unaware of the –implied consent‖ aspect of implicit modelling.

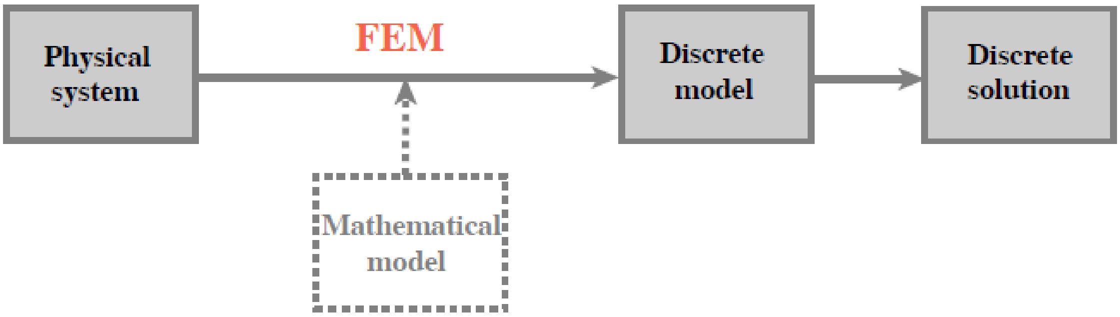


Figure 2: Implicit modelling: picking elements from an existing FEM code implicitlyaccepts an idealization. Read the fine print.

The other extreme occur when you select a mathematical model of the physical problem with your eyes wide open and *then* either shop around for finite element programs that implements that model, or write the program yourself. This is *explicit modelling*. It requires far more technical expertise, resources, experience and maturity than implicit modelling. But for problems that fall out of the ordinary it may be the right thing to do.

In practice a combination of implicit and explicit modelling is quite common. The physicalproblem to be solved is broken down into subproblems. Those subproblems that are conventional and fit existing programs may be treated with implicit modelling. Those subproblems that require special handling may yield only to explicit modelling treatment.

3.3 Discretization

3.3.1 Purpose

Mathematical modelling is a simplifying step. But models of physical systems are not necessarily easy to solve. They usually involve coupled partial differential equations in space and time subject to boundary and/or interface conditions. Such analytical models have an *infinite* number of degrees of freedom.

At this point one faces the choice of trying for analytical or numerical solutions. Analyticalsolutions, also called –closed form solutions, are more intellectually satisfying, particularly if they apply to a wide class of problems. Unfortunately they tend to be restricted to regular geometries and simple boundary conditions. Moreover a closed- form solution, expressed for example as the inverse of an integral transform, often has to be numerically evaluated to be useful.
Most problems faced by the engineer either do not yield to analytical treatment or doing sowould require a disproportionate amount of effort. The practical way out is numerical simulation. Here is where finite element methods and the digital computer enter the scene.

To make numerical simulations practical it is necessary to reduce the number of degrees of freedom to a *finite* number. The reduction is called *discretization*. The end result of thediscretization process is the *discrete model* depicted in Figures 1 and 2.Discretization can proceed in space dimensions as well as in the time dimension. Because the present course deals only with static problems, we need not consider the time dimension and are free to concentrate on *spatial discretization*.

3.3.2 Error Sources and Approximation

Figure 1 tries to convey graphically that each simulation step introduces a source of error. In engineering practice modelling errors are by far the most important. But they aredifficult and expensive to evaluate, because such *model validation* requires access to and comparison with experimental results.

Next in order of importance is the *discretization error*. Even if solution errors are ignored and usually they can, the computed solution of the discrete model is in general only an approximation in some sense to the exact solution of the mathematical model. A quantitative measurement of this discrepancy is called the *discretization error*.
The characterization and study of this error is addressed by a branch of numericalmathematics called approximation theory.

Intuitively one might suspect that the accuracy of the discrete model solution would improve as the number of degrees of freedom is increased, and that the discretization error goes to zero as that number goes to infinity. This loosely worded statement describes the *convergence* requirement of discrete approximations. One of the key goals of approximation theory is to make the statement as precise as it can be expected from a branch of mathematics.

3.3.3 Finite and Boundary Element Methods

The most popular discretization procedures in structural mechanics are finite elementmethods and boundary element methods. The finite element method (FEM) is by far the most widely used. The boundary element method (BEM) has gained in popularity for special types of problems, particularly those involving infinite domains, but remains a distant second.
In non-structural application areas such as fluid mechanics and thermal analysis, the finite element method is gradually making up ground but faces stiff competition from both the classical and energy-based *finite difference* methods. Finite difference and finite volume methods are particularly well entrenched in computational fluid dynamics.

3.4 Interpretation of FEM

The finite element method (FEM) is the dominant discretization technique in structural mechanics. FEM can be interpreted from either a physical or mathematical standpoint. Thetreatment has so far emphasized the former.

The basic concept in the physical interpretation of the FEM is the subdivision of the mathematical model into disjoint (non-overlapping) components of simple geometry called*finite elements* or *elements* for short. The response of each element is expressed in terms of a finite number of degrees of freedom characterized as the value of an unknown function, or functions, at a set of nodal points.

The response of the mathematical model is then considered to be approximated by that of the discrete model obtained by connecting or assembling the collection of all elements.

The disconnection-assembly concept occurs naturally when examining many artificial and natural systems. For example, it is easy to visualize an engine, bridge, building, airplane, orskeleton as fabricated from simpler components.
Unlike finite difference models, finite elements *do not overlap* in space. In the mathematical interpretation of the FEM, this property goes by the name *disjoint support*.

FEM Element Attributes

Just like the members in the truss example, one can take finite elements of any kind one ata time. Their local properties can be developed by considering them in isolation, asindividual entities. This is the key to the programming of element libraries.

In the Direct Stiffness Method, elements are isolated by disconnection and localization. This procedure involves the separation of elements from their neighbors by disconnecting the nodes, followed by the referral of the element to a convenient local coordinate system. After these two steps we can consider *generaic* elements: a bar element, a beam element, and so on. From the standpoint of computer implementation, it means that you can write one subroutine or module that constructs all elements of one type, instead of writing one for each element instance.

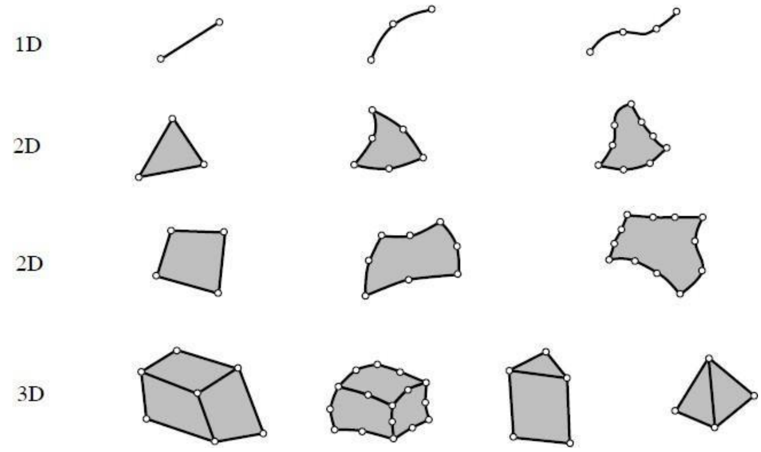


Figure 3: Typical finite element geometries in one through three dimensions

The following is a summary of the data associated with an individual finite element. This data is used in finite element programs to carry out element level calculations.

Dimensionality. Elements can have one, two or three space dimensions. (There are also special elements with zero dimensionality, such as lumped springs.)

Nodal points. Each element possesses a set of distinguishing points called *nodal points* or *nodes* for short. Nodes serve two purposes: definition of element geometry, and home for degrees of freedom. They are located at the corners or end points of elements (see Figure 3); in the so-called refined or higher-order elements nodes are also placed on sides or faces.

Geometry. The geometry of the element is defined by the placement of the nodal points. Most elements used in practice have fairly simple geometries. In one-dimension, elements are usually straight lines or curved segments. In two dimensions they are of triangular or quadrilateral shape.
In three dimensions the three common shapes are tetrahedra, pentahedra (also calledwedges or prisms), and hexahedra (also called cuboids or ―bricks‖). See Figure 3.

Degrees of freedom. The degrees of freedom (DOF) specify the *state* of the element. They also function as ―handles‖ through which adjacent elements are connected. DOFs are defined as the values (and possibly derivatives) of a primary field variable at nodal points.

The actual selection depends on criteria studied at length in Part II. Here we simply note that the key factor is the way in which the primary variable appears in the mathematical model. For mechanical elements, the primary variable is the displacement field and the DOF for many (but not all) elements are the displacement components at the nodes.

Nodal forces. There is always a set of nodal forces in a one-to-one correspondence with degrees of freedom. In mechanical elements the correspondence is established through energy arguments.

Constitutive properties. For a mechanical element these is the relation that specifies the material properties. For example, in a linear elastic bar element it is sufficient to specifythe elastic modulus E and the thermal coefficient of expansion.

Fabrication properties. For a mechanical element these are fabrication properties which have been integrated out from the element dimensionality. Examples are cross sectional properties of MoM elements such as bars, beams and shafts, as well as the thickness of a plate or shell element.
This data is used by the element generation subroutines to compute element stiffnessrelations in the local system.

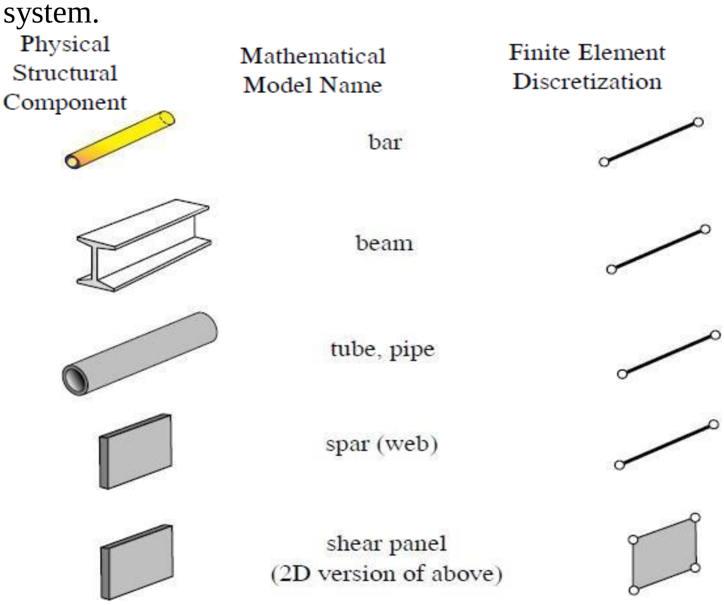


Figure 4. Examples of primitive structural elements

3.5 Assembly Procedure

The assembly procedure of the Direct Stiffness Method for a general finite element model follows rules identical in principle to those discussed for the truss example. As in that casethe process involves two basic steps:
Globalization. The element equations are transformed to a common *global* coordinate system.
Merge. The element stiffness equations are merged into the master stiffness equations by appropriate indexing and entry addition.

The computer implementation of this process is not necessarily as simple as the hand calculations of the truss example suggest. The master stiffness relations in practical cases may involve thousands (or even millions) of degrees of freedom. To conserve storage and processing time the use of sparse matrix techniques as well as peripheral storage is required. But this inevitably increases the programming complexity.

3.6 Boundary Conditions

A key strength of the FEM is the ease and elegance with which it handles arbitrary boundary and interface conditions. This power, however, has a down side. One of the biggest hurdles a FEM newcomer faces is the understanding and proper handling ofboundary conditions. Surprisingly, prior exposure to partial differential equations, withouta balancing study of variational calculus, does not appear to be of much help in this regard.

In the present Section we summarize some basic rules for treating boundary conditions.

Essential and Natural B.C.

The important thing to remember is that boundary conditions (BCs) come in two basicflavors:
Essential BCs are those that directly affect the degrees of freedom, and are imposed on theleft-hand side vector **u**.
Natural BCs are those that do not directly affect the degrees of freedom, and are imposedon the right-hand side vector **f**.

The mathematical justification for this distinction requires use of the variation calculus, and is consequently relegated to Part II of the course. For the moment, the basic recipe is:

1. If a boundary condition involves one or more degrees of freedom in a *direct* way, itis essential. An example is a prescribed node displacement.
2. Otherwise it is natural.

The term ―direct‖ is meant to exclude derivatives of the primary function, unless thosederivatives also appear as degrees of freedom, such as rotations in beams and plates.

3.6.2 Boundary Conditions in Structural Problems

In mechanical problems, essential boundary conditions are those that involve *displacements* (but not strain-type displacement derivatives). The support conditions forthe truss problem furnish a particularly simple example. But there are more general boundary conditions that occur in practice.

A structural engineer must be familiar with displacement B.C. of the following types.
Ground or support constraints. Directly restraint the structure against rigid body motions.

Symmetry conditions. To impose symmetry or antisymmetry restraints at certain points, lines or planes of structural symmetry. This allows the discretization to proceed only overpart of the structure with a consequent savings in the number of equations to be solved.
Ignorable freedoms. To suppress displacements that are irrelevant to the problem. (In classical dynamics these are called *ignorable coordinates*.) Even experienced users offinite element programs are sometimes baffled by this kind.

Connection constraints. To provide connectivity to adjoining structures or substructures,or to specify relations between degrees of freedom. Many conditions of this type fall under the label.

multipoint constraints or *multifreedom constraints*, which can be notoriously difficult tohandle from a numerical standpoint.

FEM allows detailed visualization of where structures bend or twist, and indicates the distribution of stresses and displacements. FEM software provides a wide range of simulation options for controlling the complexity of both modelling and analysis of a system. Similarly, the desired level of accuracy required and associated computational time requirements can be managed simultaneously to address most engineering applications. FEM allows entire designs to be constructed, refined, and optimized before the design is manufactured.

This powerful design tool has significantly improved both the standard of engineering designs and the methodology of the design process in many industrial applications. The introduction of FEM has substantially decreased the time to take products from concept to the production line. It is primarily through improved initial prototype designs using FEM that testing and development have been accelerated. In summary, benefits of FEM include increased accuracy, enhanced design and better insight into critical design parameters, virtual prototyping, fewer hardware prototypes, a faster and less expensive design cycle, increased productivity, and increased revenue.

3.7 Databased models

Data modelling is a method used to define and analyze data requirements needed to support the business processes of an organization. The data requirements are recorded as a conceptual data model with associated data definitions. Actual implementation of theconceptual model is called a logical data model. To implement one conceptual data model may require multiple logical data models.

Data modelling defines not just data elements, but their structures and relationships between them. Data modelling techniques and methodologies are used to model data in a standard, consistent, predictable manner in order to manage it as a resource. The use of data modelling standards is strongly recommended for all projects requiring a standard means of defining and analyzing data within an organization, e.g., using data modelling:

- to manage data as a resource;
- for the integration of information systems;
- for designing databases/data warehouses (aka data repositories)

Data modelling may be performed during various types of projects and in multiple phases of projects. Data models are progressive; there is no such thing as the final data model for abusiness or application. Instead a data model should be considered a living document that will change in response to a changing business. The data models should ideally be stored ina repository so that they can be retrieved, expanded, and edited over time.

Types of Data Models

Whitten (2004) determined two types of data modelling:

- **Strategic data modelling**: This is part of the creation of an information systems strategy, which defines an overall vision and architecture for information systems isdefined. Information engineering is a methodology that embraces this approach.
- **Data modelling during systems analysis**: In systems analysis logical data modelsare created as part of the development of new databases.

  Data modelling is also a technique for detailing business requirements for a database. It is sometimes called *database modelling* because a data model is eventually implemented in a database.

  The main aim of data models is to support the development of data-base by providing the definition and format of data. According to West and Fowler (1999) "if this is done consistently across systems then compatibility of data can be achieved. If the same data structures are used to store and access data then different applications can share data.

However, systems and interfaces often cost more than they should, to build, operate, and maintain. They may also constrain the business rather than support it. A major cause is thatthe quality of the data models implemented in systems and interfaces is poor. As aconsequence:

- Business rules, specific to how things are done in a particular place, are often fixed in the structure of a data model. This means that small changes in the way business is conducted lead to large changes in computer systems and interfaces
- Entity types are often not identified, or incorrectly identified. This can lead to replication of data, data structure, and functionality, together with the attendant costs of that duplication in development and maintenance
- Data models for different systems are arbitrarily different. The result of this is that complex interfaces are required between systems that share data. These interfaces can account for between 25-70% of the cost of current systems
- "Data cannot be shared electronically with customers and suppliers, because the structure and meaning of data has not been standardised. For example, engineering design data and drawings for process plant are still sometimes exchanged on paper

The reason for these problems is a lack of standards that will ensure that data models will both meet business needs and be consistent.

3.8 The three perspectives of Data model

The perspectives shows that a data model can be an external model (or view), a conceptual model, or a physical model. This is not the only way to look at data models, but it is a useful way, particularly when comparing models.

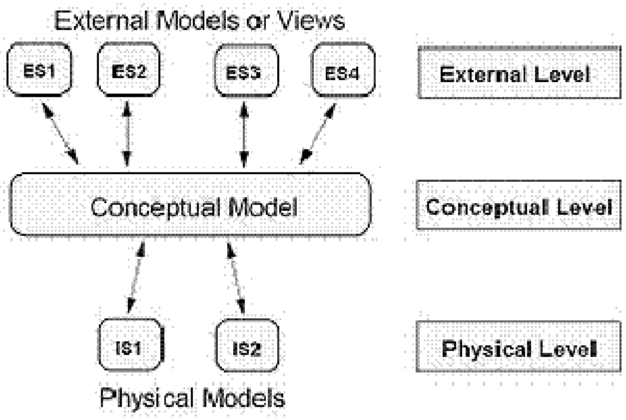


Fig. 6: ANSI/SPARC **three level architecture**.

A data model *instance* may be one of three kinds according to ANSI in 1975:

Conceptual schema - A conceptual schema specifies the kinds of facts or propositions thatcan be expressed using the model. In that sense, it defines the allowed expressions in an artificial 'language' with a scope that is limited by the scope of the model. It describes the semantics of a domain. For example, it may be a model of the interest area of an organization or industry. This consists of entity classes, representing kinds of things of significance in the domain, and relationships assertions about associations between pairs of entity classes The use of conceptual schema has evolved to become a powerful communication tool with business users. Often called a subject area model (SAM) or high-level data model (HDM), this model is used to communicate core data concepts, rules, and definitions to a business user as part of an overall application development or enterprise initiative. The number of objects should be very small and focused on key concepts. Try to limit this model to one page, although for extremely large organizations or complex projects, the model might span two or more pages

Logical schema - describes the semantics, as represented by a particular data manipulationtechnology. This consists of descriptions of tables and columns, object oriented classes, and XML tags, among other things.

Physical schema - describes the physical means by which data are stored. This is concerned with partitions, CPUs, tablespaces, and the like.

The significance of this approach, according to ANSI, is that it allows the three perspectives to be relatively independent of each other.

Storage technology can change without affecting either the logical or the conceptual model. The table/column structure can change without (necessarily) affecting the conceptual model. In each case, of course, the structures must remain consistent with the other model. The table/column structure may be different from a direct translation of the entity classes and attributes, but it must ultimately carry out the objectives of the conceptual entity class structure. Early phases of many software development projects emphasize the design of a **conceptual data model** . Such a design can be detailed into a **logical data model** . In later stages, this model may be translated into **physical data model**. However, it is also possible to implement a conceptual model directly

3.9 Database Model

A **database model** is a theory or specification describing how a database is structuredand used. Several such models have been suggested. Common models include:

Flat model: This may not strictly qualify as a data model. The flat (or table) model consistsof a single, two-dimensional array of data elements, where all members of a given column are assumed to be similar values, and all members of a row are assumed to be related to one another.

Hierarchical model: In this model data is organized into a tree-like structure, implying a single upward link in each record to describe the nesting, and a sort field to keep the records in a particular order in each same-level list.

Network model: This model organizes data using two fundamental constructs, calledrecords and sets. Records contain fields, and sets define one-to-many relationships betweenrecords: one owner, many members.

Relational model: is a database model based on first-order predicate logic. Its core idea isto describe a database as a collection of predicates over a finite set of predicate variables, describing constraints on the possible values and combinations of values.

Object-relational model: Similar to a relational database model, but objects, classes andinheritance are directly supported in database schemas and in the query language.

Modelling & Simulation - Database

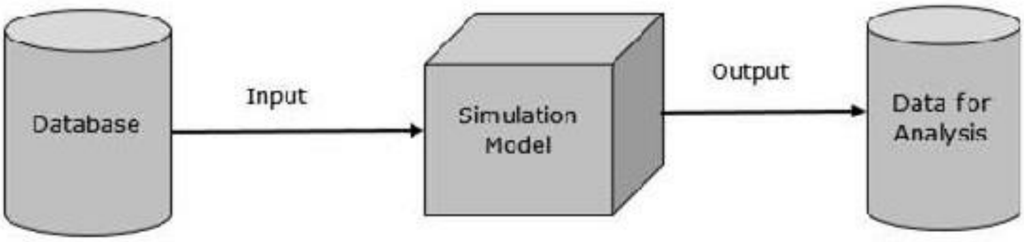
The objective of the database in Modelling & Simulation is to provide data representation and its relationship for analysis and testing purposes. The first data model was introduced in 1980 by Edgar Codd. Following were the salient features of the model.

- Database is the collection of different data objects that defines the information and their relationships.
- Rules are for defining the constraints on data in the objects.
- Operations can be applied to objects for retrieving information.

Initially, Data Modelling was based on the concept of entities & relationships in which the entities are types of information of data, and relationships represent the associations between the entities.

The latest concept for data modeling is the object-oriented design in which entities are represented as classes, which are used as templates in computer programming. A class having its name, attributes, constraints, and relationships with objects of other classes.

Its basic representation looks like −



Data Representation

Data Representation for Events

A simulation event has its attributes such as the event name and its associated time information. It represents the execution of a provided simulation using a set of input data associated with the input file parameter and provides its result as a set of output data, stored in multiple files associated with data files.

Data Representation for Input Files

Every simulation process requires a different set of input data and its associated parameter values, which are represented in the input data file. The input file is associated with the software which processes the simulation. The data model represents the referenced files by an association with a data file.

Data Representation for Output Files

When the simulation process is completed, it produces various output files and each output file is represented as a data file. Each file has its name, description and a universal factor. A data file is classified into two files. The first file contains the numerical values and the second file contains the descriptive information for the contents of the numerical file.

4.0 Self-Assessment Exercise(s)

Answer the following questions:

1. Draw the FEM physical simulation process

2. Briefly discuss each of the data assciated with FE used in programs for elementarycalculations.
3. With the aid of diagrams differentiate between the common data model
4. Briefly discuss the basic rules for the treatment of boundary conditions
5. What is FEM and how/where can it be applied?
6. How does the ANSI three perspectives to data model allows for relativelyindependent of each.
7. What is the purpose of FEM discretization

5.0 Conclusion

The development of systems and interfaces often cost more than they should, to build, operate, and maintain. A major cause is that the quality of the data models implemented in systems and interfaces is poor. This usually is as a result of:

- Violation of business rules, as a result small changes in the way business is conducted lead to large changes in computer systems and interfaces.
- Unidentified or incorrect identification of entity which can lead to replication of data, data structure, and functionality, and increased costs of development and maintenance

Consequently data cannot be shared electronically with customers and suppliers due tounstructured and lack of standard data that can meet business needs.

6.0 Summary

In this unit, we have discussed elaborately,

- the Physics-based Finite Element Method (FEM) which is defined as a numerical technique for finding approximate solutions of partial differential equations (PDE) aswell as of integral equations.
  - Here we stated the uses, traced its origin and discussed its aplications
  - Carried out the overview of the FEM basics including:
    - Idealization
    - Discretization; purpose and error sources
    - Finite and Boundary element methods
    - Interpretations of FEM and
    - Elementary attributes used in FEM
- Stated the three perspectives of data models; Conceptual, logical and Physicalschemas.
- The different types of database models; Flat, hierarchical, network, Relational andObject oriented models.

7.0 Further Readings

- Gordon, S. I., & Guilfoos, B. (2017). *Introduction to Modeling and Simulation withMATLAB® and Python*. Milton: CRC Press.
- Zeigler, B. P., Muzy, A., & Kofman, E. (2019). *Theory of modeling and simulation: Discrete event and iterative system computational foundations*. San Diego (Calif.):Academic Press.
- Law, A. M. (2015). *Simulation modeling and analysis*. New York: McGraw-Hill.
- Verschuuren, G. M., & Travise, S. (2016). *100 Excel Simulations: Using Excel to ModelRisk, Investments, Genetics, Growth, Gambling and Monte Carlo Analysis*. Holy Macro! Books.
- Grigoryev, I. (2015). *AnyLogic 6 in three days: A quick course in simulation modeling*.Hampton, NJ: AnyLogic North America.
- Dimotikalis, I., Skiadas, C., & Skiadas, C. H. (2011). *Chaos theory: Modeling, simulationand applications: Selected papers from the 3rd Cghaotic Modeling and Simulation International* Unit

Unit 4: Statistics for Modelling and Simulation

Contents

1.0 Introduction

In this unit we will discuss two ways statistics are computed and applied in modelling and simulations these include: inference and descriptive processes. Statistical inference is generally distinguished from descriptive statistics. In simple terms, descriptive statistics can be thought of as being just a straightforward presentation of facts, in which modelling decisions made by a data analyst have had minimal influence. Statistical inference is the process of drawing conclusions from data that are subject to random variation, for example, observational errors or sampling variation. A complete statistical analysis will nearly always include both descriptive statistics and statistical inference, and will often progress in a series of steps where the emphasis moves gradually from description to inference.

2.0 Intended Learning Outcomes (ILOs)

By the end of this unit you should be able to:

- Differentiate between Descriptive and Inference statistics
- Describe the features of descriptive statistics
- Describe features of Inference statistics
- Compute the essential statistics for simulation

3.0 Main Content

3.1 Descriptive and Inference statistics

Descriptive statistics describe the main features of a collection of data quantitatively. Descriptive statistics are distinguished from inferential statistics (or inductive statistics), in that descriptive statistics aim to summarize a data set quantitatively without employing a probabilistic formulation, rather than use the data to make inferences about the
population that the data are thought to represent. Even when a data analysis draws its main conclusions using inferential statistics, descriptive statistics are generally also presented. For example in a paper reporting on a study involving human subjects, there typically appears a table giving the overall sample size, sample sizes in important subgroups (e.g., for each treatment or exposure group), and demographic or clinical characteristics such as the averageage, the proportion of subjects of each sex, and the proportion of subjects with related co morbidities.

Inferential statistics tries to make inferences about a population from the sample data. We also use inferential statistics to make judgments of the probability that an observed differencebetween groups is a dependable one, or that it might have happened by chance in this study. Thus, we use inferential statistics to make inferences from our data to more general conditions; we use descriptive statistics simply to describe what's going on in our data.

3.1 Descriptive Statistics

Descriptive statistics provide simple summaries about the sample and the measures. Togetherwith simple graphics analysis, they form the basis of quantitative analysis of data. Descriptive statistics summarize data. For example, the shooting percentage in basketball is adescriptive statistics that summarizes the performance of a player or a team. The percentageis the number of shots made divided by the number of shots taken. A player who shoots 33% is making approximately one shot in every three. One making 25% is hitting once in four. The percentage summarizes or describes multiple discrete events. Or, consider the score of many students, the grade point average. This single number describes the general performance of a student across the range of their course experiences.

One that describes a large set of observations with a single indicator risks distorting the original data or losing important detail. For example, the shooting percentage doesn't tell you whether the shots are three-pointers or lay-ups, and GPA doesn't tell you whether the student was in difficult or easy courses. Despite these limitations, descriptive statistics provide a powerful summary that may enable comparisons across people or other units.

3.1.1 Univariate Analysis

Univariate analysis involves the examination across cases of a single variable, focusing on three characteristics: the distribution; the central tendency; and the dispersion. It is common to compute all three for each study variable.

a. Distribution

The distribution is a summary of the frequency of individual or ranges of values for a variable. The simplest distribution would list every value of a variable and the number of cases that had that value. For instance, computing the distribution of gender in the study population means computing the percentages that are male and female. The gender variable has only two, making it possible and meaningful to list each one. However, this does not work for a variable such as income that has many possible values. Typically, specific values are not particularly meaningful (income of 50,000 is typically not meaningfully different from 51,000). Grouping the raw scores using ranges of values reduces the number of categories to something more meaningful. For instance, we might group incomes into ranges of 0-10,000, 10,001-30,000, etc.

b. Central tendency

The central tendency of a distribution locates the "center" of a distribution of values. The three major types of estimates of central tendency are the *mean*, the *median*, and the *mode*.

The **mean** is the most commonly used method of describing central tendency. To computethe mean, take the sum of the values and divide by the count. For example, the mean quiz score is determined by summing all the scores and dividing by the number of students taking the exam. For example, consider the test score values:

15, 20, 21, 36, 15, 25, 15

The sum of these 7 values is 147, so the mean is 147/7 =21.

The mean is computed using the formular: □ X_i / n, where the sum is over i = 1 to n.

The **median** is the score found at the middle of the set of values, i.e., that has as many cases with a larger value as have a smaller value. One way to compute the median is to sort the values in numerical order, and then locate the value in the middle of the list. For example, if there are 500 values, the value in 250th position is the median. Sorting the 8 scores above produces:

15, 15, 15, 20, 21, 25, 36

There are 7 scores and score #4 represents the halfway point. The median is 20. If there is an even number of observations, then the median is the mean of the two middle scores. In the example, if there were an 8th observation, with a value of 25, the median becomes the average of the 4th and 5th scores, in this case 20.5:

15, 15, 15, 20, 21, 25, 25, 36

The **mode** is the most frequently occurring value in the set. To determine the mode, compute the distribution as above. The mode is the value with the greatest frequency. In the example, the modal value 15, occurs three times. In some distributions there is a "tie" for the highest frequency, i.e., there are multiple modal values. These are called **multi- modal** distributions.

Notice that the three measures typically produce different results. The term "average" obscures the difference between them and is better avoided. The three values are equal if the distribution is perfectly "**normal**" (i.e., bell-shaped).

Dispersion

Dispersion is the spread of values around the central tendency. There are two common measures of dispersion, the range and the standard deviation. The **range** is simply the highestvalue minus the lowest value. In our example distribution, the high value is 36 and the low is 15, so the range is 36 − 15 = 21.

The **standard deviation** is a more accurate and detailed estimate of dispersion because an outlier can greatly exaggerate the range (as was true in this example where the single outlier value of 36 stands apart from the rest of the values). The standard deviation shows the relation that set of scores has to the mean of the sample. Again let's take the set of scores:

15, 20, 21, 36, 15, 25, 15

to compute the standard deviation, we first find the distance between each value and the mean. We know from above that the mean is 21. So, the differences from the mean are:

15 − 21 = −6

20 − 21 = −1

21 − 21 = 0

36 − 21 = 15
15 − 21 = −6
25 − 21 = +4
15 − 21 = −6

Notice that values that are below the mean have negative differences and values above ithave positive ones. Next, we square each difference:
$(6)^2 = 36$
$(-1)^2 = 1$
$(+0)^2 = 0$
$(15)^2 = 225$
$(-6)^2 = 36$
$(+4)^2 = 16$
$(-6)^2 = 36$
Now, we take these "squares" and sum them to get the **sum of squares** (SS) value. Here, thesum is 350. Next, we divide this sum by the number of scores minus 1. Here, the result is 350 / 6 = 58.3. This value is known as the **variance**. To get the standard deviation, we take thesquare root of the variance (remember that we squared the deviations earlier). This would be $\sqrt{58.3} = 7.63$.

Although this computation may seem intricate, it's actually quite simple. In English, we candescribe the standard deviation as:
the square root of the sum of the squared deviations from the mean divided by the number ofscores minus one given as: $\sqrt{}(\square\ (x_i - u)^2)/\ n$; where x = observed value and u = the mean

The standard deviation allows us to reach some conclusions about specific scores in ourdistribution. Assuming that the distribution of scores is close to "normal", the following conclusions can be reached:

a. approximately 68% of the scores in the sample fall within one standard deviation ofthe mean (u - SD) and (u + SD)
b. approximately 95% of the scores in the sample fall within two standard deviations ofthe mean (u-2SD) and (u+2SD)
c. approximately 99% of the scores in the sample fall within three standard deviations ofthe mean (u - 3SD) and (u + 3SD)

For example, since the mean in our example is 21 and the standard deviation is 7.63, we can from the above statement estimate that approximately 95% of the scores will fall in the range of 21 − (2×7.63) to 21 + (2×7.63) or between 5.74 and 36.26. Values beyond two standard deviations from the mean can be considered "outliers". 36 is the only such value in our distribution.

Outliers help identify observations for further analysis or possible problems in the observations. Standard deviations also convert measures on very different scales, such as height and weight, into values that can be compared.

d. Other Statistics

In research involving comparisons between groups, emphasis is often placed on the **significance level** for the **hypothesis** that the groups being compared differ to a degree greater than would be expected by chance. This significance level is often represented as a **p-value**, or sometimes as the standard score of a test statistic. In contrast, an **effect size** conveys the estimated magnitude and direction of the difference between groups, without regard to whether the difference is statistically significant. Reporting significance levels without effect sizes is problematic, since for large sample sizes even small effects of little practical importance can be statistically significant.

3.1.2 Examples of descriptive statistics

Most statistics can be used either as a descriptive statistic, or in an inductive analysis. For example, we can report the average reading test score for the students in each classroom in aschool, to give a descriptive sense of the typical scores and their variation. If we perform a formal *hypothesis test* on the scores, we are doing *inductive* rather than descriptive analysis.

The following is a list some statistical common in descriptive analyses:Measures of central tendency
Measures of dispersion Measures of association
Cross-tabulation, contingency tableHistogram
Quantile, Q-Q plotScatter plot
Box plot

3.3 Inference Statistics

The terms **statistical inference**, **statistical induction** and **inferential statistics** are used to describe systems of procedures that can be used to draw conclusions from datasets arising from systems affected by random variation. Initial requirements of such a system of procedures for *inference* and *induction* are that the system should produce reasonable answers when applied to well-defined situations and that it should be general enough to be applied across a range of situations.

The outcome of statistical inference may be an answer to the question "what should be done next?", where this might be a decision about making further experiments or surveys, or aboutdrawing a conclusion before implementing some organizational or governmental policy.

For the most part, statistical inference makes propositions about populations, using data drawn from the population of interest via some form of random sampling. More generally, data about a random process is obtained from its observed behaviour during a finite period of time. Given a parameter or hypothesis about which one wishes to make inference, statistical inference most often uses:

- a statistical model of the random process that is supposed to generate the data, and
- a particular realization of the random process; i.e., a set of data.The conclusion of a **statistical inference** is a statistical **proposition**.

3.3.1 Some common forms of statistical proposition

- An **estimate** - a particular value that best approximates some parameter of interest,
- A **confidence interval** (or set estimate) - an interval constructed from the data in sucha way that, under repeated sampling of datasets, such intervals would contain the true parameter value with the probability at the stated confidence level,
- A **credible interval** - a set of values containing, for example, 95% of posterior belief,

- Rejection of an **hypothesis**
- **Clustering** or classification of data points into groups

3.3.2 Models/Assumptions

Any statistical inference requires some assumptions. A **statistical model** is a set of assumptions concerning the generation of the observed data and similar data. Descriptions of statistical models usually emphasize the role of population quantities of interest, about which we wish to draw inference.

3.3.3 Degree of models/assumptions

Statisticians distinguish between three levels of modelling assumptions;

- **Fully parametric**: The probability distributions describing the data-generation process are assumed to be fully described by a family of probability distributions involving only a finite number of unknown parameters. For example, one may assume that the distribution of population values is truly Normal, with unknown mean and variance, and that datasets are generated by 'simple' random sampling. The family of *generalized linear models* is a widely-used and flexible class of parametric models.
- **Non-parametric**: The assumptions made about the process of generating the data are much less than in parametric statistics and may be minimal. For example, every continuous probability distribution has a median, which may be estimated using the sample median orthe Hodges-Lehmann-Sen estimator, which has good properties when the data arise from simple random sampling.
- **Semi-parametric**: This term typically implies assumptions 'between' fully and non- parametric approaches. For example, one may assume that a population distribution have a finite mean. Furthermore, one may assume that the mean response level in the population depends in a truly linear manner on some covariate (a parametric assumption) but does not make any parametric assumption describing the variance around that mean. More generally, semi-parametric models can often be separated into 'structural' and 'random variation' components. One component is treated parametrically and the other non-parametrically.

3.3.4 Importance of valid models/assumptions

Whatever level of assumption is made, correctly-calibrated inference in general requires these assumptions to be correct; i.e., that the data-generating mechanisms really has beencorrectly specified.

- Incorrect assumptions of *'simple' random sampling* can invalidate statistical inference. More complex semi- and fully-parametric assumptions are also cause for concern. For example, incorrect assumptions of Normality in the population can invalidates some forms of regression-based inference.
- The use of **any** parametric model is viewed skeptically by most experts in sampling human populations: "most sampling statisticians, when they deal with confidence intervals at all, limit themselves to statements about [estimators] based on very large samples, wherethe central limit theorem ensures that these [estimators] will have distributions that are nearlynormal." Here, the central limit theorem states that the distribution of the sample mean "for very large samples" is approximately normally distributed, if the distribution is not heavy tailed.

3.3.5 Approximate distributions

Given the difficulty in specifying exact distributions of sample statistics, many methods havebeen developed for approximating these.

With ***finite samples***, approximation results measure how close a limiting distribution approaches the statistic's **sample distribution**: For example, with 10,000 independent samples the normal distribution approximates (to two digits of accuracy) the distribution of the sample mean for many population distributions. Yet for many practical purposes, the normal approximation provides a good approximation to the sample-mean's distribution when there are 10 (or more) independent samples, according to simulation studies, and statisticians' experience. Following Kolmogorov's work in the 1950s, advanced statistics uses approximation theory and functional analysis to quantify the error of approximation: In this approach, the metric geometry of probability distributions is studied; this approach quantifies approximation error.

With ***infinite samples***, limiting results like the central limit theorem describe the sample statistic's limiting distribution, if one exists. Limiting results are not statements about finite samples, and indeed are logically irrelevant to finite samples. However, the asymptotic theory of limiting distributions is often invoked for work in estimation and testing. For example, limiting results are often invoked to justify the generalized method of moments and the use of generalized estimating equations, which are popular in econometrics and biostatistics. The magnitude of the difference between the limiting distribution and the true distribution (formally, the 'error' of the approximation) can be assessed using simulation. Theuse of limiting results in this way works well in many applications, especially with low- dimensional models with log-concave likelihoods (such as with one-parameter exponential families).

3.3.6 Randomization-based models

For a given dataset that was produced by a randomization design, the randomization distribution of a statistic (under the null-hypothesis) is defined by evaluating the test statistic for all of the plans that could have been generated by the randomization design.

In frequentist inference, randomization allows inferences to be based on the randomization distribution rather than a subjective model, and this is important especially in survey sampling and design of experiments. Statistical inference from randomized studies is also more straightforward than many other situations.

In Bayesian inference, randomization is also of importance:
In survey sampling –*sampling without replacement* ensures the *exchangeability* of the sample with thepopulation; in randomized experiments, randomization warrants a *missing at random* assumption for covariate information.

Objective randomization allows properly inductive procedures. Many statisticians preferrandomization-based analysis of data that was generated by well-defined randomizationprocedures. However, it is has been observed that in fields of science with developedtheoretical knowledge and experimental control, randomized experiments may increase the costs of experimentation without improving the quality of inferences. Similarly, results from randomized experiments are recommended by leading statistical authorities as allowing inferences with greater reliability than do observational studies of the same phenomena. However, a good observational study may be better than a bad randomized experiment.

The statistical analysis of a randomized experiment may be based on the randomization scheme stated in the experimental protocol and does not need a subjective model. However, not all hypotheses can be tested by randomized experiments or random samples, which often require a large budget, a lot of expertise and time, and may have ethical problems.

3.3.7 Modes of inference

Different schools of statistical inference have become established. These schools (or 'paradigms') are not mutually-exclusive, and methods which work well under one paradigm often have attractive interpretations under other paradigms. The two main paradigms in use are frequentist and Bayesian inference, which are both summarized below.

a. Frequentist inference

This paradigm regulates the production of propositions by considering (notional) repeated sampling of datasets similar to the one at hand. By considering its characteristics under repeated sample, the frequentist properties of any statistical inference procedure can be described - although in practice this quantification may be challenging. Examples of frequentist inference are: P-value and Confidence interval

The frequentist calibration of procedures can be done without regard to utility functions. However, some elements of frequentist statistics, such as statistical decision theory, doincorporate utility functions. Loss functions must be explicitly stated for statistical theoriststo prove that a statistical procedure has an optimality property. For example, median- unbiased estimators are optimal under absolute value loss functions, and least squares estimators are optimal under squared error loss functions.

While statisticians using frequentist inference must choose for themselves the parameters of interest, and the estimators/test statistic to be used, the absence of obviously-explicit utilities and prior distributions has helped frequentist procedures to become widely- viewed as 'objective'.

b. Bayesian inference

The Bayesian calculus describes degrees of belief using the 'language' of probability; beliefs are positive, integrate to one, and obey probability axioms. Bayesian inference uses the available *posterior beliefs* as the basis for making statistical propositions. There are several different justifications for using the Bayesian approach. Examples of Bayesian inference are: *Credible intervals* for interval estimation and *Bayes factors* for model comparison

Many informal Bayesian inferences are based on "intuitively reasonable" summaries of the posterior. For example, the posterior mean, median and mode, highest posterior density intervals, and Bayes Factors can all be motivated in this way. While a user's utility function need not be stated for this sort of inference, these summaries do all depend (to some extent) on stated earlier beliefs, and are generally viewed as subjective conclusions.

Formally, Bayesian inference is calibrated with reference to an explicitly stated utility, or loss function; the 'Bayes rule' is the one which maximizes expected utility, averaged over the subsequent uncertainty. Formal Bayesian inference therefore automatically provides optimal decisions in a decision theoretic sense. Given assumptions, data and utility, Bayesian inference can be made for essentially any problem, although not every statistical inference need have a Bayesian interpretation. Some advocates of Bayesian inference assert that inference *must* take place in this decision-theoretic framework, and that Bayesian inference should not conclude with the evaluation and summarization of posterior beliefs.

3.4 Other Essential Statistics for Simulations

3.4.1 Sample Size Determination

A common goal of survey research is to collect data representative of a population. The researcher uses information gathered from the survey to generalize findings from a drawn sample back to a population, within the limits of random error. However, when critiquing business education research, Wunsch (1986) stated that ―two of the most consistent flaws included:

1. disregard for sampling error when determining sample size, and
2. disregard for response and non-response bias‖.

Within a quantitative survey design, determining sample size and dealing with no response bias is essential. ―One of the real advantages of quantitative methods is their ability to use smaller groups of people to make inferences about larger groups that would be prohibitively expensive to study‖. The question then is, how large of a sample is required to infer research findings back to a population?

Standard textbook authors and researchers offer tested methods that allow studies to take full advantage of statistical measurements, which in turn give researchers the upper hand in determining the correct sample size. Sample size is one of the four inter-related features of a study design that can influence the detection of significant differences, relationships or interactions (Peers, 1996). Generally, these survey designs try to minimize both alpha error (finding a difference that does not actually exist in the population) and beta error (failing to find a difference that actually exists in the population) (Peers, 1996).
However, improvement is needed. Researchers are learning experimental statistics fromhighly competent statisticians and then doing their best to apply the formulas and approaches

Foundations for Sample Size Determination

Primary Variables of Measurement

The researcher must make decisions as to which variables will be incorporated into formula calculations. For example, if the researcher plans to use a seven-point scale to measure a continuous variable, e.g., job satisfaction, and also plans to determine if the respondentsdiffer by certain categorical variables, e.g., gender, tenured, educational level, etc., which variable(s) should be used as the basis for sample size? This is important because the use of gender as the primary variable will result in a substantially larger sample size than if one used the seven-point scale as the primary variable of measure.
Cochran (1977) addressed this issue by stating that ―One method of determining sample size is to specify margins of error for the items that are regarded as most vital to the survey. An estimation of the sample size needed is first made separately for each of these important items‖. When these calculations are completed, researchers will have a range of n‘s, usually ranging from smaller n‘s for scaled, continuous variables, to larger n‘s for dichotomous or categorical variables.
The researcher should make sampling decisions based on these data. If the n‘s for the variables of interest are relatively close, the researcher can simply use the largest n as the sample size and be confident that the sample size will provide the desired results.

More commonly, there is a sufficient variation among the n‘s so that we are reluctant to choose the largest, either from budgetary considerations or because this will give an over- all standard of precision substantially higher than originally contemplated. In this event, the desired standard of precision may be relaxed for certain of the items, in order to permit the use of a smaller value of n. The researcher may also decide to use this information in deciding whether to keep all of the variables identified in the study. ―In some cases, the n‘s are so discordant that certain of them must be dropped from the inquiry; . . .‖.

Error Estimation

Cochran‘s (1977) formula uses two key factors:

(1) the risk the researcher is willing to accept in the study, commonly called the margin of error, or the error the researcher is willing to accept, and

(2) the alpha level, the level of acceptable risk the researcher is willing to accept that the true margin Alpha Level.

The alpha level used in determining sample size in most educational research studies is either .05 or .01 (Ary, Jacobs, & Razavieh, 1996). In Cochran‘s formula, the alpha level is incorporated into the formula by utilizing the t-value for the alpha level selected (e.g., t-value for alpha level of .05 is 1.96 for sample sizes above 120). Researchers should ensure they use the correct t-value when their research involves smaller populations, e.g., t-value for alpha of .05 and a population of 60 is 2.00.

In general, an alpha level of .05 is acceptable for most research. An alpha level of .10 or lower may be used if the researcher is more interested in identifying marginal relationships, differences or other statistical phenomena as a precursor to further studies.

An alpha level of .01 may be used in those cases where decisions based on the research are critical and errors may cause substantial financial or personal harm, e.g., major programmaticchanges.

Acceptable Margin of Error

The general rule relative to acceptable margins of error in educational and social research is as follows: For categorical data, 5% margin of error is acceptable, and, for continuous data, 3% margin of error is acceptable (Krejcie & Morgan, 1970). For example, a 3% margin of error would result in the researcher being confident that the true mean of a seven point scaleis within ±.21 (.03 times seven points on the scale) of the mean calculated from the research sample. For a dichotomous variable, a 5% margin of error would result in the researcherbeing confident that the proportion of respondents who were male was within ±5% of the proportion calculated from the research sample. Researchers may increase these values when a higher margin of error is acceptable or may decrease these values when a higher degree of precision is needed.

Variance Estimation

A critical component of sample size formulas is the estimation of variance in the primary variables of interest in the study. The researcher does not have direct control over variance and must incorporate variance estimates into research design. Cochran (1977) listed four ways of estimating population variances for sample size determinations:

(1) take the sample in two steps, and use the results of the first step to determine how many additional responses are needed to attain an appropriate sample size based on the variance observed in the first step data;

(2) use pilot study results;

(3) use data from previous studies of the same or a similar population; or

(4) estimate or guess the structure of the population assisted by some logical mathematical results.

The first three ways are logical and produce valid estimates of variance; therefore, they donot need to be discussed further. However, in many educational and social research studies, itis not feasible to use any of the first three ways and the researcher must estimate variance using the fourth method.

A researcher typically needs to estimate the variance of scaled and categorical variables. To estimate the variance of a scaled variable, one must determine the inclusive range of the scale, and then divide by the number of standard deviations that would include all possible values in the range, and then square this number. For example, if a researcher used a seven- point scale and given that six standard deviations (three to each side of the mean) would capture 98% of all responses, the calculations would be as follows:

7 (number of points on the scale) S = ...

6 (number of standard deviations)

When estimating the variance of a dichotomous (proportional) variable such as gender, Krejcie and Morgan (1970) recommended that researchers should use .50 as an estimate of the population proportion. This proportion will result in the maximization of variance, which will also produce the maximum sample size. This proportion can be used to estimate variancein the population. For example, squaring .50 will result in a population variance estimate of
.25 for a dichotomous variable.

Basic Sample Size Determination

a. Continuous Data

Before proceeding with sample size calculations, assuming continuous data, the researcher should determine if a categorical variable will play a primary role in data analysis. If so, the categorical sample size formulas should be used. If this is not the case, the sample size formulas for continuous data described in this section are appropriate.

Assume that a researcher has set the alpha level a priori at .05, plans to use a seven point scale, has set the level of acceptable error at 3%, and has estimated the standard deviation of the scale as 1.167. Cochran's sample size formula for continuous data and an example of its use is presented here along with the explanations as to how these decisions were made.

$$n_0 = \frac{(t)^2 * (s)^2}{(d)^2} = \frac{(1.96)^2(1.167)^2}{(7*.03)^2} = 118$$

Where t = value for selected alpha level of .025 in each tail = 1.96 (the alpha level of .05 indicates the level of risk the researcher is willing to take that true margin of error mayexceed the acceptable margin of error.)

s = estimate of standard deviation in the population = 1.167 (estimate of variance deviation for 7 point scale calculated by using 7 [inclusive range of scale] divided by 6 [number of standard deviations that include almost all (approximately 98%) of the possible values in the range]).

d = acceptable margin of error for mean being estimated = .21 (number of points on primary scale * acceptable margin of error; points on primary scale = 7; acceptable margin of error =
.03 [error researcher is willing to except]).

Therefore, for a population of 1,679, the required sample size is 118. However, since this sample size exceeds 5% of the population (1,679*.05=84), Cochran's (1977) correctionformula should be used to calculate the final sample size. These calculations are as follows:

$$n = \frac{no}{(1 + no / Population)} = \frac{(118)}{(1 + 118/1679)} = 111$$

Where population size = 1,679.

n0 = required return sample size according to Cochran‘s formula= 118. n1 = required returnsample size because sample > 5% of population.

These procedures result in the minimum returned sample size. If a researcher has a captiveaudience, this sample size may be attained easily.

However, since many educational and social research studies often use data collection methods such as surveys and other voluntary participation methods, the response rates are typically well below 100%. Salkind (1997) recommended over-sampling when he stated that

‒If you are mailing out surveys or questionnaires . . . count on increasing your sample size by 40%-50% to account for lost mail and uncooperative subjects‖. But Over- sampling can add costs to the survey but is often necessary. A second consequence is, of course, that the variances of estimates are increased because the sample actually obtained is smaller than the target sample.

However, many researchers criticize the use of over-sampling to ensure that this minimum sample size is achieved and suggestions on how to secure the minimal sample size are scarce.If the researcher decides to use over-sampling, four methods may be used to determine the anticipated response rate:

(1) take the sample in two steps, and use the results of the first step to estimate how manyadditional responses may be expected from the second step;

(2) use pilot study results;

(3) use responses rates from previous studies of the same or a similar population; or

(4) estimate the response rate. The first three ways are logical and will produce validestimates of response.

b. Categorical Data

The sample size formulas and procedures used for categorical data are very similar, but some variations do exist. Assume a researcher has set the alpha level a priori at .05, plans to use aproportional variable, has set the level of acceptable error at 5%, and has estimated the standard deviation of the scale as .5. Cochran‘s sample size formula for categorical data andan example of its use is presented here along with explanations as to how these decisions were made.

(t)2 * (p)(q)

no= ----------------

(d)2

(1.96)2(.5)(.5)

---------------- = 384(.05)2

Where t = value for selected alpha level of .025 in each tail = 1.96 (the alpha level of .05 indicates the level of risk the researcher is willing to take that true margin of error mayexceed the acceptable margin of error).

Where (p)(q) = estimate of variance = .25 (maximum possible proportion (.5) * 1- maximum possible proportion (.5) produces maximum possible sample size).

Where d = acceptable margin of error for proportion being estimated = .05 (error researcheris willing to except).

Therefore, for a population of 1,679, the required sample size is 384. However, since this sample size exceeds 5% of the population (1,679*.05=84), Cochran‘s (1977) correctionformula should be used to calculate the final sample size. These calculations are as follows:

n_0

$n_1 =$ .... $(1 + n_o / \text{Population})$

(384)

$n_1 =$ .... $= 313$

(1 + 384/1679)

Where population size = 1,679,

n_0 = required return sample size according to Cochran's formula= 384,n_1 = required return sample size because sample > 5% of population

These procedures result in a minimum returned sample size of 313. Using the same oversampling procedures as cited in the continuous data example, and again assuming aresponse rate of 65%, a minimum drawn sample size of 482 should be used. These calculations were based on the following:

Where anticipated return rate = 65%.

Where n2 = sample size adjusted for response rate. Where minimum sample size (corrected) =313.

Therefore, n2 = 313/.65 = 482.

3.4.2 The Central Limit Theorem

The main idea of the central limit theorem (CLT) is that the average of a sample of observations drawn from some population with any shape-distribution is approximately distributed as a normal distribution if certain conditions are met. In theoretical statistics there are several versions of the central limit theorem depending on how these conditions are specified. These are concerned with the types of assumptions made about the distribution of the parent population (population from which the sample is drawn) and the actual sampling procedure.

One of the simplest versions of the theorem says that if is a random sample of size n (say, n larger than 30) from an infinite population, finite standard deviation, then the standardized sample mean converges to a standard normal distribution or, equivalently, the sample mean approaches a normal distribution with mean equal to the population mean and standard deviation equal to standard deviation of the population divided by the square root of sample size n. In applications of the central limit theorem to practical problems in statistical inference, however, statisticians are more interested in how closely the approximate distribution of the sample mean follows a normal distribution for finite sample sizes, than the limiting distribution itself. Sufficiently close agreement with a normal distribution allows statisticians to use normal theory for making inferences about population parameters (such as the mean) using the sample mean, irrespective of the actual form of the parent population.

It is well known that whatever the parent population is, the standardized variable will have a distribution with a mean 0 and standard deviation 1 under random sampling. Moreover, if the parent population is normal, then it is distributed exactly as a standard normal variable for any positive integer n. The central limit theorem states the remarkable result that, even when the parent population is non-normal, the standardized variable is approximately normal if the sample size is large enough (say > 30). It is generally not possible to state conditions under which the approximation given by the central limit theorem works and what sample sizes are needed before the approximation becomes good enough. As a general guideline, statisticians have used the prescription that if the *parent distribution is symmetric and relatively short-tailed,* then the sample mean reaches approximate normality for smallersamples than if the parent population is skewed or long-tailed.

Under certain conditions, in large samples, the sampling distribution of the sample mean can be approximated by a normal distribution. The sample size needed for the approximation to be adequate depends strongly on the shape of the parent distribution. Symmetry (or lack thereof) is particularly important. For a symmetric parent distribution, even if very different from the shape of a normal distribution, an adequate approximation can be obtained with small samples (e.g., 10 or 12 for the uniform distribution). For symmetric short-tailed parent distributions, the sample mean reaches approximate normality for smaller samples than if the parent population is skewed and long-tailed. In some extreme cases (e.g. binomial) samples sizes far exceeding the typical guidelines (e.g., 30) are needed for an adequate approximation. For some distributions without first and second moments (e.g., Cauchy), the central limit theorem does not hold.

3.4.3 The Least Squares Model

Many problems in analyzing data involve describing how variables are related. The simplest of all models describing the relationship between two variables is a linear, or straight-line, model. The simplest method of fitting a linear model is to "eye-ball" a line through the dataon a plot. A more elegant, and conventional method is that of "least squares", which finds the line minimizing the sum of distances between observed points and the fitted line. With this you will:

- Realize that fitting the "best" line by eye is difficult, especially when there is a lot ofresidual variability in the data.
- Know that there is a simple connection between the numerical coefficients in theregression equation and the slope and intercept of regression line.
- Know that a single summary statistic like a correlation coefficient does not tell the whole story. A scatter plot is an essential complement to examining the relationshipbetween the two variables.

3.4.4 ANOVA: Analysis of Variance

Analysis of Variance or ANOVA enables us to test the difference between 2 or more means. ANOVA does this by examining the ratio of variability between two conditions and variability within each condition. For example, if we give a drug that we believe will improvememory to a group of people and give a placebo to another group of people, we might measure memory performance by the number of words recalled from a list we ask everyoneto memorize. A **t-test** would compare the likelihood of observing the difference in the mean number of words recalled for each group. An ANOVA test, on the other hand, wouldcompare the variability that we observe between the two conditions to the variability observed within each condition. We measure variability as the sum of the difference of each score from the mean. When we actually calculate an ANOVA we use a short-cut formula. Thus, when the variability that we predict (between the two groups) is much greater than the variability we don't predict (within each group) then we will conclude that our treatments produce different results.

3.4.5 Exponential Density Function (EDF)

EDF is use to take important class of decision problems under uncertainty such as the chance between events. For example, the chance of the length of time to next breakdown of a machine not exceeding a certain time, such as the photocopying machine in your office not tobreak during this week.

Exponential distribution gives distribution of time between independent events occurring at aconstant rate.

Its density function is:
f(t) = □ exp(-□t),
where □ is the average number of events per unit of time, which is a positive number. Themean and the variance of the random variable t (time between events) are 1/ □, and $1/□^2$, respectively
Applications include probabilistic assessment of the time between arrival of patients to theemergency room of a hospital, and arrival of ships to a particular port.

3.4.6 Poisson Process

An important class of decision problems under uncertainty is characterized by the small chance of the occurrence of a particular event, such as an accident. Poisson gives probability of exactly x independent occurrences during a given period of time if events take place independently and at a constant rate. It may also represent number of occurrences over constant areas or volumes. The following statements describe the *Poisson Process*:

1. The occurrences of the events are independent.
2. The occurrence of events from a set of assumptions in an interval of space or time hasno effect on the probability of a second occurrence of the event in the same, or any other, interval.
3. Theoretically, an infinite number of occurrences of the event must be possible in the interval.
4. The probability of the single occurrence of the event in a given interval isproportional to the length of the interval.
5. In any infinitesimally small portion of the interval, the probability of more thanoneoccurrence of the event is negligible.

Poisson processes are often used, for example in quality control, reliability, insuranceclaim, incoming number of telephone calls, and queuing theory.

An Application: One of the most useful applications of the Poisson Process is in the field of queuing theory. In many situations where queues occur it has been shown that the number of people joining the queue in a given time period follows the Poisson model. For example, if the rate of arrivals to an emergency room is □ per unit of time period (say 1hr), then:
P (n arrivals) = $□^n e^-$ / n!

The mean and variance of random variable n are both □. However if the mean and variance of a random variable having equal numerical values, then it is not necessary that its distribution is a Poisson.

Applications:
P (0 arrival) = $e^{-□}$
P (1 arrival) = □ e^- □ / 1!P (2 arrival) = $□^2e^-$ / **2**
and so on. In general:

P (n+1 arrivals) = □Pr (n arrivals) / n.

3.4.7 Uniform Density Function (UDF)

This function gives the probability that observation will occur within a particular interval when probability of occurrence within that interval is directly proportional to interval length.

For example, it is used to generate random numbers in sampling and Monte Carlosimulation.

The mass function of geometric mean of n independent uniforms [0,1]is: P(X = x) = n $x^{(n - 1)}$

$(Log[1/x^n])^{(n-1)} / (n - 1)!$.

$z_L = [U^L-(1-U)^L] / L$ is said to have Tukey's symmetrical distribution.

You may like to use *Uniform Applet* to perform your computations, then visit also: http://home.ubalt.edu/ntsbarsh/Business-stat/otherapplets/pvalues.htm

3.4.7 Test for Randomness

We need to test for both randomness as well as uniformity. The tests can be classified in 2 categories: Empirical or statistical tests, and theoretical tests.

Theoretical tests deal with the properties of the generator used to create the realization with desired distribution, and do not look at the number generated at all. For example, we would not use a generator with poor qualities to generate random numbers.

Statistical tests are based solely on the random observations produced.

A. Test for independence:

Plot the x_i realization vs x_{i+1}. If there is independence, the graph will not show any distinctive patterns at all, but will be perfectly scattered.

B. Runs tests.(run-ups, run-downs):

This is a direct test of the independence assumption. There are two test statistics to consider: one based on a normal approximation and another using numerical approximations.

Test based on Normal approximation:

Suppose you have N random realizations. Let K be the total number of runs in a sequence. If the number of positive and negative runs are greater than say 20, the distribution of K is reasonably approximated by a Normal distribution with mean (2N - 1) /3 and (16N - 29) / 90. Reject the hypothesis of independence or existence of runs if | Zo| < Z(1-alpha/2) where Zo is the Z score.

C. Correlation tests:

Do the random numbers exhibit discernible correlation? Compute the sample Autcorrelation Function.

Frequency or Uniform Distribution Test:

Use Kolmogorov-Smirimov test to determine if the realizations follow a U(0,1).

3.4.8 Some Useful SPSS Commands

a. Test for Binomial:

NPAR TEST BINOMIAL(p)=GENDER(0, 1)

b. Gooness-of-fit for discrete r.v.:

NPAR TEST CHISQUARE=X (1,3)/EXPECTED=20 30 50

C. Two population t-test

T-TEST GROUPS=GENDER(1,2)/VARIABLES=X

3.5 Neural Networks in Modelling & Simulation

Neural network is the branch of artificial intelligence. Neural network is a network of many processors named as units, each unit having its small local memory. Each unit is connected by unidirectional communication channels named as connections, which carry the numeric data. Each unit works only on their local data and on the inputs they receive from the connections.

History

The historical perspective of simulation is as enumerated in a chronological order.

The first neural model was developed in **1940** by McCulloch & Pitts.

In **1949**, Donald Hebb wrote a book "The Organization of Behavior", which pointed to the concept of neurons.

In **1950**, with the computers being advanced, it became possible to make a model on these theories. It was done by IBM research laboratories. However, the effort failed and later attempts were successful.

In **1959**, Bernard Widrow and Marcian Hoff, developed models called ADALINE and MADALINE. These models have Multiple ADAptive LINear Elements. MADALINE was the first neural network to be applied to a real-world problem.

In **1962**, the perceptron model was developed by Rosenblatt, having the ability to solve simple pattern classification problems.

In **1969**, Minsky & Papert provided mathematical proof of the limitations of the perceptron model in computation. It was said that the perceptron model cannot solve X-OR problem. Such drawbacks led to the temporary decline of the neural networks.

In **1982**, John Hopfield of Caltech presented his ideas on paper to the National Academy of Sciences to create machines using bidirectional lines. Previously, unidirectional lines were used.

When traditional artificial intelligence techniques involving symbolic methods failed, then arises the need to use neural networks. Neural networks have its massive parallelism techniques, which provide the computing power needed to solve such problems.

Application Areas

Neural network can be used in speech synthesis machines, for pattern recognition, to detect diagnostic problems, in robotic control boards and medical equipments.

Fuzzy Set in Modelling & Simulation

As discussed earlier, each process of continuous simulation depends on differential equations and their parameters such as a, b, c, d > 0. Generally, point estimates are calculated and used in the model. However, sometimes these estimates are uncertain so we need fuzzy numbers in differential equations, which provide the estimates of the unknown parameters.

What is a Fuzzy Set?

In a classical set, an element is either a member of the set or not. Fuzzy sets are defined in terms of classical sets **X** as −

$\mathbf{A = \{(x,\mu A(x)) | x \in X\}}$

Case 1 − The function **μA(x)** has the following properties −

$\mathbf{\forall x \in X \ \mu A(x) \geq 0}$

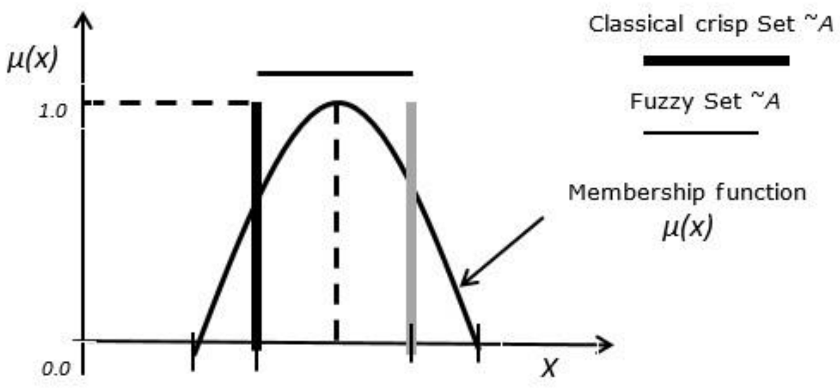
$\mathbf{\sup x \in X \{\mu A(x)\} = 1}$

Case 2 − Let fuzzy set **B** be defined as **A = {(3, 0.3), (4, 0.7), (5, 1), (6, 0.4)}**, then its standard fuzzy notation is written as **A = {0.3/3, 0.7/4, 1/5, 0.4/6}**

Any value with a membership grade of zero doesn't appear in the expression of the set.

Case 3 − Relationship between fuzzy set and classical crisp set.

The following figure depicts the relationship between a fuzzy set and a classical crisp set.



Self-Assessment Exercise(s)

Answer the following questions:

1. State the Cochran's sample size formula for continuous and categorical data
2. Assume a researcher has set the alpha level a priori at 10%, plans to use a proportional variable, has set the level of acceptable error at 5%, and has estimated the standard deviation of the scale as .5. Find the sample size for a population of 2500
3. What are the Cochran's key factors for error estimation?
4. State the essential feature of Poisson process
5. List four ways of estimating population variances for sample size determinationsaccording to Cochran.
6. What are the objectives of randomization, and state the importance of randomization in frequentist and Bayesian inferences

5.0 Conclusion

Statistics is the basis of simulation. In this unit we have simply introduced some basic statistics in modelling and simulations. We hope that the reader will broaden his/herunderstanding by consulting the referenced texts or othe statistics books.

6.0 Summary

In this unit we were able to

- Differentiate between the two broad components of statistics: descriptive andinference statistics
- Have concise discussions of descriptive statistics on; Univarite statistics measures: the distribution, central tendency, dispersion, etc. and gave someexamples.
- Discuss Inference statistics under the following subheads: definition, Model/assumptions, approximate distributions, random-based models andmodes of inference.
- Introduce some essential statistical measures in simulation such as;
  - sample size determination
  - central limit theorem

- least square model
- Analysis of variance
- Exponential distribution function
- Poisson distribution
- Uniform distribution
- Test for randomness
- Some commands of Special package for statistical analysis (SPSS)

7.0 Further Readings

- Gordon, S. I., & Guilfoos, B. (2017). *Introduction to Modeling and Simulation withMATLAB® and Python*. Milton: CRC Press.
- Zeigler, B. P., Muzy, A., & Kofman, E. (2019). *Theory of modeling and simulation: Discrete event and iterative system computational foundations*. San Diego (Calif.):Academic Press.
- Kluever, C. A. (2020). *Dynamic systems modeling, simulation, and control*. Hoboken,N.J: John Wiley & Sons.
- Law, A. M. (2015). *Simulation modeling and analysis*. New York: McGraw-Hill.
- Verschuuren, G. M., & Travise, S. (2016). *100 Excel Simulations: Using Excel to ModelRisk, Investments, Genetics, Growth, Gambling and Monte Carlo Analysis*. Holy Macro! Books.
- Grigoryev, I. (2015). *AnyLogic 6 in three days: A quick course in simulation modeling*.Hampton, NJ: AnyLogic North America.
- Dimotikalis, I., Skiadas, C., & Skiadas, C. H. (2011). *Chaos theory: Modeling, simulationand applications: Selected papers from the 3rd Cghaotic Modeling and Simulation International Conference (CHAOS2010), Chania, Crete, Greece, 1-4 June, 2010*. Singapore: World Scientific.
- Velten, K. (2010). *Mathematical modeling and simulation: Introduction for scientists andengineers*. Weinheim: Wiley-VCH.