COMP4030

DATA MODELLING AND ANALYSIS

Lecture 3: Simulation

LECTURE OUTLINE

1. Lecture Outcomes
2. Previously in DMA
    1. Mathematical Modelling
    2. DMA at a glance
3. Simulation: Definition
4. Types of Simulation
    1. Continuous Simulation
    2. Discrete Event Simulation
    3. Monte Carlo Simulation
5. Model and simulation example

LECTURE OUTCOMES

- At the end of this lecture, you should be able to answer these questions:
  - What is simulation?
  - Why is it useful?
  - Where can it be applied?
  - When should it be applied?
  - What types of simulation are there?
  - What is discrete event simulation?
  - What is a deterministic simulation?

PREVIOUSLY IN DMA: MATHEMATICAL MODELS (DEFINITION)

- Mathematical models are used to solve problems (or answer questions) about a **system**.

> ***Definition 2.4 (Mathematical Model)*** *A mathematical model is a triplet (S, Q, M) where S is a system, Q is a set of questions* $Q = \{Q_1, Q_2, \ldots, Q_m\}$ *relating to S, and M is a set of mathematical statements* $M = \{\Sigma_1, \Sigma_2, \ldots, \Sigma_n\}$ *which can be used to answer Q.*

- The order of (S,Q,M) follows chronology
- The model its linked to its ***purpose***

PREVIOUSLY IN DMA: STEPS IN MODEL CONSTRUCTION

1 Identify the Question or Problem

↓

2 System Analysis

↓

3 Formulate Mathematical Model

↓

4 Obtain Mathematical Solution

↓

5 Interpret the Solution

↓

6 Compare with reality

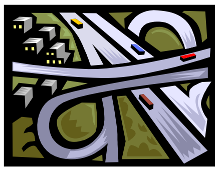
AND REMEMBER!

- The best model is the **simplest model that still serves its purpose:**
  - Complex enough to help us understand a system and to solve problems.
- **Seen in terms of a simple model:** the complexity of the system will no longer obstruct our view
- We will be able to look through the complexity of the system at the heart of things.

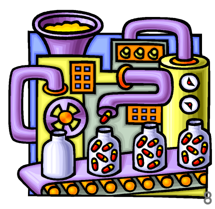
SIMULATION

- Experiments are a key component of model building
  - Collect data in order to understand the system
  - To compare model behaviour
- In some cases, experiments will not be possible:
  - Too complex:
    - What is the optimal configuration of road directions (one and two way roads) in order to minimize traffic?
  - Too expensive:
    - Optimal arrangement of machines in a plant to maximise productivity and minimise energy consumption?
  - Too time consuming:
    - How does Ehlers-Danlos syndrome (**EDS**) affect the nervous system of patients as they grow old?
  - Too dangerous!
    - What are the effects of a new species entering a new ecosystem?

SIMULATION



- Solution: use simulation!
- Simulation is a fundamental method in engineering and science
- Reasons for simulating models:
  - Uncertainty of components (e.g. customer arrivals)
  - Interconnectedness of components (e.g. customers in a three stage service process)
  - Combinatorial complexity
  - Dynamic complexity (e.g. actions have different effects in the short and long term)
- Can take into account system uncertainty, interconnectedness and complexity. It helps to:
  - Predict system performance,
  - Compare alternative system designs and
  - Determine the effects of alternative policies on system performance.



SIMULATION DEFINITION

> ***Definition 3.1 (Simulation)*** *Simulation is the application of a model with the objective to derive strategies that help solve a problem or answer a question pertaining to a system.*

- The term *simulation* originates from the Latin word "*simulare*", which means "to pretend"
  - In a simulation, the model *pretends* to be the real system.
- This definition of a *simulation* is again purpose oriented
  - Similar to our previous definition of a model
- Different authors might define the term differently according to their objectives

SIMULATION DEFINITION

- Another possible definition (by Fritzon [1]) is:

> ***Definition 3.2 (Simulation)*** *A simulation is an experiment performed on a model.*

- The role of simulations can be two-fold:
  - Simulate the complete model, when all parameters have been measured
    - Example: simulate the disease spread on a population with a known infection rate
  - Simulate parameters or variables whose values cannot be experimented with:
    - Simulate the disease spread on a population with an unknown infection rate

[1] Fritzon, P. (2004) *Principles of Object Oriented Modeling and Simulation with Modelica 2.1*, John Wiley & Sons, Ltd.

WHY SIMULATION?

- Advantages:
  - Approximate complex systems easily
  - Can test on a wide range of conditions
    - Experimental approach to modelling (‘What-if’ analysis tool)
  - Easy to modify
  - Can control level of detail quite easily
  - Easy to do sensitivity analysis
  - High-level simulation languages

WHY SIMULATION?

- Disadvantages:
  - Can be expensive to develop and operate
  - Requires expertise
  - Computational power and memory
  - Sometimes we need to conduct thousands of experiments to draw conclusions
    - Particularly if you are in an stochastic system
  - It is not *“real”*

TYPES OF SIMULATION:

- Simulations can be classified like models:
  - Lumped vs distributed
  - Deterministic vs. stochastic;
  - Discrete vs. continuous;
  - …
- However, there are two types of classification that merit a closer look
  - Discrete vs. continuous
  - Deterministic vs. stochastic

TYPES OF SIMULATION: CONTINUOUS VS DISCRETE

- **Continuous simulation:** simulation based on continuous time. Models will use numerical integration of differential equations.
  - Example: simulation of a rocket trajectory model
- **Discrete-event simulation:** models in which systems' states change their values at discrete times.
- Discrete simulations can be applied to represent a continuous system, but the resulting simulations produce approximate results.
- Continuous simulations can be applied to represent discrete phenomena, but the resulting simulations can produce impossible results for some cases.
  - Example: using a continuous simulation to model a live population of animals may produce the impossible result of 1/3 of a sheep.

DISCRETE EVENT SIMULATION

> ***Definition 3.3 (Discrete event simulation)*** *The modelling of a system as it evolves over time in which state variables change instantaneously at a (countable number of) separate points in time. These points in time are the ones at which an event occurs, where event is defined as an instantaneous occurrence that may change the state of a system.*

- Discrete event simulations can be used to simulate a great number of systems in various fields.
- To optimise and increase the efficiency of systems in:
  - Manufacturing, material handling, logistics, transportation, and healthcare
- We will be using Discrete Event Simulation in this course

DISCRETE EVENT SIMULATION: EXAMPLE 1 – RENTAL CAR COMPANY

- General problem statement:
  - You are given a set of *M* variables and their values at *N* times
  - You are asked to, given the variable patterns in the past, predict future values
- For example:
  - You are the CEO of Malaysia Rental Cars LLC, a rental car company.
  - You have two stores, one in Kuala Lumpur and one in Melaka.
  - You have 2000 cars in your fleet, that you need to distribute between both cities.
  - However, demand in each city is different, and varies depending on the day.
  - Looking at past patters of demand and general city requirements can you predict how many cars you must deploy to each city?

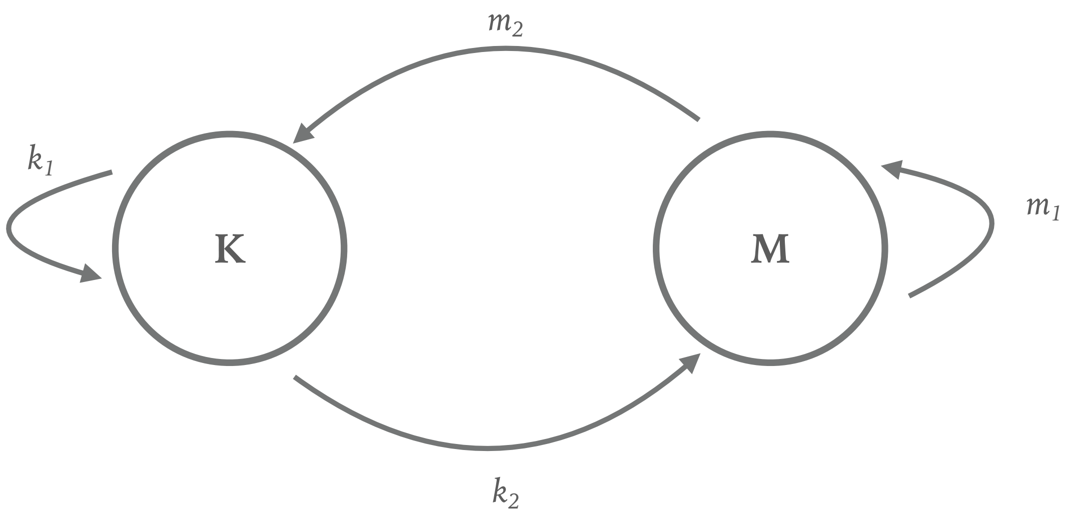
DISCRETE EVENT SIMULATION: EXAMPLE 1 – RENTAL CAR COMPANY

- Car demand and deployment in the last 6 days is as follows.
- On day zero the KL branch has 700 of the company's cars and the Melaka branch has 1300. On day one, the KL branch has 1210 and the Melaka branch has 790. And so on …
- Suppose we also know that generally, 80% cars in KL stay in KL and that 50% cars of Melaka remain in Melaka.
- What should you do for days 6 and 7?

Day	KL	Melaka
0	700	1300
1	1210	790
2	1363	637
3	1409	591
4	1423	577
5	1427	573

DISCRETE EVENT SIMULATION: EXAMPLE 1 – RENTAL CAR COMPANY

How can you model this problem?



DISCRETE EVENT SIMULATION: EXAMPLE 1 – RENTAL CAR COMPANY

$$K_{n+1} = k_1 K_n + m_2 M_n$$

$$M_{n+1} = k_2 K_n + m_1 M_n$$

- Where:
  - K_n represents the number of cars in KL at time n
  - M_n represents the number of cars in Melaka at time n
  - k_1 and m_1 represent the % of cars that stay in KL and M, respectively
  - k_2 and m_2 represent the % of cars that move between cities
    - where: $k_2 = 1 - k_1$ and $m_2 = 1 - m_1$
- To find the number of cars in KL and Melaka on days 6 and 7 we can write a simple simulation:
  - Imagine that 80% of KL cars stay there and 50% of Melaka cars leave Melaka at the end of the day.

DISCRETE EVENT SIMULATION: EXAMPLE 1 – RENTAL CAR COMPANY

$$K_{n+1} = 0.8K_n + 0.5M_n$$

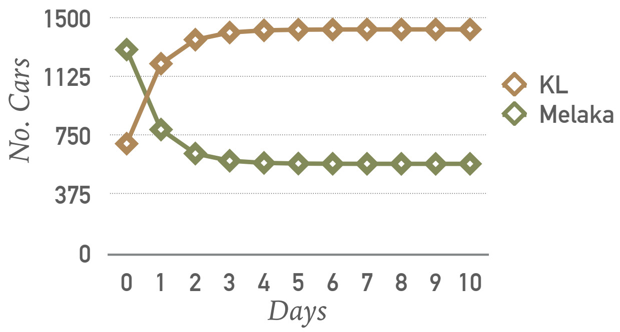
$$M_{n+1} = 0.2K_n + 0.5M_n$$

- Where
  - K_n represents the number of cars in KL at time n
  - M_n represents the number of cars in Melaka at time n
- To find the number of cars in KL and Melaka on days 6 and 7 we can write a simple simulation using a function in R

DISCRETE EVENT SIMULATION: EXAMPLE 1 – RENTAL CAR COMPANY

```
cars_exchange = function(kuala, melaka, time){
  cars = data.frame(t = 1, kl = kuala, mk = melaka)
  for (i in 2:time){
    cars[i,]$t = i
    cars[i,]$kl = 0.8* cars[i-1,]$kl + 0.5 *cars[i-1,]$mk
    cars[i,]$mk = 0.2*cars[i-1,]$kl + 0.5 * cars[i-1,]$mk
    }
  return(cars)}
```

Day	KL	Melaka
0	700	1300
1	1210	790
2	1363	637
3	1409	591
4	1423	577
5	1427	573
6	1428	572
7	1428	572



DISCRETE EVENT SIMULATION: EXAMPLE 1 – RENTAL CAR COMPANY

- Some of the categories of this simulation:
  - Mechanistic: we know something about the system (car rental and drop off)
  - Deterministic: no random effects
  - Discrete - event: involves time but in discrete states
  - Distributed: involves two different spatial locations
  - Direct problem: we know the inputs (i.e. day 0 to 5) and want to know the outputs (i.e. car stock in day 6 and 7)
  - Management: the goal of the system is to predict stock and facilitate managerial decisions

DISCRETE EVENT SIMULATION: EXAMPLE 1 - RENTAL CAR COMPANY

DEMO TIME

TYPES OF SIMULATION: DETERMINISTIC VS STOCHASTIC

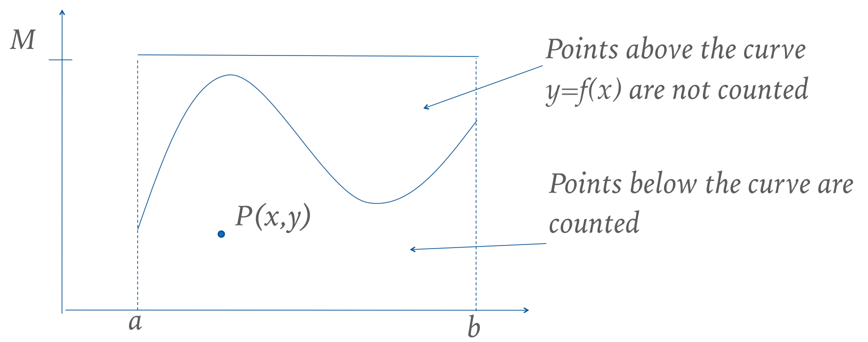
- **Deterministic simulation:** simulation in which all variables are regulated by deterministic algorithms.
  - Replicated runs with same conditions always produce identical results.
- **Stochastic simulation:** simulation where some variable (or process) is subject to randomness.
  - Uses Monte Carlo techniques or, when applicable, random number generators.
  - Replicated runs will produce different results within a specific confidence band.
- In this course, we will focus on Deterministic Simulations.

STOCHASTIC SIMULATION: MONTE CARLO SIMULATION

- Monte Carlo simulation: If it is necessary to know the value of a random variable and :
  - We don't know its distribution
  - But it is possible to take samples from the distribution
- We can estimate it by taking the numerous samples, independently, and averaging them.
- Points to consider, however are:
  - How many samples do you need?
  - How should you generate those samples?

STOCHASTIC SIMULATION: MONTE CARLO SIMULATION

- Example: compute the area under a curve whose function we do not know
- Assume:
  - $y = f(x)$ is a given continuous function satisfying $0 \leq f(x) \leq M$ over the closed interval $a \leq x \leq b$



STOCHASTIC SIMULATION: MONTE CARLO SIMULATION

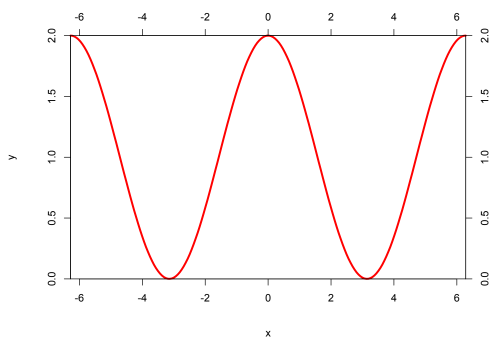
- Basic idea:
  - Generate random points within the rectangular boundary
  - Classify each based on whether it is located under or above the curve
  - Compute the ratio of under/total points
  - Multiply this ratio by the rectangular area

Input *Total number of n of random points to be generated in the simulation*
Output *AREA = approximate area under curve, between a and b*

Step 1 *Initialize: counter = 0*
Step 2 *For i=1, 2, …, n (do steps 3-4)*
Step 3 *Generate a random point (x_i,y_i) within [a,b] and [0,M]*
Step 4 *If $y_i \leq f(x_i)$ then increment counter*
Step 5 *Calculate AREA = M(b-a) × counter/n*

STOCHASTIC SIMULATION: EXAMPLE – MONTE CARLO SIMULATION

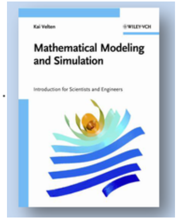
- Test $y = \cos(x)+1$ over the interval $-2\pi \leq x \leq 2\pi$, where $-1 \leq \cos(x) < 1$
- The actual area is 4π (12.566), half of the total area (8π)
- Results from Monte Carlo simulation:

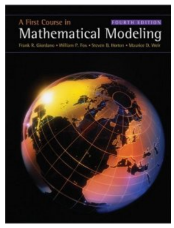


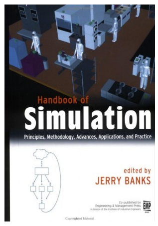
Number of points	Approximate area
100	12.817
200	12.692
400	11.750
600	13.572
800	12.723
1,000	12.290
10,000	12.360
100,000	12.564

RESOURCES

- Books used in this lecture
  - Velten, K., "*Mathematical Modeling and Simulation: Introduction for Scientists and Engineers*"; Wiley-VCH, 2009
  - Giordano, F.R., et al, "*A First Course in Mathematical Modeling*" (4th ed.); Brooks/Cole, Cengage Learning, 2009
  - Banks, J., (1998) "*Handbook of Simulation: Principles, Methodology, Advances, Applications*", Wiley-Interscience







LECTURE OUTCOMES

- At the end of this lecture, you should be able to answer these questions:
  - What is simulation?
  - Why is it useful?
  - Where can it be applied?
  - When should it be applied?
  - What types of simulation are there?
  - What is discrete event simulation?
  - What is a deterministic simulation?

THE END

Questions