Week 3 - Concepts in Mathematical Modelling

Week 3 - Concepts in Mathematical Modelling - Sampling - understanding the problem, choosing variables, making assumptions, solving the equations, interpreting the solution, validating the model, and criticizing and improving the model.

Introduction:

Mathematical modeling is a critical tool for understanding, analyzing, and solving complex real-world problems across various disciplines. In this module, we will delve into the core concepts of mathematical modeling and how it is applied in practice.

Learning objectives

- Understand what models and mathematical models are.
- Think about a real-world problem from a modelling perspective
- Framework of modelling
- Have a broad understanding of mathematical models and the scope of its application

What is a model?

There are three components of a model.

- Represents a phenomena
- Simple
- Flexible

An example of a Physical Model - Normally a prototype (constructed copy) of a real object - Examples: model building, bridges, airplane, solar system

An example of a Schematic Model - Normally used to visualize a system structure - Examples: Organisation charts, Maps, Diagrams, Process charts

An example of a Mathematical Model - A simple and flexible representation of a real object or phenomena using mathematical language - Examples: Algebraic equations, Differential equation, Algorithm - For data science: descriptive and inferential statistics

Why Mathematics?

1. Mathematics is a precise language

- Forces us to formulate concrete ideas and assumptions in an unambiguous way.

2. Mathematics is a concise language

- One equation says more than 1000 words

3. Mathematics is a universal language

- Same mathematical techniques can be applied over a range of scales.

4. Mathematics is the language that computers understand best.

So the objective of mathematical and statistical modelling is to analyse, make predictions and provide insights of real-world phenomena and the applications of mathematical modelling stretch far and wide.

Examples: - Epidemiology - Analysis of financial markets - Strategies for cooperative and noncooperative interactions - Software development - Economics (term of trades, rates, surplus analysis) - Electoral studies - Astronomy - Organisational behavious

However, there are some limitations of mathematical modelling. Mathematical modelling is only as good as the data and assumptions you provide it and even then you will never be able to model all the variables required for a model to be 100% accurate.

Definition of Mathematical Modeling:

Mathematical modeling is the process of using mathematical equations, algorithms, and simulations to represent, analyze, and make predictions about real-world phenomena. It involves creating simplified, abstract representations of complex systems to gain insights and solve practical problems.

Definition of Statistical Modeling

Statistical modeling is a process used to describe and understand the relationships between variables in a dataset. It involves the use of statistical techniques to analyze data, make predictions, and draw conclusions about the underlying patterns and structures within the data.

Key Concepts in Mathematical Modeling:

1. Abstraction:

Abstraction is the process of simplifying a complex real-world problem by focusing on its essential features while ignoring irrelevant details. Models are abstractions that capture the critical aspects of a system.

2. Variables:

Variables are quantities that can change in a model. They represent important parameters or properties of the system. Variables can be categorized as independent (inputs), dependent (outputs), or control variables.

3. Equations:

Mathematical models are typically represented by mathematical equations that describe the relationships between variables. Equations can be algebraic, differential, integral, or a combination of these.

4. Assumptions:

Assumptions are simplifications and constraints applied to the model to make it tractable. Assumptions should be justified and clearly stated in the modeling process.

5. Parameters:

Parameters are constants in the model that affect its behavior. Determining parameter values often requires data collection or expert knowledge.

6. Validation and Verification:

Validation is the process of testing a model's accuracy by comparing its predictions to real-world observations. Verification ensures that the model is implemented correctly and behaves as intended.

7. Simulation and Analysis:

Models are used to perform simulations to study the system's behavior under different conditions. Analysis of model results helps in drawing conclusions and making decisions.

8. Sensitivity Analysis:

Sensitivity analysis assesses how changes in parameters or assumptions affect the model's predictions. It helps identify the most critical factors influencing the system.

9. Model Selection:

Depending on the problem, different modeling techniques and approaches may be suitable, such as deterministic or stochastic models, continuous or discrete models, etc.

10. Interpretation:

The interpretation of model results is crucial for drawing meaningful conclusions and making decisions. It involves translating mathematical findings into actionable insights.

11. Communication:

Effective communication of modeling results to stakeholders is essential. Reports, presentations, and visualizations help convey the model's findings.

Key Concepts in Statistical Modelling

1. Variables:

In statistical modeling, variables are characteristics or attributes that can take different values. They can be classified into different types, such as categorical (e.g., gender, color) or continuous (e.g., age, temperature).

2. Data Distribution:

Understanding the distribution of the data is crucial in statistical modeling. This includes assessing the shape, center, spread, and any potential outliers or anomalies in the data.

3. Probability Distributions:

Probability distributions describe the likelihood of observing different outcomes in a dataset. Common probability distributions used in statistical modeling include the normal (Gaussian), binomial, Poisson, and exponential distributions.

4. Parameters:

Parameters are the unknown quantities in statistical models that need to be estimated from the data. They represent the characteristics of the population being studied and are typically denoted by Greek letters (e.g., for population mean, for population standard deviation).

5. Estimation:

Estimation involves the process of determining the values of unknown parameters based on observed data. Estimators are statistical techniques used to estimate parameters, and they can be point estimators (providing a single value estimate) or interval estimators (providing a range of plausible values).

6. Hypothesis Testing:

Hypothesis testing is a statistical method used to make inferences about population parameters based on sample data. It involves formulating null and alternative hypotheses, calculating test statistics, and determining whether there is enough evidence to reject the null hypothesis in favor of the alternative hypothesis.

7. Regression Analysis:

Regression analysis is a statistical technique used to model the relationship between one or more independent variables (predictors) and a dependent variable (outcome). It is widely used for prediction and inference in various fields.

8. Model Assumptions:

Every statistical model is based on certain assumptions about the data. These assumptions need to be carefully assessed to ensure the validity of the model and the reliability of the inferences drawn from it.

9. Model Selection and Validation:

Model selection involves choosing the most appropriate model among competing alternatives based on criteria such as goodness-of-fit, simplicity, and interpretability. Model validation involves assessing the performance of the selected model using techniques such as cross-validation or independent data validation.

10. Uncertainty and Confidence Intervals:

Uncertainty quantifies the degree of variability or randomness in statistical estimates. Confidence intervals provide a range of plausible values for a parameter estimate, along with a level of confidence that the true parameter value lies within that range.

Four stages of mathematical modelling

Stage 1: Building the model

1. Define the model objectives

- Be clear about what you want your model to do

2. Determine the appropriate level & key model components

- What level of simplification is required?
- Apply the principle of Ockham's razor

3. Define your assumptions

- Arguably the most important step in building a model
- Assumptions reflect our understanding how the phenomena occurs
- Remember: the model results are only as valid as the asssumptions!
- Different assumptions can lead to fundamental differences

4. Create a flow diagram

- Useful especially if your objective is to understand the relationship between variables
- A visual tool to simplify an otherwise complex model

5. Develop your model

- Depends on the objectives and your modelling approach – Statistical models normally use a single linear or non-linear function – Deterministic models of dynamical system normally use differential equations – Stochastic models require expression for the probability of events

Stage 2: Analysing the model

1. Analytical method (using mathematical principle)

- Precise! Provide an exact solution
- Usually only possible for simple systems or problems with strict assymptions
- In pure mathematics, the analysis of a pattern may lead to a proposition!

2. Numerical method (using computers)

- Applies to moth mathematical model: Widely use in many fields
- Requires the use of numerical algorithms implemented in computational systems (e.g. euler methods, monte-carlo simulations)
- Provide an approximate solution

- Use establish code. Avoid writing your own numerical solver as much as possible. You will be able to use existing libraries to create mathematical models. This might save a lot of time and cost especially in the experimental study.

3. Uncertainty analysis

- Normally found in stochastic (probabiliistic) model.
- Analysing the variable in output from the variablility in input e.g. How confident are we on the model output?

4. Sensitivity analysis

- Analyse how sensitive the output of the model when certain parameters (input) are changed. e.g. When we change our input, by how much does the output change?

Stage 3. Validating the model

1. Pure mathematics

- The process of writing out the proof of our proposition
- There are many techniques in proving a theorem: direct proof, proof by cases, proof by contradiction, proof by induction, etc.)

2. Applied mathematics (Predominantly statistical model)

- Comparison of model prediction to the real data (observations).
- In most cases, these statistical techniques are useful to check the model's robustness.
- Hypothesis testing
- Analysis of standard deviation and Mean-Squared Error (MSE)
- Identify why the model prediction may differ with the real data: natural variablility (2008 financial crash), important variables ignored in the model, etc

Stage 4. Applying the model

1. Pure mathematics

- Decide on the term of a mathematical statement. Does your solution consitute as a Theorem, Lemma, Corollary, or simply a generalization?
- Define the possible applications and extension to your model and findings.

2. Applied mathematics (Predominantly statistical model)

- Valuable decision support tools!
- In Economics and Finance, a model may offer several options for public decision-making
- Mathematical models such as linear programming and game theory are very popular because they have a wide range of applications.

What makes a good model?

The key attributes of a good model.

1. Fit for purpose

- Precisely and concisely solves the real-world problem for the end users
- Appropriate balance between accuracy, simplicity and flexibility

2. For statistical predictive model: How easy is it to use your model for different types of data (replicability)

Drawbacks of Modelling

Assumptions: Many statistical and mathematical models rely on strict assumptions about the data, such as linearity, normality, independence, and homoscedasticity. Violations of these assumptions can lead to biased estimates and incorrect inferences.

Overfitting: Overfitting occurs when a model captures noise or random fluctuations in the data rather than the underlying patterns. This can result in a model that performs well on the training data but fails to generalize to new, unseen data.

Underfitting: Underfitting occurs when a model is too simplistic to capture the underlying structure of the data. This can result in poor predictive performance and biased parameter estimates.

Complexity and Interpretability: Some models, especially those based on machine learning algorithms, can be highly complex and difficult to interpret. This lack of interpretability can make it challenging to understand the underlying mechanisms driving the model's predictions.

Data Requirements: Many statistical and mathematical models require large amounts of data to produce reliable estimates and predictions. In situations where data are scarce or incomplete, these models may perform poorly or fail altogether.

Data Quality: The quality of the data used to train and validate statistical models can significantly impact their performance. Missing values, measurement errors, and outliers can all distort the results and lead to incorrect conclusions.

Computational Complexity: Some models, particularly those involving large datasets or complex algorithms, can be computationally intensive and require substantial computational resources and time to train and evaluate.

Assumption of Causality: Statistical models often establish associations between variables but may not necessarily imply causality. Inferring causal relationships from observational data

can be challenging and may require additional methods such as experimental design or causal inference techniques.

Model Uncertainty: Statistical models provide point estimates or predictions along with measures of uncertainty, such as confidence intervals or prediction intervals. However, these measures may not fully capture all sources of uncertainty, leading to overly optimistic or conservative estimates.

Subjectivity and Bias: The choice of model, variables, and analysis techniques can be influenced by subjective judgments and biases. Researchers may inadvertently introduce bias into the modeling process, leading to erroneous conclusions.

Conclusion:

Both Mathematical modeling and Statistical modelling provide powerful and versatile approaches for tackling complex real-world problems. It involves abstraction, equations, assumptions, parameter estimation, validation, simulation, and interpretation. Whilst Mathematical models provide insights, support decision-making, and drive innovation in various fields. Statistical modelling also does this but relies on real world data to create and validate models.