ECS509U - Probability & Matrices

Tassos Tombros
Week 1

What is this module about?

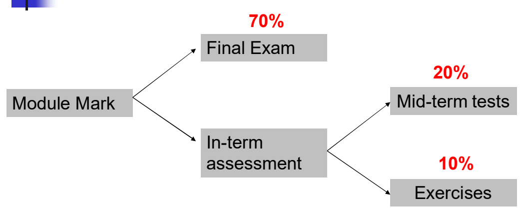
- The module aims to:
  - introduce (mathematical) topics that are relevant to computer applications, including **probability** and **basic linear algebra**
  - increase the students' capability to think abstractly and rigorously
  - focus on tools & problem solving

Lectures & Exercise Classes

- **Lectures**
  - Monday 15:00-16:00 @ **Arts 2 LT**
  - Thursday 14:00-15:00 @ **the Great Hall**
- **Tutorial/Exercise Classes**
  - Fridays, there are timetabling issues for now watch this space
- **Exercise Classes start THIS FRIDAY**

You will be assigned to a 1-hour slot, no changes will be negotiated in the assignment

Module Assessment



- Mid-Term tests: Friday of **Weeks 7 and 12**, times TBC

- Exercises: Weekly at the Exercise Classes

Why Probability? Why Matrices?

- Both provide useful **tools** for tackling many CS problems
  - network traffic modelling
  - software risk assessment
  - machine learning
  - computer graphics
  - computer vision & image processing
  - etc. etc.
- Both topics have extensive links to future CS modules

Modules that may use probabilistic techniques and/or matrices

- Data Mining
- Bayesian Decision & Risk Analysis
- Semi-structured Data and Advanced Data Modelling
- Artificial Intelligence
- Distributed Systems
- Computer graphics
- Image Processing
- etc. etc.

Week 1: Learning Objectives

- **Introduction to Probability**
- At the end of Week 1 you should be able to:
  - understand the concept of probability
  - differentiate between the various approaches for calculating probability values
  - work with sample spaces and events
  - understand the basic probability laws
  - apply these laws for solving simple problems

Introduction to Probability

- What does a probability value really represent?
  - chance? likelihood? odds? percentage? proportion?
- Interpreting probabilities can be interesting:
  - "The chance of winning a prize in an instant lottery is 7.74%"
    - How can it be so precise? Where does the number come from?
  - If the chance of rain for tomorrow is 30%, does it mean it will not rain because it is less than 50%?
  - 70% of students will pass this module on the first sit

So, what is probability?

- Typically, a value **between 0 and 1** that reflects the **likelihood** of the **occurrence** of a specific **event**
- Calculating probability values is harder some times than others
  - e.g. probability of observing a specific volume of network traffic in a given network connection for a given time period vs. probability or rolling a 3 with a die
- Probability values can be calculated in a variety of ways

The 4 approaches to calculating probabilities

- **Subjective**
  - The most vague and least scientific way, based on personal views, hopes, etc.
  - What do you think the chance that your favorite football team will win the Premier League is?
- **Classical**
  - Mathematical approach, using rules and formulas (more in the coming weeks)
  - Roll the dice!!!

The 4 approaches to calculating probabilities

- **Frequency-based**
  - Base calculations on observed data, and calculate the percentage of times that the event has occurred in the observed data (*relative frequency*)
  - Probabilities are estimates, since they are based on finite sample size (your estimates will be only as good as the data you collect)
- **Simulation-based**
  - We create the data by setting up a scenario, playing out the scenario a large number of times, and counting the percentage of times a certain outcome occurs

Some first formulas

- For the classical approach:

$$P(A) = \frac{\text{number of ways A can occur}}{\text{number of ways the experiment can proceed}}$$

  - This only works **if all outcomes are equally likely**
- For the frequency-based approach:

$$P(A) = \frac{\text{number of times event A occurred}}{\text{number of ways the experiment was run}}$$

Think…

a) We roll a die:

- What is the probability of getting a 2?
- What is the probability of getting a 2 or a 6?

P(rolling a 2) = 1/6
P(rolling a 2 or a 6) = 1/6 + 1/6 = 2/6

b) How would I be able to calculate the probability of any student passing a module that I have been teaching for the past 6 years?

Basic terms & definitions

- It all comes down to:
  - understanding the **event** for which you want to estimate the probability
  - calculating all the **possible outcomes** of the process at hand
- Think of the single die example:
  - **experiment:** throwing one die
  - **event:** e.g. throwing an odd number
  - **all possible outcomes** (sample space): 1, 2, 3, 4, 5, 6
- Now let's define these concepts in a more formal way

Basic terms: Experiment

- A 'probabilistic' **experiment** is basically an activity that we do not know what will happen for sure but we 'observe' what happens
  - a random process
- Some examples:
  - we toss one coin and observe whether it shows heads or tails
  - we will observe the temperature at mid-day tomorrow
  - we will ask someone whether or not they have bought our product in the last 12 months

Outcomes: Sample Spaces

- **The list all the possible outcomes** of an experiment is called the **Sample Space**
  - we will use the letter *S* for sample spaces
  - any collection of items in probability is called a set, and so S is also a set
- If your experiment is rolling a single die: $S = \{1, 2, 3, 4, 5, 6\}$
- If your experiment is tossing a coin twice: $S = \{HH, HT, TH, TT\}$
- Note that an outcome is essentially one of the possible things that can happen in the experiment
  - in any experiment, one and only one outcome occurs

Think…

- What we perceive as random and what is actually random are two separate things
- Assume we flip a coin 10 times, which of the following two sequences is more random?
  - H T H H T H T T H T
  - H T H T T T T T T H

Both outcomes are equally likely, each with probability 1/1024 (we flip a coin 10 times so the total number of possible outcomes is $2^{10} = 1024$)

Types of sample spaces

- **Finite**
  - If you can write & count all elements in S. **Example:** rolling a single die, $S = \{1, 2, 3, 4, 5, 6\}$
- **Countably infinite**
  - There is a way to show the progression of the values in S, but they can go to infinity
  - **Example:** the number of accesses to a web server during a week's time, $S = \{0, 1, 2, 3, 4, \ldots\}$
- **Uncountably infinite**
  - The possible outcomes are too numerous to write down in a listing, so we use an interval to describe them
  - **Example:** length of time it takes a computer to complete a task, with a max of 5 sec. S = {all real numbers x such that $0<x\leq5$}

Events: subsets of sample spaces

- An **event is a set of outcomes and is a subset of the sample space S**
  - The empty set Ø is called the **impossible event**
  - The subset *S* is called the **certain event**
- **We measure the probability of an event: P(event)**
  - probability of the temperature tomorrow mid-day being less than 20 degrees – the event is made up of all outcomes where the temperature is less than 20 degrees – P(temp<20)
  - probability of observing a sum of 3 when rolling 2 dice – the event is made up of all outcomes where the sum of the 2 dice is equal to 3 – P(sum of 2 dice = 3)
  - probability of observing the number 4 when rolling one die – the event is made up of only one outcome – P(rolling a 4)

Sample spaces and events

- We roll a single die, S = {1, 2, 3, 4, 5, 6} :
  - Event A: that we roll an odd number: A = { 1, 3, 5 }
  - Event B: that we roll a number greater than 2: B = { 3, 4, 5, 6 }
- We toss one coin 3 times. What is the sample space S? Use H, T to show the two possible outcomes

  S = {HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}
- In general, when in an 'experiment you have ***n*** **possible outcomes and *k* repetitions, the sample space will consist of n^k events**
  - apply this to the example with the coin above

For the next few slides…

- You will need to remember some of the basic **set theory** covered in Logic & Discrete Structures:
  - **Union** of two events (sets): A ∪ B (i.e. A **or** B)
  - **Intersection** of two events (sets): A ∩ B (i.e. A **and** B)
  - **Complement** of an event (set) A: A′ (i.e. **not** A)
- Example: Roll a single die
  - Event A: roll an odd number
  - Event B: roll a number larger than 2
  - Event C: roll an even number

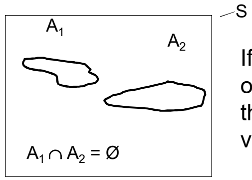
Example continued

- Sample space S = { 1, 2, 3, 4, 5, 6}
- Turn events into sets
  - A = {1, 3, 5 }, B = { 3, 4, 5, 6 }, C = {2, 4, 6 }
- A ∪ B = {1, 3, 4, 5, 6}
- A ∪ C = { 1, 2, 3, 4, 5, 6} = S
- A ∩ B = { 3, 5 }
- A ∩ C = Ø
- A′ = {2, 4, 6} = C

Mutually exclusive events

- Two events A_1 and A_2 are mutually exclusive if and only if $\mathbf{A_1 \cap A_2 = Ø}$
- Events A_1, A_2, A_3, … are mutually exclusive if and only if $\mathbf{A_i \cap A_j = Ø}$ for all i≠j
- **In plain words:**
  - Two events are mutually exclusive if they can't occur at the same time
  - Three or more events are mutually exclusive if every two of them are mutually exclusive

Mutually exclusive events



If you know A_1 has occurred, then you know that A_2 can not occur, and vice versa

One special case of mutually exclusive events is events that are complements of each other - **Why?**

Complement events **BY DEFINITION** have no intersection (i.e. no common elements, they are the 'opposite' of each other) and therefore they are mutually exclusive events

Some examples

- You roll a die once. Let A be the event that the die comes 2, B the event that the die comes an even number and C the event that the die comes an odd number
- A and B are not mut.excl. because $A \cap B = \{2\}$
- A and C are mut.excl. because $A \cap C = \varnothing$
- B and C are mut.excl. (B and C are actually complements of each other)

Some probability laws

- Kolmogorov's axioms of probability
  - If *S* the sample space for an experiment, then ***P*(*S*) = 1**
  - For every event A, **P(*A*)≥0**
  - Let A_1, A_2, A_3 … be a **countable** collection of **mutually exclusive events.** Then: $P(A_1 \cup A_2 \cup A_3 \ldots) = P(A_1) + P(A_2) + P(A_3) + \ldots$
- These 3 axioms are building blocks on which to develop a more complete system of probability theory

Exercise

- One study on the location of pages found on the Web, showed that **35%** of pages were hosted in the USA, **15%** in the UK, **25%** in the rest of Europe and **25%** in the rest of the world. If we pick one page at random, **what is the probability that it will be from the UK or from the US?**

We are given: P(USA)=0.35, P(UK)=0.15, P(EUR)=0.25, P(REST_OF_WORLD)=0.25

We are asked to find $P(UK \cup US)$. Because our events are mutually exclusive, we apply the 3rd axiom, and get that:

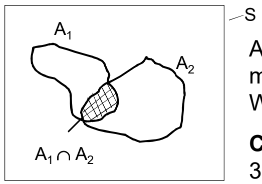
$P(UK \cup US) = P(UK) + P(US) = 0.35 + 0.15 = \mathbf{0.5}$

Some consequences of the axioms

- **P(Ø) = 0**
  - the probability assigned to the impossible event is zero
- **P(A′) = 1 - P(A)**
  - the probability that an event will not occur, is equal to 1 minus the probability that the event will occur
- For any event A, **0 ≤ P(A) ≤ 1**
- These can be proven based on the axioms

The general addition rule

- For any two events A_1 and A_2:
  - **$P(A1 \cup A2) = P(A1) + P(A2) - P(A1 \cap A2)$**



A_1 and A_2 are now **NOT** mutually exclusive events - WHY?

Contrast this rule with the 3rd of Kolmogorov's axioms

Example

- Suppose you roll a die once, and consider the events:
  - **A:** the die comes up an even number
  - **B:** the die comes up greater than 4

Find the probability that the die comes up **an even number OR greater than 4**

A={2, 4, 6}, B={5, 6}

A and B are not mutually exclusive, $A \cap B = \{6\}$

We use the general addition rule to find the probability of the union of the two events:

$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 3/6 + 2/6 - 1/6 = 4/6$

A quick note: Independent events

- Two events are called **independent** if knowledge that one has occurred does not affect the probability of the other event occurring
- For now (in detail next week), we will say that if **P(A $\cap$ B) = P(A)*P(B)** then A and B are independent (**and vice-versa**)
  - This is the **multiplication rule for independence**

Some examples

- Rolling a single die, S = {1, 2, 3, 4, 5, 6}
  - Event A: die comes up odd
  - Event B: die comes up 1
  - Event C: the die is 1 or 2
- Are A and B independent?
- Are A and C independent?

Sample answer

A={1,3,5}, so P(A)=3/6

B={1}, so P(B)=1/6

C={1,2}, so P(C)=2/6

To check if A and B are independent we will check to see if **P(A$\cap$B)=P(A)xP(B)**

A$\cap$B = {1} so P(A$\cap$B) = 1/6

P(A)xP(B) = (3/6)x(1/6) = 3/36 $\neq$ P(A$\cap$B) so A and B are not independent

Similarly for A and C, you will check to see if **P(A$\cap$C)=P(A)xP(C)**

A$\cap$C = {1}, P(A$\cap$C)=1/6

P(A)xP(C) = (3/6)x(2/6)=6/36 = 1/6 = P(A$\cap$C) so A and C are independent

Summary of lecture

- In Week 1 we covered:
  - Basic notion of probability
  - Approaches to calculating probability values
  - Sample spaces and events
  - Basic probability laws
  - How to use these in solving problems
- Don't forget:
  - Work on the exercises for Friday **BEFORE** you come to the class