Lecture 5

INTRODUCTION TO PROBABILITY

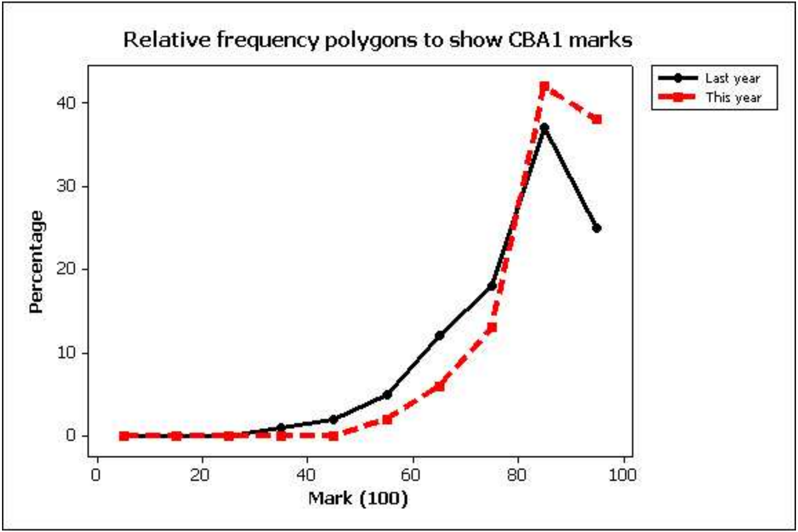
Computer Based Assessments

419 students successfully completed CBA1 in exam mode by midnight last night.

I have chosen to summarise the marks both **graphically** and **numerically**.

n	n^*	Mean	Median	St. dev.	IQR
419	**26**	**93.51**	**96.18**	**8.5**	**10.05**
543	28	93.40	95.32	9.10	11.22
438	28	90.53	93.43	11.60	13.15

Computer Based Assessments



Computer Based Assessments

- Marks will be given out soon
- Be careful – don't try the CBA in exam mode until you've had plenty practice!
- **120** students didn't sit the CBA in exam mode until after 5pm on Friday...
- ...**76** of these students were first–time users!
- Some students still didn't have access to the CBA by the end of last week. These students can be split into two groups...
  - Genuine students who registered late/transferred from other programmes
  - Students who just left it too late to let me know!

Other stuff

Calculators – check the website for a list of University–approved calculators

Videos – now up–and–running!

Introduction

Probability is the language we use to model uncertainty.

We all intuitively understand that few things in life are certain. There is usually an element of **uncertainty** or **randomness** around outcomes of our choices.

In business this uncertainty can make all the difference between a good investment and a poor one.

Hence, an understanding of probability and how we might incorporate this into our decision–making processes is important.

Definitions

We often use the letter P to represent a probability.

For example, $P(Rain)$ would be the probability that it rains.

An **Experiment** is an activity where we do not know for certain what will happen, but we can *observe* what happens. **For example:**

- We will ask someone whether or not they have used our product.
- We will observe the temperature at midday tomorrow.
- We will toss a coin and observe whether it shows "heads" or "tails".

An **Outcome** is one of the possible things that can happen. For example, suppose that we are interested in the shoe size of the next customer to come into a shoe shop.

Possible outcomes include:

- "eight"
- "twelve"
- "nine and a half"

In any experiment, one, *and only one*, outcome occurs.

The **Sample space** is the set of all possible outcomes. For example, it could be the set of all shoe sizes.

An **Event** is a set of outcomes. For example "the shoe size of the next customer is less than 9" is an event. It is made up of all of the outcomes where the shoe size is less than 9.

More stuff...

Probabilities are usually expressed in terms of **fractions**, **decimals** or **percentages**.

Therefore we could express the probability of it raining today as

$$P(Rain) = \frac{1}{20} = 0.05 = 5\%.$$

All probabilities are measured on a scale from zero to one.

- An **impossible** event has a probability of zero
- A **certain** event has a probability of one
- An **evens** event has a probability of 0.5
- Can you imagine where abouts on this scale a **likely** event will lie? Or an **extremely unlikely event**?

The collection of all possible outcomes – the **sample space** – has a probability of 1.

For example:

- Suppose an event has only two outcomes – *success* or *failure*
- Then $P(\textit{success} \text{ or } \textit{failure}) = 1$

Another example:

- Suppose we have a fair six–sided die
- Then $P(1 \text{ or } 2 \text{ or } 3 \text{ or } 4 \text{ or } 5 \text{ or } 6) = 1$

Two events are said to be **mutually exclusive** if both can not occur simultaneously. In the example above, the outcomes *success* and *failure* are mutually exclusive.

Two events are said to be **independent** if the occurrence of one does not affect the probability of the second occurring.

For example, if you toss a coin and look out of the window, the events “get heads” and “it is raining” would be independent.

An example of non–independence

Imagine I go into the Student's Union and pick a student at random.

The events "**the student is female**" and "**the student is studying engineering**" are not independent.

there is a greater proportion of male students on engineering courses than on other courses at the University!

How do we measure Probability?

There are **three** main ways in which we can measure probability:

- **Classical**
- **Frequentist**
- **Subjective/Bayesian**

All three obey the basic rules described so far.

Different people argue in favour of the different views of probability and some will argue that each kind has its uses depending on the circumstances.

Classical probability

This view is based on the concept of **equally likely events**.

If we toss a fair coin, we have two possible outcomes – **Heads** or **Tails**. Both outcomes are **equally likely**. Thus

$$P(Head) = \frac{1}{2} \quad \text{and} \quad P(Tail) = \frac{1}{2}$$

The underlying idea behind this view of probability is **symmetry**.

In this example, there is no reason to think that the outcome *Head* and the outcome *Tail* have different probabilities.

Since there are two outcomes and one of them must occur, both outcomes must have probability $1/2$.

Another commonly used example is **rolling dice**.

There are **six** possible outcomes – (1, 2, 3, 4, 5, 6) – if the die is fair, each of them should have an equal chance of occurring.

Hence $P(1) = \frac{1}{6}$, $P(2) = \frac{1}{6}$, $\ldots$.

What about other calculations, such as $P(\text{Even Number})$?

$$P(\text{Even Number}) = \frac{3}{6} = \frac{1}{2}.$$

This follows from the formula

$$P(\text{Event}) = \frac{\text{Total number of outcomes in which event occurs}}{\text{Total number of possible outcomes}}.$$

Frequentist probability

When the outcomes of an experiment are not equally likely, we can **conduct experiments** to give us an idea of how likely the different outcomes are.

Examples

- Probability of producing a defective item in a manufacturing process
  - We could monitor the process over a long period of time
  - The probability of a defective could be measured by the proportion of defectives in our sample
- Imagine we believed a coin was **unfair**
  - Toss the coin a large number of times
  - See how many heads you obtain, and express $P(Head)$ as a proportion

By conducting experiments the probability of an event can easily be estimated using the following formula:

$$P(\text{Event}) = \frac{\text{Number of times an event occurs}}{\text{Total number of times experiment done}}.$$

The larger the experiment, the closer this probability is to the "**true**" probability.

The frequentist view of probability regards probability as the long run relative frequency (or proportion).

In the defects example, the "true" probability of getting a defective item is the proportion obtained in a very large experiment (strictly an **infinitely** long sequence of trials).

In the frequentist view, probability is a **property of nature**.

In practice we cannot conduct infinite sequences of trials, and so we never know the "true" values of probabilities.

We also have to be able to imagine a long sequence of "**identical**" trials.

This does not seem to be appropriate for "**one-off**" experiments like the launch of a new product.

Subjective/Bayesian Probability

We are probably all intuitively familiar with this method of assigning probabilities:

- When we board an aeroplane, we judge the probability of it crashing to be sufficiently small that we are happy to undertake the journey
- The odds given by bookmakers on a horse race reflect people's beliefs about which horse will win

This view of probability does not fit within the frequentist definition as the race cannot be run a large number of times.

Such ways of thinking about probabilities are **subjective**.

This in itself, some think, is a problem – probabilities which two people assign to the same event can be (very) different.

This becomes important if these probabilities are to be used in **decision making**.

- Suppose you need to decide whether to launch a new product
- You might get market researchers to investigate how likely it is that the product will be a success
- This would affect your decision about whether to launch or not
- We would need a way to reconcile any differences

Example 1

A fast–food chain with 700 outlets describes the geographic location of its restaurants with the following table:

		Region			
		NE	SE	SW	NW
Population	Under 10,000	35	42	21	70
	10,000–100,000	70	105	84	35
	Over 100,000	175	28	35	0

A health and safety organisation selects a restaurant at random for a hygiene inspection.

Example 1

(a) Which of the three approaches to probability would you use to find:
   (i) $P(\text{NE restaurant chosen})$
   (ii) $P(\text{Restaurant chosen from a city with a population} > 100{,}000)$
   (iii) $P(\text{SW and city with a population} < 10{,}000)$?

(b) Now use this approach to find the probabilities given above.

Example 1

Do we have equally likely outcomes?

No!

Therefore we **cannot** use the classical approach for probability.

But we **can** use the

frequentist approach!

Example 1

Using the frequentist approach, we get:

$$
\begin{aligned}
P(\text{NE restaurant}) &= \frac{35+70+175}{700} \\
&= \frac{280}{700} \\
&= 0.4 \qquad \text{and}
\end{aligned}
$$

$$
\begin{aligned}
P(\text{population} > 100{,}000) &= \frac{175+28+35}{700} \\
&= \frac{238}{700} \\
&= 0.34
\end{aligned}
$$

Example 1

Similarly,

$$\begin{aligned} P(\text{SW and population} < 10{,}000) &= \frac{21}{700} \\ &= 0.03 \end{aligned}$$

Example 2

The spinner shown below is spun once.

Example 2

(a) Which of the three approaches to probability would you use to find:
   (i) P(lands red),
   (ii) P(lands triangle), and
   (iii) P(lands quadrilateral)?

(b) Now use this approach to find the probabilities given above.

Example 2

We can use the

Classical approach

to probability, since the six outcomes should be equally likely!

Example 2

Thus, we have:

$$\begin{aligned} P(\text{lands red}) &= \frac{1}{6} \\ P(\text{lands triangle}) &= \frac{1}{6} \\ P(\text{lands quadrilateral}) &= \frac{4}{6} = \frac{2}{3} \end{aligned}$$

LAWS OF PROBABILITY

Multiplication Law

The probability of two **independent** events E_1 and E_2 both occurring is

$$P(E_1 \text{ and } E_2) = P(E_1) \times P(E_2).$$

For example, the probability of throwing a six followed by another six on two rolls of a die is calculated as follows:

- The outcomes of the two rolls of the die are independent
- Let E_1 denote a six on the first roll and E_2 a six on the second roll
- $P(\text{two sixes}) = P(E_1 \text{ and } E_2) = P(E_1) \times P(E_2) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$

This is the **multiplication rule** of probability.

This method of calculating probabilities extends to when there are *many* **independent** events:

$$P(E_1 \text{ and } E_2 \text{ and } \cdots \text{ and } E_n) = P(E_1) \times P(E_2) \times \cdots \times P(E_n).$$

There is a more complicated rule for multiplying probabilities when the events are *not* independent, and we will see this next week.

An elderly woman was assaulted and robbed in an alley in San Pedro, California. A witness saw a blonde woman with a pony–tail running out of the alley and get into a yellow car driven by a hispanic male with a beard and a moustache. A couple answering that description were arrested nearby and brought to trial. The prosecutor calculated

$$\Pr(\text{blonde}) = \frac{1}{3}, \qquad \Pr(\text{pony–tail}) = \frac{1}{10},$$

$$\Pr(\text{beard}) = \frac{1}{10}, \qquad \Pr(\text{moustache}) = \frac{1}{4},$$

$$\Pr(\text{yellow car}) = \frac{1}{10}, \qquad \Pr(\text{hispanic male with white female}) = \frac{1}{1000},$$

so that

$$\Pr(\text{coincidence}) = \frac{1}{3} \times \frac{1}{10} \times \frac{1}{10} \times \frac{1}{4} \times \frac{1}{10} \times \frac{1}{1000} = 1 \text{ in } 12 \text{ million}.$$

Not surprisingly, the verdict was guilty. This evidence was challenged on appeal, however, and the verdict reversed. Why?

Addition Law

The **addition law** describes the probability of any of two **or** more events occurring.

The addition law for two events E_1 and E_2 is

$$P(E_1 \text{ or } E_2) = P(E_1) + P(E_2) - P(E_1 \text{ and } E_2).$$

This describes the probability of *either* event E_1 *or* event E_2 happening.

Example

50% of families in a certain city subscribe to the morning newspaper, 65% subscribe to the afternoon newspaper, and 30% of the families subscribe to both newspapers. **What proportion of families subscribe to at least one newspaper**?

We are told that

- $P(\text{Morning}) = 0.5$
- $P(\text{Afternoon}) = 0.65$ and
- $P(\text{Morning and Afternoon}) = 0.3$

Therefore

$$\begin{aligned} P(\text{at least one paper}) &= P(\text{Morning or Afternoon}) \\ &= P(\text{Morning}) + P(\text{Afternoon}) \\ &\quad - P(\text{Morning and Afternoon}) \\ &= 0.5 + 0.65 - 0.3 \\ &= 0.85. \end{aligned}$$

So 85% of of the city subscribe to at least one of the newspapers.

A more basic version of the rule works where events are **mutually exclusive**.

If events E_1 and E_2 are mutually exclusive then

$$P(E_1 \text{ or } E_2) = P(E_1) + P(E_2).$$

This simplification occurs because when two events are mutually exclusive they cannot happen together and so $P(E_1 \text{ and } E_2) = 0$.

These two laws are the basis of more complicated problem solving we will see later.

A tricky example...

A building has three rooms. Each room has two separate electric lights. There are thus six electric lights altogether. After a certain time there is a probability of 0.1 that a given light will have failed and all lights are independent of all other lights. Find the probability that, after this time, there is at least one room in which both lights have failed.