Core course for *App.AI*, *Bioinfo.*, *Data Sci.&Eng.*, *Dec.Analytics*, *Q.Fin.*, *Risk Mgmt.* and *Stat.* Majors:

STAT2601A Probability and Statistics I (2023-2024 First Semester)

Chapter 1: Introduction

1.1 Introduction to Probability

The study of *probability* is highly related to *uncertainty* and *variability*. Probability has a diverse application in different fields, ranging from medicine, gambling, weather forecast to various scientific and social problems.

Presence of *uncertainty*:

- In finance, an investor would like to assess the *risks* of different investments, especially during uncertain economic periods.
- In casino, a gambler considers what he should put his bet on.
- In climate forecasting, scientists make projections based on past data and probability models.
- In engineering, a technician investigates the chance of electrical circuit overloads.
- In biology, an offspring may or may not inherit a specific genetic disease from the parents.
- In quality control, a manager needs to make sure certain percentage of product are under satisfaction.
- When you throw ten dice together, you may be interested in the possible patterns of points combinations.
- You may decide on how much time you should spend on studying this course in order to pass it.
- You may be graduating with First Class Honours, or you may be forced to discontinue your study due to poor GPA. =(

Probability

- Objective and quantitative assessment of how certain (or how likely) an event will occur.
- Basic tool in statistical inference.

What is the difference between *probability* and *statistics*?

Example 1.1. (Probability vs Statistics)

	Probability Theory	Statistical Inference
1.	We have a fair coin. (i.e., Pr(Head) = Pr(Tail) = 0.5)	We have a coin.
2.	The fair coin is tossed ten times.	This coin is tossed ten times.
3.	Pr(all are head) = ?	All are head. $\Longrightarrow$ Is it a fair coin?

From probability theory, if the coin is fair, then Pr(all are head) $= (0.5)^{10}$.

If it is a fair coin, it is very unlikely for us to observe ten heads in a row.

Thus, if we observe that happens, for statistical inference, we may reach a conclusion that the coin is not a fair coin.

Probability and statistics

- science or art of understanding variability and uncertainty.

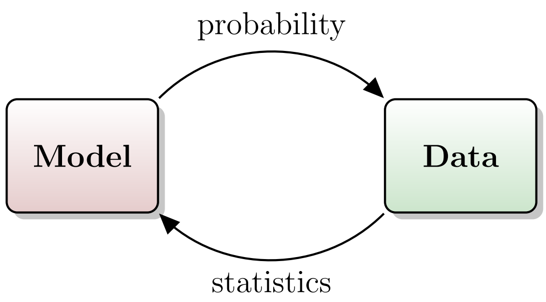
The subject of *probability* (focus of STAT2601) seeks to

- study properties of **models** thought to describe/approximate real-life random phenomena, usually in a mathematical framework;
- provide a rigorous mathematical basis for rules that govern calculation of uncertainty, defining a "calculus of variability".

The subject of *statistics* (focus of STAT2602) seeks to

- study data obtained from real-life experiments/observation studies, draw evidence to support certain modelling conjectures, infer from them general conclusions and/or make decisions;
- provide tools and methodologies for data analysis, which are firmly justified with the help of probability theory.

A simplified link between probability and statistics:



1.2 Origin of Probability

1.2.1 Historical Event

The well-known gambler's problem:

Time:	Year 1654
Place:	France
Persons:	C. de Méré, B. Pascal, P. de Fermat

1.2.1.1 Version 1

- Game 1

  Méré bets even money that he would throw at least one six in four throws of the die. (i.e., Méré will win if at least one six turn out in four throws.)

- Game 2

  Méré bets even money that a double-six would appear at least once if he were given twenty-four throws of two dice.

It is believed that these two games are indifferent to Méré according to the following argument:

	No. of possible outcomes	No. of throws
Game 1	6	4
Game 2	$6 \times 6 = 36$	$6 \times 4 = 24$

However, experience showed that Méré will win more likely in Game 1 than in Game 2.

1.2.1.2 Version 2

- Two persons, Fermat and Pascal, are involved in the following game:

  A fair coin is to be tossed. If the outcome is Head, Fermat can get one point; otherwise Pascal can get one point. The one who first gets 4 points can win a prize of \$100. Suppose after five tosses Fermat got 3 points and Pascal got 2 points and the game was interrupted for some reason. What will be the reasonable division of the prize to the players Fermat and Pascal?

1.2.1.3 Solutions

In 1654, Méré posed the problem to a young French mathematician named B. Pascal. Pascal had discussed this problem with another mathematician, P. de Fermat. Through their discussions they did not only come up with a convincing and self-consistent solution to this problem, but also developed concepts that continue to be fundamental in probability to this day.

- Version 1 Solution

$$\begin{aligned}
\Pr(\text{Méré wins in Game 1}) &= 1-\left(\frac{5}{6}\right)^4 = 0.51775,\\
\Pr(\text{Méré wins in Game 2}) &= 1-\left(\frac{35}{36}\right)^{24} = 0.4914.
\end{aligned}$$

Therefore it is reasonable that the player has more advantage in Game 1 than in Game 2.

- Version 2 Solution

According to Fermat and Pascal, the division should depend on the possible ways the game might have continued, i.e., depend on the chance that each player would win if the game were not interrupted.

$$\begin{aligned}
\Pr(\text{Pascal will win if the game continues}) &= \Pr(\text{two tails}) = 0.25,\\
\Pr(\text{Fermat will win if the game continues}) &= 1-\Pr(\text{two tails}) = 0.75.
\end{aligned}$$

Hence Fermat should get \$75 and Pascal should get \$25.

Based on modern probability theory, the general form of such problem (nowadays known as the *problem of points*) can be solved by using the negative binomial distribution, which will be introduced in a later chapter.

1.2.2 The Classical School and the Principle of Indifference

Definition 1.1.
Suppose a single trial in a chance situation can have one of N equally likely outcomes such that for each trial, one and only one outcome will occur. If f of the N possibilities are favourable to a specified event E, then the probability of the occurrence of event E is defined and denoted as

$$\Pr(E) = \frac{f}{N}.$$

◀

Example 1.2.

The probability of drawing a diamond from a deck of 52 cards is

$$\Pr(\text{a diamond is drawn}) = \frac{13}{52} = \frac{1}{4}.$$

★

Example 1.3. (Birthday problem)

Suppose we randomly draw n people from the whole population. What is the probability that no two persons in this sample have the same birthday?

Solution:

$$\begin{aligned} N &= 365 \times 365 \times \cdots \times 365 = 365^n \\ f &= 365 \times 364 \times 363 \times \cdots \times (365 - n + 1) = {}_{365}P_n \quad \text{(notation of permutation)} \end{aligned}$$

Therefore,
$$\Pr(\text{all different birthdays}) = \frac{{}_{365}P_n}{365^n}.$$

If $n = 23$, then the probability is 0.4927 which is less than a half.

In other words, with 23 people, the probability that some pair of them will have the same birthday would be greater than half (0.5073).

Thought Question:

From the above calculations, what are the implicit assumptions behind the solution or problem setting?

1.2.2.1 Principle of Indifference

How do we know that the outcomes are equally likely to begin with?

1. There is natural symmetry of the experiment and thus the outcome must be equal probable.
2. *Principle of indifference / Principle of insufficient reason*

   If there is no reason to believe that one outcome is any more likely than any other, then we should assume that they are equally likely.

In applying the classical assumption of equally likely outcomes, it is critically important that we use the "correct" information on the experiment.

Example 1.4.

Two fair coins are tossed together.

Possible outcomes:

$$\{\text{two heads, one head one tail, two tails}\} \qquad \Pr(\text{two heads}) = \tfrac{1}{3}?$$

or

$$\{(\text{head, head}), (\text{head, tail}), (\text{tail, head}), (\text{tail, tail})\} \quad \Pr(\text{two heads}) = \tfrac{1}{4}?$$

Experimental evidence suggests that the probability of getting two heads is $1/4$, not $1/3$.

The set of possible outcomes {two heads, one head one tail, two tails} is not appropriate because it suppresses information about the random experiment by not distinguishing between the coins when it is possible to do so.

This suggests that to apply the classical assumption of equally likely outcomes correctly, one must choose the set of possible outcomes in such a way that all information about the random experiment is captured.

1.2.3 The Frequency School (The Law of Average)

Definition 1.2.
Probability was viewed in terms of *relative frequency* when the basic process is repeated over and over again, independently and under the same conditions. If an experiment is repeated n times and event E occurs in n_E times, then the probability of occurrence of E is defined as the limit of the relative frequency:

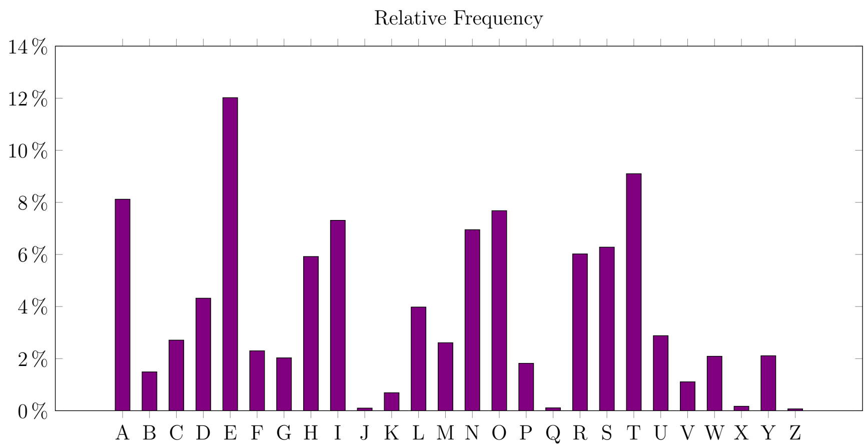
$$\frac{n_E}{n} \to \Pr(E) \qquad \text{as } n \to \infty. \qquad \blacktriangleleft$$

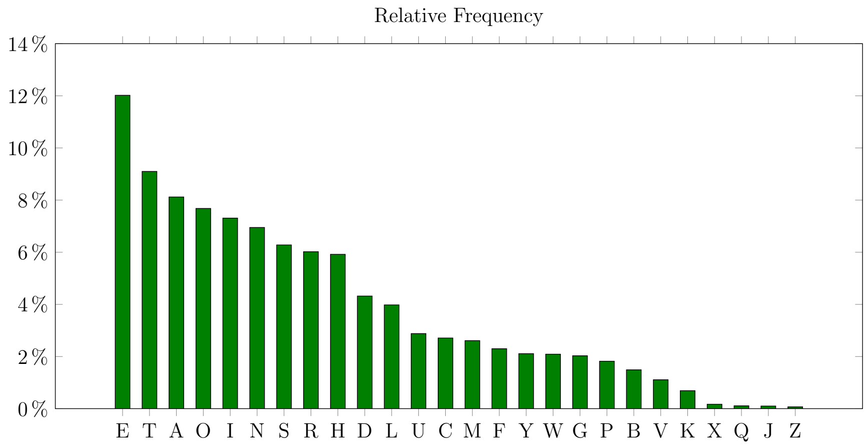
Example 1.5. (English letter count)

We may examine the frequencies of letters used in typical English words. The probability of occurrence of each letter can be estimated by the corresponding relative frequency.

The following data were obtained based on a sample of 40000 English words.

Letter	Count	Relative Frequency
A	14810	8.12%
B	2715	1.49%
C	4943	2.71%
D	7874	4.32%
E	21912	12.02%
F	4200	2.30%
G	3693	2.03%
H	10795	5.92%
I	13318	7.31%
J	188	0.10%
K	1257	0.69%
L	7253	3.98%
M	4761	2.61%
N	12666	6.95%
O	14003	7.68%
P	3316	1.82%
Q	205	0.11%
R	10977	6.02%
S	11450	6.28%
T	16587	9.10%
U	5246	2.88%
V	2019	1.11%
W	3819	2.09%
X	315	0.17%
Y	3853	2.11%
Z	128	0.07%
Total	182303	100.00%





The probabilities may be estimated by the relative frequencies:

$$\begin{aligned} \Pr(\mathrm{A}) &= 8.12\%, \\ \Pr(\mathrm{B}) &= 1.49\%, \\ \Pr(\mathrm{C}) &= 2.71\%, \\ &\vdots \end{aligned}$$

★

Remarks

1. It is impossible to observe an infinite series in reality and to provide empirical confirmation.
2. In this point of view, probability can only be assigned to results of experiments that can, at least in principle, be conducted repeatedly.
3. For fixed number of experiments, the relative frequency is called the *empirical probability*. It provides a way to estimate the probability (by relative frequency). The estimate will be "good" if n is large.

1.2.4 The Subjective School

Definition 1.3.
Probability is quantitative expressions of uncertainty about a person's knowledge of the occurrence of some event.

◀

From the subjective point of view, we emphasize the uncertainty of our knowledge rather than the uncertainty of the event's occurrence. Even if the event has occurred, if we do not know about it, the event will still be "uncertain" in our view.

Example 1.6.

Consider the following statements.

"I thought there is 25% chance that tomorrow will rain."

"Manchester United will have 60% chance to win the match against Manchester City next week."

"I wish I had worked hard in last semester, I think I should have 80% chance to get better grade if I did so."

★

Classical statisticians argue that subjective interpretation suffer from a lack of objectivity because different individuals are free to assign different probabilities to the same event according to their own personal opinions. However, the classical interpretation has built-in subjectivity (e.g., assuming equally likely outcomes) and that the advantage of the subjective approach is that the subjectivity is made explicitly.

1.2.4.1 Elementary Gambling Situation (EGS)

An elementary gambling situation is an agreement in which you receive W if E occurs and you pay L dollars if the event E does not occur. The amounts L and W are called the *stakes* in the gamble; the ratio $W : L$ is called the *odds* (against); and the fraction that your stake bears to the total is called the *betting quotient*,

$$q = \frac{L}{L+W}.$$

Example 1.7.

In watching a particular soccer match between Manchester United and Manchester City, you proposed to your friend that if Manchester United wins, then you can obtain \$100 from your friend; otherwise you should pay \$150 to your friend.

Event E:	Manchester United wins the match
Your stake:	\$150
Your friend's stake:	\$100
Odds (against) of your bet:	2/3
Betting quotient:	0.6

There is a place where you would be indifferent as to which side of the bet you would take: bet on the occurrence of E at odds $W : L$ or bet on the non-occurrence of E at odds $L : W$. For such a value of q, the bet is called a *fair bet*, and the stakes set up as $W : L$ are called *fair-betting odds*. The value of q is then your value of (subjective) probability of occurrence of E, i.e., $\Pr(E)$.

Reason:

If the game is played N times, we will expect:

$$\begin{array}{rll} E \text{ occurs} & N \times \Pr(E) \text{ times} & \Longrightarrow \text{Gain is } W \times N \times \Pr(E) \\ E \text{ does not occur} & N \times [1-\Pr(E)] \text{ times} & \Longrightarrow \text{Loss is } L \times N \times [1-\Pr(E)] \end{array}$$

For a fair game,

$$\text{Gain} = \text{Loss}.$$

Therefore,

$$\begin{aligned} W \times N \times \Pr(E) &= L \times N \times [1-\Pr(E)] \\ \Pr(E) &= \frac{L}{L+W} = q. \end{aligned}$$

★

~ **End of Chapter 1** ~