THE CHINESE UNIVERSITY OF HONG KONG
Department of Mathematics

Mathematical Modelling Project Team
mathmodel@math.cuhk.edu.hk

HSMMC Pre-workshop Learning Materials (Probability and Statistics)

Last updated: March 23, 2026

Background 背景

Probability and statistics are frequently used in mathematical modelling for:
概率與統計在數學建模中經常用於：

- Modelling events with randomness
  對具有隨機性的事件進行建模
- Analyzing the correlation between factors
  分析因素之間的相關性
- ...

Here, we cover some basic concepts of probability and statistics.
在此，我們涵蓋概率與統計的一些基本概念。

1 Probability 概率

Probability gives a numerical measure for the degree of uncertainty or the occurrence of an event.
概率是對不確定性程度或事件發生可能性的一種數值度量。

If there are n total possible outcomes in a sample space S, and m of those are favorable for an event A, then the probability of event A is given as
若樣本空間 S 中共有 n 個可能結果，其中 m 個結果對事件 A 有利，則事件 A 的概率為

$$P(A) = \frac{\text{number of favorable outcomes}}{\text{total number of possible outcomes}} = \frac{n(A)}{n(S)} = \frac{m}{n}$$

Example 例子

Find the probability of getting a 3 or 5 when throwing a die.
求擲一顆骰子擲出 3 或 5 的概率。

Solution. Sample space $S = \{1, 2, 3, 4, 5, 6\}$ and event $A = \{3, 5\}$.
We have $n(A) = 2$ and $n(S) = 6$. So, $P(A) = \frac{n(A)}{n(S)} = \frac{1}{3}$.
解：樣本空間 $S = \{1, 2, 3, 4, 5, 6\}$，事件 $A = \{3, 5\}$。
我們有 $n(A) = 2$，$n(S) = 6$。因此 $P(A) = \frac{n(A)}{n(S)} = \frac{1}{3}$。

2 Set Notation 集合記號

2.1 Complement 餘集

The **complement** of event A is the set of all outcomes in a sample that are not included in the event A. The complement of event A is denoted by A'.
事件 A 的**餘集**是樣本中不包含在事件 A 內的所有結果所組成的集合。事件 A 的餘集記為 A'。

2.2 Intersection 交集

The event $A \cap B$ is the **intersection** of the events A and B and consists of outcomes that are contained within both events A and B.
事件 $A \cap B$ 是事件 A 與 B 的**交集**，由同時包含於事件 A 和事件 B 中的結果所組成。

2.3 Mutually Exclusive 互斥

Two events are said to be **mutually exclusive** if $A \cap B = \emptyset$.
若 $A \cap B = \emptyset$，則稱兩事件互斥。

2.4 Union 聯集

The event $A \cup B$ is the **union** of events A and B and consists of the outcomes that are contained within at least one of the events A and B.
事件 $A \cup B$ 是事件 A 與 B 的**聯集**，由包含於事件 A 或事件 B（或兩者）中的結果所組成。

2.5 (Theorem) Distributive Laws 分配律

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

and

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

2.6 (Theorem) De Morgan's Law 笛摩根法則

$$(A \cap B)' = A' \cup B'$$
$$(A \cup B)' = A' \cap B'$$

3 Law of probability 概率法則

3.1 Axioms of Probability 概率公理

For an experiment with sample space $S = \{e_1, e_2, \cdots, e_n\}$, we can assign probability $P(e_1), P(e_2), \cdots, P(e_n)$ provided that
對於樣本空間 $S = \{e_1, e_2, \cdots, e_n\}$ 的試驗，我們可賦予概率 $P(e_1), P(e_2), \cdots, P(e_n)$，前提是

1. $0 \leq P(e_i) \leq 1$,
2. $P(e_1) + P(e_2) + \cdots + P(e_n) = 1$.

3.2 (Theorem) Complement Rule 餘集法則

$$P(A') = 1 - P(A)$$

3.3 (Theorem) Addition Law 加法法則

If A and B are two different events then
若 A 和 B 為兩個不同的事件，則

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

4 Counting rules 計數法則

4.1 (Theorem) Permutations 排列

The number of permutations of n distinct objects taken r at a time is
從 n 個相異物件中取出 r 個的排列數為

$$_nP_r = \frac{n!}{(n-r)!}$$

4.2 (Theorem) Combinations 組合

The number of distinct subsets or combinations of size r that can be selected from n distinct objects, $(r \leq n)$, is given by
從 n 個相異物件中選取大小為 r（$r \leq n$）的不同子集或組合的數目為

$$_nC_r = \binom{n}{r} = \frac{n!}{r!(n-r)!}$$

5 Conditional probability and independence 條件概率與獨立性

5.1 Conditional Probability 條件概率

The **conditional probability** of event A given B is
事件 A 在給定 B 下的**條件概率**為

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \text{ for } P(B) > 0$$

5.2 Independence 獨立性

The events A and B are called **independent** if
事件 A 與 B 稱為**獨立**若

$$P(A \cap B) = P(A)P(B)$$

5.3 Partition 分割

Events $B_1, B_2, \cdots, B_k$ are said to be a **partition** of the sample space S if the following two conditions are satisfied.
事件 $B_1, B_2, \cdots, B_k$ 稱為樣本空間 S 的一個**分割**，若滿足以下兩個條件：

1. $B_i \cap B_j = \emptyset$ for each pair i, j
2. $B_1 \cup B_2 \cup \cdots \cup B_k = S$

5.4 (Theorem) Bayes Rule 貝氏定理

If $B_1, B_2, \cdots, B_k$ form a partition of the sample space S such that $P(B_i) \neq 0$ for $i = 1, 2, \cdots, k$, then for any event A of S,
若 $B_1, B_2, \cdots, B_k$ 構成樣本空間 S 的一個分割，且對 $i = 1, 2, \cdots, k$ 有 $P(B_i) \neq 0$，則對 S 中的任意事件 A，

$$P(A) = \sum_{i=1}^{k} P(B_i \cap A) = \sum_{i=1}^{k} P(B_i)P(A|B_i)$$

Subsequently,
進而，

$$P(B_j|A) = \frac{P(B_j)P(A|B_j)}{P(A)}$$

6 Discrete probability distributions 離散概率分佈

6.1 Random Variable 隨機變量

A **random variable (RV)** is a number associated with each outcome of some random experiment. A random variable X is said to be **discrete** if it can take on only a finite or countable number of possible values x.
隨機變量 (RV)是與某隨機試驗的每個結果相關聯的一個數值。若隨機變量 X 只能取有限個或可數個可能值 x，則稱 X 為**離散**隨機變量。

Example 例子

Toss two coins and record the number of heads: $0, 1$ or 2. Then, the following outcomes can be observed.
投擲兩枚硬幣並記錄正面出現的次數：$0, 1$ 或 2。則可觀察到以下結果。

Outcome	TT	HT	TH	HH
Number of heads	0	1	1	2

The random variables will be denoted by capital letters $X, Y, Z, \cdots$, and the lowercase x will represent a particular value of X. For the above example, $x = 2$ if heads comes up twice. Now, we want to look at the probabilities of the outcomes. For the probability that the random variable X has the value x, we write $P(X = x)$, or just $p(x)$.
隨機變量會以大寫字母 $X, Y, Z, \cdots$ 表示，小寫 x 則代表 X 的某個特定值。在上述例子中，若正面出現兩次，則 $x = 2$。現在，我們希望考察結果的概率。隨機變量 X 取值 x 的概率記為 $P(X = x)$，或簡記為 $p(x)$。
For the coin flipping random variable X, we can make the table:
對於上述擲硬幣的隨機變量 X，我們可列出下表：

x	0	1	2
$p(x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

The table represents the **probability distribution** of the random variable X.
該表格代表了隨機變量 X 的**概率分佈**。

6.2 Probability Mass Function 概率質量函數

The function $p_X(x)$ or simply $p(x)$ is called **probability mass function (PMF)** of X if it satisfies:
函數 $p_X(x)$ 或簡記為 $p(x)$ 稱為 X 的**概率質量函數 (PMF)**，若它滿足：

1. $P(X = x) = p_X(x) \geq 0$
2. $\sum_x P(X = x) = 1$, where the sum is over all possible x
   $\sum_x P(X = x) = 1$，其中求和是對所有可能的 x 進行的

Example 例子

A shipment of 8 computers contains 3 that are defective. If a school makes a random purchase of 2 of these computers, find the PMF for the number of defectives.
一批 8 部電腦中有 3 部是壞的。若一所學校隨機購買其中 2 部電腦，求壞電腦數量的概率質量函數 (PMF)。

Solution. Let X be a random variable whose values x are the possible numbers of defective computers purchased by the school. Then x must be $0, 1$ or 2. For each case, we have
解：設 X 為隨機變量，其值 x 為學校購買的壞電腦的可能數量。則 x 必須為 $0, 1$ 或 2。對每種情況，我們有

$$P(X=0) = \frac{\binom{3}{0}\binom{5}{2}}{\binom{8}{2}} = \frac{10}{28}$$

$$P(X=1) = \frac{\binom{3}{1}\binom{5}{1}}{\binom{8}{2}} = \frac{15}{28}$$

$$P(X=2) = \frac{\binom{3}{2}\binom{5}{0}}{\binom{8}{2}} = \frac{3}{28}$$

Thus, the PMF of X is given by
因此，X 的概率質量函數為

x	0	1	2
$p(x)$	$\frac{10}{28}$	$\frac{15}{28}$	$\frac{3}{28}$

6.3 Cumulative Distribution Function 累積分佈函數

The **cumulative distribution function (CDF)** $F(x)$ for a random variable X is defined as
隨機變量 X 的**累積分佈函數 (CDF)** $F(x)$ 定義為

$$F(x) = P(X \leq x)$$

If X is discrete,
若 X 是離散的，則

$$F(x) = \sum_{y \leq x} p(y)$$

where $p(x)$ is the probability mass function.
其中 $p(x)$ 為概率質量函數。

Example 例子

Find the CDF of the random variable in the last example.
求上例中隨機變量的累積分佈函數 (CDF)。

Solution. The CDF of the random variable X is:
解：隨機變量 X 的累積分佈函數 (CDF) 為：
$F(0) = p(0) = \frac{10}{28}$
$F(1) = p(0) + p(1) = \frac{10}{28} + \frac{15}{28} = \frac{25}{28}$
$F(2) = p(0) + p(1) + p(2) = \frac{10}{28} + \frac{15}{28} + \frac{3}{28} = 1$
Hence,
因此，

$$F(x) = \begin{cases} 0 & \text{for } x < 0 \\ \frac{10}{28} & \text{for } 0 \leq x < 1 \\ \frac{25}{28} & \text{for } 1 \leq x < 2 \\ 1 & \text{for } x \geq 2 \end{cases}$$

6.4 Expected Value 期望值

The **mean** or **expected value** of a discrete random variable X with probability mass function $p(x)$ is given by
具有機率質量函數 $p(x)$）的離散隨機變量 X 的**平均值**或**期望值**為

$$\mathbb{E}(X) = \sum_x xp(x)$$

6.5 Variance 方差

The **variance** of a random variable X with expected value μ is given by
期望值為 μ 的隨機變量 X的**方差**為

$$V(X) = \sigma^2 = \mathbb{E}(X - \mu)^2 = \mathbb{E}(X^2) - \mu^2,$$

where
其中

$$\mathbb{E}(X^2) = \sum_x x^2 p(x)$$

6.6 Standard Deviation 標準差

The **standard deviation** of a random variable X is the square root of the variance, and is given by
隨機變量 X 的**標準差**是方差的平方根，即

$$\sigma = \sqrt{V(X)} = \sqrt{\mathbb{E}(X - \mu)^2}$$

Remark. The mean describes the *center* of the probability distribution, while the standard deviation describes the *spread*.
注釋： 均值描述概率分佈的中心，而標準差描述離散程度。

Example 例子

The number of fire emergencies at a rural county in a week has the following distribution
某鄉村地區一星期內發生火災緊急事件的次數有如下分佈

x	0	1	2	3	4
$P(X=x)$	0.52	0.28	0.14	0.04	0.02

Find $\mathbb{E}(X)$, $V(X)$ and σ.
求 $\mathbb{E}(X)$、$V(X)$ 和 σ。

Solution. By definition, we see that
解：根據定義，可得

$$\mathbb{E}(X) = 0(0.52) + 1(0.28) + 2(0.14) + 3(0.04) + 4(0.02) = 0.76 = \mu$$

$$\mathbb{E}(X^2) = 0^2(0.52) + 1^2(0.28) + 2^2(0.14) + 3^2(0.04) + 4^2(0.02) = 1.52$$

$$V(X) = 1.52 - (0.76)^2 = 0.9424$$

$$\sigma = \sqrt{0.9424} \approx 0.9708$$

7 Continuous probability distributions 連續概率分佈

All of the random variables discussed previously were discrete, meaning they can take only a finite (or, at most, countable) number of values. However, many of the random variables seen in practice have more than a countable collection of possible values. For example, the metal content of ore samples may run from 0.10 to 0.80. Such random variables can take any value in an interval of real numbers. Since the random variables of this type have a continuum of possible values, they are called **continuous random variables**.
前面討論的所有隨機變量都是離散的，即它們只能取有限個（或至多可數個）值。然而，許多實際中遇到的隨機變量有超過可數個的可能值。例如，礦石樣本的金屬含量可能在 0.10 到 0.80 之間。這類隨機變量可以取實數區間中的任何值。由於這類隨機變量有連續多個可能值，它們被稱為**連續隨機變量**。

7.1 Probability Density Function 概率密度函數

The function $f(x)$ is a **probability density function (PDF)** for the continuous random variable X, defined over the set of real numbers, if
函數 $f(x)$ 是定義在實數集上的連續隨機變量 X 的**概率密度函數 (PDF)**，若滿足

1. $f(x) \geq 0$ for all x,
2. $\int_{-\infty}^{\infty} f(x)dx = 1$, and
3. $P(a \leq x \leq b) = \int_a^b f(x)dx$.

Remark. What does this actually mean? Since continuous probability functions are defined for an infinite number of points, the probability at a single point is always **zero**! Probabilities are measured over intervals, not single points. That is, the area under the curve between two distinct points defines the probability for that interval. This means that the height of the probability function can in fact be greater than one. The property that the integral must equal one is equivalent to the property for discrete distributions that the sum of all the probabilities must equal one.
注釋： 這實際意義為何？由於連續概率函數定義在無限多個點上，在單一點上的概率總是**零**！概率是在區間上而非單點上度量的。也就是說，曲線下兩點之間的面積定義了該區間的概率。這意味著概率函數的高度實際上可以大於一。積分必須等於一的性質等價於離散分佈中所有概率之和必須等於一的性質。

7.2 Cumulative Distribution Function 累積分佈函數

The **cumulative distribution function (CDF)** $F(x)$ of a continuous random variable X, with density function $f(x)$, is

密度函數為 $f(x)$ 的**累積分佈函數 (CDF)** $F(x)$ 的連續隨機變量 X 為

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)dx$$

Remark. As an immediate consequence of the above equation, one can write these two results:
注釋： 作為上述等式的直接推論，可以得到以下兩個結果：

1. $P(a \leq x \leq b) = F(b) - F(a)$.
2. $f(x) = F'(x)$, if the derivative exists.
   $f(x) = F'(x)$，若導數存在

Further readings 延伸閱讀

- Probability and Statistics self-learning materials on Khan Academy:
  Khan Academy的概率與統計自學材料：
  https://www.khanacademy.org/math/statistics-probability
- Probability and Statistics self-learning materials on MIT OpenCourseWare:
  MIT OpenCourseWare 的概率與統計自學材料：
  https://ocw.mit.edu/courses/18-05-introduction-to-probability-and-statistics-spring-2022/