Ch1 Probability and Distribution

Jih-Chang Yu

August 8, 2023

Set Theory

Definition of Outcome

The outcome refers to a possible result or event that can occur in an experiment or study. In the context of probability and statistics, outcomes are the different possible values that a random variable can take.

For example, if we toss a fair coin, the possible outcomes are "heads" and "tails". In a six-sided die roll, the outcomes are the numbers 1 to 6.

Definition of Sample Space

- The sample space refers to the set of all possible outcomes in an experiment or study. It represents the complete range of outcomes that can occur.

For example, when rolling a six-sided die, the sample space is {1, 2, 3, 4, 5, 6}. Each number represents a possible outcome.

{1}
{1, 2}
{1, 3}, {1.4}, {1.5}, {1.6}
{2.3} . . . {5, 6}

Definition of Event

An event is a specific outcome or a collection of outcomes that we are interested in studying. It is a subset of the sample space, consisting of one or more outcomes.

For instance, in rolling a six-sided die, the event of getting an even number can be represented as $\{2, 4, 6\}$.

- In set theory, a subset is a collection of elements that are all part of a larger set. A subset is formed by selecting some or all of the elements from the original set.

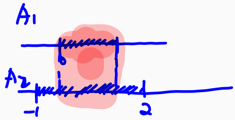
Formally, if we have a set A and another set B, A is said to be a subset of B if every element in A is also an element in B. We represent this as $A \subseteq B$. For example, let's consider the set of even numbers $E = \{2, 4, 6, 8, 10, ...\}$ and the set of natural numbers $N = \{1, 2, 3, 4, 5, ...\}$. Here, E is a subset of N because every element in E (even numbers) is also an element in N (natural numbers).

$E \subseteq N$

Examples 1 and 2

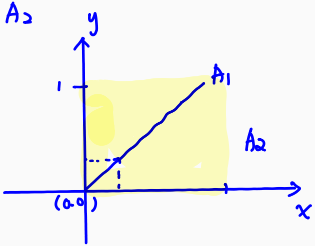
Ex1: $A_1 = \{x ; 0 \le x \le 1\}$, $A_2 = \{x ; -1 \le x \le 2\}$

Q: $A_1 \subseteq A_2$ or $A_2 \subseteq A_1$



Ex2: $A_1 = \{(x,y) ; 0 \le x = y \le 1\}$, $A_2 = \{(x,y) ; 0 \le x \le 1, \ 0 \le y \le 1\}$

Q: $A_1 \subseteq A_2$



Definition of Null Set

ϕ

In set theory, a null set (or empty set) is a set that contains no elements. It is denoted by the symbol $\emptyset$ or $\{\}$. Formally, for any given set A, the null set is a subset of A if and only if no element belongs to the null set.

Definition of Union

In set theory, the union of two sets A and B is a set that contains all the elements present in either A, B, or both. It is denoted by the symbol $\cup$.

Formally, the union of sets A and B is defined as:

$$C = A \cup B = \{x : x \in A \text{ or } x \in B\}$$

For example, let's consider the sets $A = \{1, 2, 3\}$ and $B = \{3, 4, 5\}$. The union of A and B would be $A \cup B = \{1, 2, 3, 4, 5\}$.

$\{1, 2, 3, \cancel{3}\ 4\ 5\}$

Examples 3-7

Ex3,

$A_1 = \{1, 2, 3\}$, $A_2 = \{3, 4, 5\}$

$A_1 \cup A_2 = \{1, 2, 3, 4, 5\}$

而不是 $\{1, 2, 3, \underline{\underline{3}}, 4, 5\}$

Ex4. $A_1 \subseteq A_2$

$A_1 \cup A_2 = A_2$

Ex5: $A_1 = \phi$, $A_2 =$ 隨意.

$A_1 \cup A_2 = A_2$

$\therefore \phi \subseteq A_2$

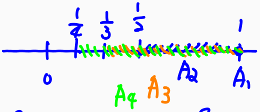
Ex6: $A \cup A = A$

Ex7: $A_k = \{x ; \frac{1}{1+k} \le x \le 1\}$

Then $A_1 \cup A_2 \cup A_3 \cup \cdots$

$= \{x : 1 \le x \le 1\} \cup \{x : \frac{1}{2} \le x \le 1\} \cup$

$\{x : \frac{1}{3} \le x \le 1\} \cdots$



$= \{x ; 0 < x \le 1\}$

Definition of Intersection

In set theory, the intersection of two sets A and B is a set that contains only the elements that are present in both A and B. It is denoted by the symbol $\cap$.

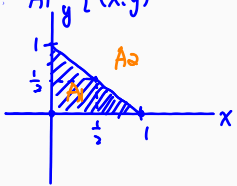
Formally, the intersection of sets A and B is defined as:

$$A \cap B = \{x : x \in A \text{ and } x \in B\}$$

For example, let's consider the sets $A = \{1, 2, 3\}$ and $B = \{3, 4, 5\}$. The intersection of A and B would be $A \cap B = \{3\}$.

Examples 9-11

Ex9: $A_1 = \{(x,y) : 0 \le x+y \le 1\}$, $A_2 = \{(x,y) : 1 < x+y\}$



$A_1 \cap A_2 = \phi$.

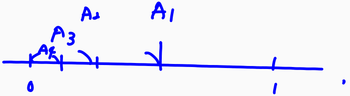
$A_1 \cup A_2 = \mathbb{R}^2$

Ex10: For every set A, $A \cap A = A$. and $A \cap \phi = \phi$

Ex11: $A_k = \{x : 0 < x < 1/k\}$, $k = 1, 2, 3, \cdots$

$A_1 \cap A_2 \cap A_3 \cap A_4 \cdots$

$\{x : 0 < x < 1\} \cap \{x : 0 < x < \frac{1}{2}\} \cap \{x : 0 < x < \frac{1}{3}\} \cap \cdots$



$A_{100} = \{x : 0 < x < 0.01\}$

$A_\infty = \{x : 0 < x < 0\} = \phi$.

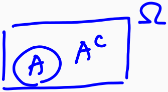
Definition of Space in Probability Theory

- In probability theory, a sample space represents the set of all possible outcomes of an experiment or a random phenomenon. It is denoted by the symbol Ω.
- The sample space is a fundamental concept that provides the foundation for probability calculations. It captures the complete range of possible results and forms the basis for defining events and calculating probabilities associated with those events.
- For example, consider a coin flip experiment. The sample space for this experiment is $\Omega = \{H, T\}$, where H represents heads and T represents tails. The sample space includes all possible outcomes of the coin flip.

  In more complex scenarios, the sample space may consist of multiple dimensions or elements, depending on the nature of the experiment.

Definition of Complement

- In set theory and probability theory, the complement of a set A refers to the elements that do not belong to A within a given universal set U. It is denoted by the symbol $\cancel{\bar{A}}$ or A^c.
- Formally, the complement of set A with respect to universal set U is defined as:



$$A' = \{x : x \in U \text{ and } x \notin A\}$$

- In probability theory, the complement of an event E represents all the outcomes that are not part of event E. It is often denoted by $\cancel{\bar{E}}$ or E^c. For example, let's consider a universal set U of integers from 1 to 10, and a set $A = \{2, 4, 6\}$. The complement of A would be $A^c = \{1, 3, 5, 7, 8, 9, 10\}$.

Probability of Set

Definition of Probability

In probability theory, the probability of an event C_i, denoted by $P(C_i)$, is a measure of the likelihood or chance that the event will occur. It quantifies the uncertainty associated with the occurrence of the event, which satisfies:

Axiom of Probability

If $P(C)$ is defined for a type of subset of space Ω and if

(a) $P(C) \geq 0$

(b) C_i, C_j, ∃ $C_i \cap C_j = \emptyset$

$$P(C_1 \cup C_2 \cup C_3 \cup \cdots) = P(C_1) + P(C_2) + P(C_3) + \cdots$$

(c) $P(\Omega) = 1$.

then P is called probability set function of the outcome of the random variable.

Theorem 1

For each $C \subset \Omega$, $P(C) = 1 - P(C^c)$

$C^c = \{x : \underset{\sim\sim}{x \in \Omega} \text{ and } \underline{x \notin C}\}$

ex: 紅 $O \times 7$

藍 $O \times 3$

若 C 表抽到紅球，

C^c 表示抽到藍球。

$\because C^c \cup C = \Omega$, $C^c \cap C = \phi$.

$P(C^c \cup C) = P(C^c) + P(C)$

$\parallel$

$P(\Omega)$

$\parallel$

1

移項

$\Rightarrow P(C) = 1 - P(C^c)$.

Theorem 2

$P(\phi) = 0$

By theorem 1.

$P(\phi) = 1 - P(\Omega) = 0.$

$\phi^c = \Omega$

$\Omega^c = \phi$

Theorem 3

If C_1 and $C_2 \subseteq \Omega$, 且 $C_1 \subset C_2$.

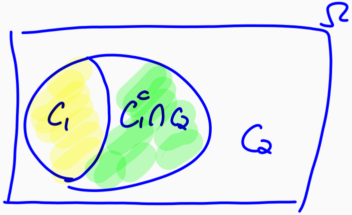
则 $P(C_1) \le P(C_2)$.

pf: $C_2 = C_1 \cup (C_1^c \cap C_2)$

$\because C_1 \cap (C_1^c \cap C_2) = \phi$

$P(C_2) = P(C_1) + \underbrace{P(C_1^c \cap C_2)}_{\ge 0}$.

$\Rightarrow P(C_2) \ge P(C_1)$.



Theorem 4

For $C \subset \Omega$, $0 \leq P(C) \leq 1$.

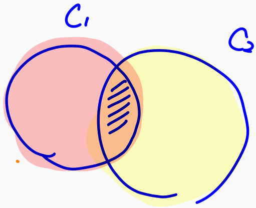
pf: $\phi \subset C \subset \Omega$

$0 = P(\phi) \leq P(C) \leq P(\Omega) = 1$.

Theorem 5

If C_1 and $C_2 \subseteq \Omega$. 則

$$P(C_1 \cup C_2) = P(C_1) + P(C_2) - P(C_1 \cap C_2).$$



pf: $^{*}\ C_1 \cup C_2 = C_1 \cup (C_1^c \cap C_2)$

and $^{**}\ C_2 = (C_1 \cap C_2) \cup (C_1^c \cap C_2)$

$$P(C_1 \cup C_2) = P(C_1) + P(\cancel{C_1^c \cap C_2}) \quad \text{-- } ①$$

$$P(C_2) = P(C_1 \cap C_2) + P(\cancel{C_1^c \cap C_2}) \quad \text{— } ②$$

$① - ② \Rightarrow \quad P(C_1 \cup C_2) - P(C_2) = P(C_1) - P(C_1 \cap C_2)$

$$P(C_1 \cup C_2) = P(C_1) + P(2) - P(C_1 \cap C_2).$$

Examples 1 and 2

Ex1 : 두 dice :

If $C_1 = \{(1,1), (2,1), (3,1), (4,1), (5,1)\}$

$C_2 = \{(1,2), (2,2), (3,2)\}$

$P(C_1) = 5/36$

$P(C_2) = 3/36$

$P(C_1 \cup C_2) = P(C_1) + P(C_2) = 8/36 \quad (\because C_1 \cap C_2 = \phi)$

$P(C_1 \cap C_2) = 0.$

Conditional Probability and Independent

Definition of Conditional Probability

給定已發生.

- In probability theory, conditional probability is a measure of the probability of an event occurring, given that another event has already occurred. It is denoted by $P(A|B)$, which represents the probability of event A given event B.
- The conditional probability of event A given event B is defined as the ratio of the probability of the intersection of events A and B to the probability of event B:

A|B

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

- Conditional probability allows us to update our probability assessments based on new information or conditions.

Conditional Probability Axioms

In probability theory, conditional probability is governed by three important axioms that define its properties and behavior. These axioms are:

Axiom 1: For any two events A and B, the conditional probability of A given B is always non-negative:

$$P(A|B) \geq 0$$

Axiom 2:$P(S|S) = 1$
Axiom 3:: For any countable sequence of events $A_1, A_2, A_3, \ldots$ that are pairwise mutually exclusive (i.e., $A_i \cap A_j = \emptyset$ for $i \neq j$) and their union covers the entire sample space S. Then

$$P\left(\bigcup_{i=1}^{\infty} A_i | B\right) = \sum_{i=1}^{\infty} P(A_i|B)$$

$P(A_1 \cup A_2 \cup A_3 \cup \cdots \cup A_n \mid B)$

Examples 1,2,5

① 8 張紙片

3張 紅　5張 藍

② 抽紙片 (不返還)

C_1 : 第一抽 "紅"

C_2 : 第二抽 "藍"

$P(C_1) = 3/8.$

$P(C_2 | C_1) = 5/7.$

7

2 紅　5 藍

$$P(C_2|C_1) = \frac{P(C_2 \cap C_1)}{P(C_1)}$$

$$\Rightarrow P(C_2 \cap C_1) = P(C_2|C_1)P(C_1) = 5/7 \cdot 3/8.$$

Independent "set version"- Example 7

任一個現況事件之機率 "A"

$$P(A) = P(A|H)P(H)$$

if $H \subset A$

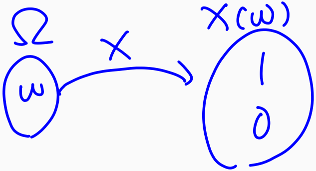
$$P(A \cap H) = P(A|H)P(H)$$

Random variables

Definition of Random Variable

- In probability theory and statistics, a random variable is a function that maps outcomes of a random experiment to numerical values. It assigns a real number to each possible outcome of the experiment.
- Formally, let X be a random variable. For each outcome w in the sample space S, the random variable X assigns a value $X(w)$ from the set of real numbers.

正面
反面



$P(X=1)$
$=$
$P(\text{正面})$

Definition of Random Variable

X : random variable , $x = X(\omega)$

X 銅板 at

x 可以為 0 反面 , 1 正面

Random variables can be categorized into two types:

$P(X=1) = \frac{1}{2}$, $P(X=0) = \frac{1}{2}$, X

- **Discrete Random Variable:** A random variable that can take on a countable set of distinct values. It is characterized by its probability mass function (PMF), which assigns probabilities to each possible value.
- **Continuous Random Variable:** A random variable that can take on any value within a specified range or interval. It is characterized by its probability density function (PDF), which gives the probability density at each point in the range.

Discrete Type Random Variables

Discrete type random variables are a category of random variables that can take on a countable set of distinct values. They are often associated with experiments or phenomena that have a finite or countably infinite number of possible outcomes.

- Examples of discrete type random variables include:
    - The number of heads obtained in a series of coin tosses.
    - The number of students in a class who pass an exam.
    - The outcome of rolling a fair six-sided d~~i~~e.

      dice

Discrete Type Random Variables

To characterize a discrete type random variable, we use its probability mass function (PMF). The PMF assigns probabilities to each possible value that the random variable can assume. It satisfies the following properties:

$P(X=1)$

- Non-negativity: The probability assigned to each value is non-negative.
- Summation: The sum of probabilities over all possible values is equal to 1.

$P(X=0)=\frac{1}{2}$ $\quad P(X=1)=\frac{1}{2}$

$0 \quad 1$

$P(X=0)+P(X=1)=1$

$1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6$

$\frac{1}{6} \quad \frac{1}{6} \quad \frac{1}{6} \quad \frac{1}{6} \quad \frac{1}{6} \quad \frac{1}{6}$

Probability Mass Function (PMF)

The PMF gives the probability of discrete random variables taking specific values.

Definition
For a discrete random variable X, the PMF is defined as:

$$P(X = x) = P(X \text{ takes the value } x) \geqslant 0 .$$

- $P(X = x) \geq 0$ for all x in the range of X.
- The sum of probabilities over all possible values is 1:

$$\sum_{\text{all } x} \underset{\sim\sim\sim\sim\sim}{P(X = x) = 1}$$

Flipping coin

Expectation of Discrete Random Variable

demo.R

The expectation, or expected value, of a discrete random variable is a measure of the average value or center of its probability distribution.

- Let X be a discrete random variable with possible values $x_1, x_2, \ldots, x_n$ and corresponding probabilities $p_1, p_2, \ldots, p_n$. The expectation of X, denoted as $\mathbb{E}[X]$ or μ, is calculated as:

$$\mathbb{E}[X] = x_1 \cdot p_1 + x_2 \cdot p_2 + \ldots + x_n \cdot p_n = \sum_{i=1}^{n} x_i \cdot p_i$$

where x_i is a possible value of X and p_i is the probability associated with x_i.

$$\underbrace{x_1}_{0} + \underbrace{x_2}_{0 \text{ or } 1} + \cdots + x_{1000} \Big/ n=1000 \approx 0.5$$

$$\frac{0\times(\text{發生了幾次}) + 1\times(\text{發生了幾次})}{n}$$

$$0\times\frac{\#0}{n} + 1\,\frac{\#1}{n} \approx 0\cdot\frac{1}{2} + 1\,\frac{1}{2}$$

Variation of Discrete Random Variable

$1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \frac{1}{6} + 4 \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} \cong 3.5.$

$0 \cdot \frac{1}{3} + 1 \frac{2}{3} = \frac{2}{3}$

Let X be a discrete random variable with expected value $\mathbb{E}[X]$. The variance of X, denoted as $\text{Var}(X)$ or σ^2, is calculated as:

$$\text{Var}(X) = \sum_{i=1}^{n} (x_i - \mathbb{E}[X])^2 \cdot p_i$$

$2 \quad 3.5 \quad \frac{1}{6}$

where x_i is a possible value of X, p_i is the probability associated with x_i, and $\mathbb{E}[X]$ is the expected value of X.

$P(X = 500) = \frac{1}{2}$

$P(X = 1000) = \frac{7}{20}$

$P(X = 3000) = \frac{3}{20}$

$E(X) = 500 \cdot \frac{1}{2} + 1000 \cdot \frac{7}{20} + 3000 \cdot \frac{3}{20}$

$=$

Variation of Discrete Random Variable

- The variance measures the average squared deviation of the values of the random variable from its expected value. It provides a measure of the spread or dispersion around the mean.
- To obtain the standard deviation, denoted as σ, we take the square root of the variance: $\sigma = \sqrt{\mathrm{Var}(X)}$.
- The concept of variation is essential in probability theory and statistics. It helps in understanding the uncertainty or variability associated with the random variable and plays a crucial role in hypothesis testing, confidence intervals, and other statistical analyses.

Continuous Random Variable

A continuous random variable is a type of random variable that can take on any value within a specified range or interval. It is often associated with measurements or observations that can take on an infinite number of possible values.

Examples of continuous random variables include:

- The height of individuals in a population.
- The time it takes for a computer program to execute.
- The amount of rainfall in a given area.

Continuous Random Variable

To characterize a continuous random variable, we use its probability density function (PDF). The PDF gives the probability density at each point in the range. Unlike the probability mass function (PMF) of a discrete random variable, the PDF does not assign probabilities to specific values but describes the probability density at each value.

The properties of a continuous random variable include:

- Non-negativity: The PDF is non-negative for all values.
- Integrates to 1: The integral of the PDF over the entire range is equal to 1.

Probability Density Function (PDF)

The PDF describes the probability distribution of continuous random variables.

Definition
For a continuous random variable X, the PDF is defined as:

$$f_X(x) = \lim_{\Delta x \to 0} \frac{P(x \leq X \leq x + \Delta x)}{\Delta x}$$

- The PDF does not give the exact probability at a specific point.
- Instead, it provides the relative likelihood of the variable being in a certain range.
- The area under the PDF curve over a range gives the probability within that range.

Expectation of Continuous Random Variable

Let X be a continuous random variable with probability density function (PDF) $f(x)$. The expectation of X, denoted as $\mathbb{E}[X]$ or μ, is calculated as:

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x \cdot f(x)\, dx$$

where x is a value of X and $f(x)$ is the probability density at x.

The expectation of a continuous random variable represents the weighted average of its possible values, where the weights are given by the PDF. It provides a summary measure of the central tendency of the random variable's distribution.

Variation of Continuous Random Variable

The variation, or variance, of a continuous random variable measures the spread or dispersion of its probability distribution.

Let X be a continuous random variable with expected value $\mathbb{E}[X]$. The variance of X, denoted as $\mathsf{Var}(X)$ or σ^2, is calculated as:

$$\mathsf{Var}(X) = \int_{-\infty}^{\infty} (x - \mathbb{E}[X])^2 \cdot f(x)\, dx$$

where $f(x)$ is the probability density function (PDF) of X and $\mathbb{E}[X]$ is the expected value of X.