THE CHINESE UNIVERSITY OF HONG KONG
Department of Mathematics
Exercises on Probability

1 Permutations and Combinations

Two core counting techniques are permutations and combinations. A **permutation** counts the number of distinct arrangements where the *order of selection matters*.

1.1 Illustrative Example: Ordered Selection of 3 Cards from 5

We have five distinct cards labeled $1, 2, 3, 4, 5$. We select three cards, and the sequence/order of selection is important.

$$5 \times 4 \times 3 = 60$$

①	②	③
(1 of 5 cards)	(1 of 4 remaining cards)	(1 of 3 remaining cards)

1.2 Step-by-Step Logical Reasoning:

We place the three selected cards into ordered positions ①, ②, ③:

1. Position ①: There are 5 available numbers: $1, 2, 3, 4, 5$ (5 total choices).
2. Position ②: One card is already used for position ①, leaving 4 unused cards (4 total choices).
3. Position ③: Two cards are already used for positions ① and ②, leaving only 3 unused cards (3 total choices).

By the multiplication principle of counting, multiply the number of choices for each position:

$$\text{Total ordered outcomes} = 5 \times 4 \times 3 = \mathbf{60}$$

1.3 General Permutation Definition and Formula

When selecting k distinct items from a set of n distinct items with order preserved, this count is called a permutation, denoted $\boldsymbol{{}_nP_k}$.

Base product form (finite descending product):

$${}_nP_k = n\,(n-1)\,(n-2)\cdots(n-k+1), \quad \text{for } k \leq n$$

1.4 Factorial Special Case ($k = n$):

If we select all n items ($k = n$):

$${}_nP_n = n\,(n-1)\,(n-2)\cdots 2 \cdot 1$$

This product is defined as the **factorial** of n, written $n!$ (read "n factorial").

1.5 Factorial Equivalent Derivation (Step-by-Step Algebra):

Multiply numerator and denominator by the trailing product $(n-k)(n-k-1)\cdots 2\cdot 1 = (n-k)!$:

$$\begin{aligned} {}_nP_k &= n\,(n-1)\cdots(n-k+1) \\ &= \frac{\big[n\,(n-1)\cdots(n-k+1)\big]\cdot\big[(n-k)\cdots 2\cdot 1\big]}{(n-k)\cdots 2\cdot 1} \\ &= \frac{n!}{(n-k)!} \end{aligned}$$

1.6 Numerical Verification for the Card Example ($n = 5$, $k = 3$)

$$\begin{aligned} {}_5P_3 &= \frac{5!}{(5-3)!} = \frac{5!}{2!} \\ 5! &= 5\times 4\times 3\times 2\times 1 = 120 \\ 2! &= 2\times 1 = 2 \\ {}_5P_3 &= \frac{120}{2} = \mathbf{60} \end{aligned}$$

Matches the direct product calculation result, confirming correctness.

2 Combination Definition and Formula

A **combination** counts the number of ways to select items where the *order of selection does not matter*.

2.1 Example 1: Select 3 unordered cards from 5 cards labeled $1, 2, 3, 4, 5$

Direct computation:

$$\frac{5\times 4\times 3}{3!} = 10$$

1: Explain order equivalence Suppose the selected values are $1, 2, 3$.

- If order matters (permutation), there are $3! = 3\times 2\times 1 = 6$ distinct ordered arrangements:

$$(1,2,3),\ (1,3,2),\ (2,1,3),\ (2,3,1),\ (3,1,2),\ (3,2,1)$$

- For combinations, all 6 ordered groupings count as **one identical unordered group**. To eliminate redundant permutations of the k selected items, we divide the total permutation count by $k!$.

2: Rewrite numerator using factorials The numerator $5\times 4\times 3$ is the permutation ${}_5P_3 = \frac{5!}{(5-3)!} = \frac{5!}{2!}$. Expand factorial definitions:

$$5! = 5\times 4\times 3\times 2\times 1,\quad 2! = 2\times 1$$

Canceling 2! from numerator and denominator leaves $5 \times 4 \times 3$. Substitute back into the combination expression:

$$\frac{5 \times 4 \times 3}{3!} = \frac{\frac{5!}{2!}}{3!} = \frac{5!}{3! \cdot 2!}$$

3: Verify numerical value step-by-step

$$3! = 3 \times 2 \times 1 = 6$$
$$5 \times 4 \times 3 = 60$$
$$\frac{60}{6} = \mathbf{10}$$

General Combination Formula When choosing k items from n distinct items with no regard to order, the combination count is denoted ${}_nC_k$ (also written binomial coefficient $\binom{n}{k}$).

$${}_nC_k = \frac{{}_nP_k}{k!} = \frac{n!}{(n-k)!\,k!} = \binom{n}{k}, \quad \text{for } k \le n$$

2.2 Example 2: Choose 5 ties out of 500 ties

Problem: Find the number of unordered ways to pick 5 ties from 500 total ties. By combination formula, $n = 500,\ k = 5$:

$$\binom{500}{5} = \frac{500!}{5! \cdot (500-5)!} = \frac{500!}{5! \cdot 495!}$$

1: Simplify the fraction for manual computation (cancel 495!):

$$\binom{500}{5} = \frac{500 \times 499 \times 498 \times 497 \times 496 \times 495!}{5! \times 495!}$$
$$= \frac{500 \times 499 \times 498 \times 497 \times 496}{5 \times 4 \times 3 \times 2 \times 1}$$

2: Arithmetic calculation: Numerator product:

$$500 \times 499 = 249500,$$
$$249500 \times 498 = 124251000,$$
$$124251000 \times 497 = 61752747000,$$
$$61752747000 \times 496 = 30629362512000$$

Denominator $5! = 120$

$$\binom{500}{5} = \frac{30629362512000}{120} = \mathbf{255244687600}$$

3 Permutations and Combinations with Repetition

Up to this point, we have covered counting problems with no repeated selections. When repetition of items is permitted, we use separate formulas for permutations and combinations with repetition.

1) Permutations with Repetition (Repeated Permutation): Symbol ${}_n\Pi_k$

${}_n\Pi_k$ denotes ordered selection of k items from n distinct types, where items may be reused (repetition allowed).

$$\boldsymbol{{}_n\Pi_k = n^k}$$

2) Combinations with Repetition (Repeated Combination): Symbol ${}_nH_k$

${}_nH_k$ denotes unordered selection of k items from n distinct types, where items may be reused (repetition allowed). The formula transforms to a standard binomial coefficient:

$$\boldsymbol{{}_nH_k = \binom{n+k-1}{k}}$$

3.1 Example 3

Select three numbers from $\{1, 2, 3, 4, 5\}$, arrange them sequentially to form three-digit natural numbers (repetition permitted). Find how many such three-digit numbers are multiples of 5.

Step-by-Step Solution Reasoning: A natural number is a multiple of 5 if its units digit equals 0 or 5. Our available digits are only $1, 2, 3, 4, 5$, so the units digit **must be fixed as** 5.

- Position 3 (units place): Only 1 valid choice (5)

- Positions 1 (hundreds) and 2 (tens): We select 2 digits from $\{1, 2, 3, 4, 5\}$ with repetition allowed ($n = 5,\ k = 2$) This is a permutation-with-repetition problem for the first two digits: ${}_n\Pi_k = n^k$

$$\begin{aligned} n = 5, \quad k = 2 \\ {}_5\Pi_2 = 5^2 \\ 5^2 = 5 \times 5 = \mathbf{25} \end{aligned}$$

Total valid three-digit multiples of 5: **25**

3.2 Example 4

Four people cast anonymous votes for one of three candidates: A, B, C. Find the total number of distinct vote outcome distributions.

Step-by-Step Solution Reasoning: Votes are anonymous, so only the count of votes for each candidate matters (order of voters does not matter). This is a combination-with-repetition scenario:

- $n = 3$ distinct candidate types (A,B,C)

- $k = 4$ total votes to distribute Formula: ${}_nH_k = \binom{n+k-1}{k}$

$$_3H_4 = \binom{3+4-1}{4} = \binom{6}{4}$$

Binomial coefficient simplification: $\binom{6}{4} = \binom{6}{2}$

$$\binom{6}{2} = \frac{6!}{2!\,(6-2)!} = \frac{6!}{2!\,4!}$$

$$6! = 6 \times 5 \times 4!$$

$$\binom{6}{2} = \frac{6 \times 5 \times 4!}{(2 \times 1) \times 4!} = \frac{30}{2} = \mathbf{15}$$

Total unique anonymous vote distributions: **15**

4 Probability

The **probability** that a specific event occurs takes a value in the interval $[0, 1]$. For example, flipping a fair coin has a probability of $\frac{1}{2}$ for landing on heads.

When flipping a coin, only two outcomes exist: heads or tails. There are 2 total possible outcomes, and exactly 1 favorable outcome for heads.

- A probability value of 0 means the event is impossible (can never happen).

- A probability value of 1 means the event is certain (must happen).

To compute probability mathematically, we first define the full set of all possible outcomes.

- Coin flip sample space: $S = \{\text{head}, \text{tail}\}$

- Standard six-sided die sample space: $S = \{1, 2, 3, 4, 5, 6\}$

This complete collection of all possible outcomes is called the **sample space**, denoted S.

4.1 Formal Classical Probability Definition

Let:

- $n(S)$ = total number of distinct possible outcomes in sample space S

- $n(A)$ = number of favorable outcomes for a specific event A

The classical probability of event A, written $P(A)$, is defined as:

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

4.2 Geometric Probability Definition

For geometric problems where S represents a total region/area, and $A \subset S$ (event region lies entirely within sample space region):

$$P(A) = \frac{\text{Area of region } \boldsymbol{A}}{\text{Area of region } \boldsymbol{S}}$$

4.3 Statistical Probability Law of Large Numbers

In practical experimental trials: Suppose an identical trial is repeated n total times, and event A is observed to occur k of those times. The **relative frequency** (statistical probability) of event A is:

$$\frac{k}{n}$$

As the number of trials n grows infinitely large, the relative frequency converges to a fixed constant value $\mathcal{P}$, called the **mathematical probability** $P(A)$. This convergence property is the **Law of Large Numbers**, expressed formally with a limit:

$$\boldsymbol{\lim_{n\to\infty} \frac{k}{n} = P(A)}$$

4.4 Fundamental Probability Axioms Properties

Let A be any event within sample space S:

1. Bounded range: $\boldsymbol{0 \leq P(A) \leq 1}$ for any valid event $A \subseteq S$
2. Certainty of full sample space: $\boldsymbol{P(S) = 1}$
3. Impossible empty event: $\boldsymbol{P(\emptyset) = 0}$ ($\emptyset$ = empty set, no outcomes)
4. Additivity for mutually exclusive events: If A and B cannot happen simultaneously (disjoint, $A \cap B = \emptyset$):

$$\boldsymbol{P(A \cup B) = P(A) + P(B)}$$

5. Complement rule: Let A^c (also written $\overline{A}$) denote the complement event (all outcomes where A *does not* occur):

$$\boldsymbol{P(A^c) = 1 - P(A)}$$

Step-by-Step Example Calculation (Coin Flip Heads)

$$\begin{aligned} S &= \{\text{head}, \text{tail}\} \implies n(S) = 2 \\ A &= \{\text{head}\} \implies n(A) = 1 \\ P(\text{head}) &= \frac{n(A)}{n(S)} = \frac{1}{2} = \mathbf{0.5} \end{aligned}$$

Step-by-Step Example Calculation (Die Roll, Event = Roll a 3)

$$\begin{aligned} S &= \{1, 2, 3, 4, 5, 6\} \implies n(S) = 6 \\ A &= \{3\} \implies n(A) = 1 \\ P(\text{roll } 3) &= \frac{1}{6} \approx \mathbf{0.1667} \end{aligned}$$

4.5 Example 5: Probability of Drawing Two Same-Colored Balls

Problem Statement: A pocket contains 3 black balls, 2 white balls, and 1 red ball (total of $3 + 2 + 1 = 6$ balls). Two balls are drawn simultaneously at random. Calculate the probability that the two selected balls share the same color.

1: Compute total number of unordered ways to pick 2 balls from 6 Drawing two balls at once is an unordered selection, so we use combinations $\binom{n}{k} = {}_nC_k$. Total sample space outcomes:

$$n(S) = \binom{6}{2}$$

Expand and calculate step-by-step:

$$\binom{6}{2} = \frac{6!}{2!\,(6-2)!} = \frac{6!}{2! \cdot 4!}$$
$$6! = 6 \times 5 \times 4!$$
$$\binom{6}{2} = \frac{6 \times 5 \times 4!}{(2 \times 1) \times 4!} = \frac{30}{2} = \mathbf{15}$$

2: Count favorable outcomes (two balls of identical color) Only black and white colors have at least two balls; there is only 1 red ball, so

1. Ways to pick 2 black balls from 3 black balls: $\binom{3}{2}$

$$\binom{3}{2} = \frac{3!}{2! \cdot 1!} = \frac{3 \times 2!}{2! \times 1} = \mathbf{3}$$

2. Ways to pick 2 white balls from 2 white balls: $\binom{2}{2}$

$$\binom{2}{2} = \frac{2!}{2! \cdot 0!} = \frac{2!}{2! \times 1} = \mathbf{1} \quad \text{(by definition } 0! = 1\text{)}$$

Total favorable outcomes $n(A)$ = black pairs + white pairs:

$$n(A) = \binom{3}{2} + \binom{2}{2} = 3 + 1 = \mathbf{4}$$

3: Classical Probability Calculation

$$P(\text{same color}) = \frac{n(A)}{n(S)} = \frac{4}{15}$$

Example 6: Hypergeometric Probability for Defective Products

Problem Setup:

Total products: 1000

Defective products: 3

Normal (non-defective) products: $1000 - 3 = 997$ We randomly select 10 products without replacement.

Compute two probabilities:

1. Probability that none of the 10 selected items are defective.
2. Probability that at least one selected item is defective.

1: Total number of ways to pick 10 items from 1000 This is an unordered combination selection:

$$n(S) = \binom{1000}{10}$$

Part (1): Zero defective items in the 10 selected To have zero defective units, we select all 10 units from the 997 normal products, and select 0 defective units from the 3 defective ones. Number of favorable outcomes for event A_0 (0 defectives):

$$n(A_0) = \binom{997}{10} \times \binom{3}{0}$$

By classical probability formula:

$$P(A_0) = \frac{\binom{997}{10}\binom{3}{0}}{\binom{1000}{10}}$$

Recall $\binom{m}{0} = 1$ for any integer $m \geq 0$, so $\binom{3}{0} = 1$, simplifying:

$$P(A_0) = \frac{\binom{997}{10}}{\binom{1000}{10}}$$

2. Numerical Simplification for $P(A_0)$ Expand binomial ratios to avoid huge factorial values:

$$\frac{\binom{997}{10}}{\binom{1000}{10}} = \frac{\frac{997!}{10! \cdot 987!}}{\frac{1000!}{10! \cdot 990!}} = \frac{997! \cdot 990!}{987! \cdot 1000!} = \frac{990 \times 989 \times 988}{1000 \times 999 \times 998}$$

Compute step arithmetic:

$$990 \times 989 \times 988 = 967360680, \quad 1000 \times 999 \times 998 = 997002000$$

$$P(0 \text{ defective}) = \frac{967360680}{997002000} \approx \mathbf{0.9702695}$$

Part (2): At least one defective item selected Let event B = at least one defective unit. The complement of B is exactly event A_0 (zero defective units). By complement probability rule:

$$P(B) = 1 - P(A_0)$$

Formula written with binomial coefficients:

$$P(\text{at least 1 defective}) = 1 - \frac{\binom{997}{10}\binom{3}{0}}{\binom{1000}{10}}$$

Numerical Calculation for Part (2)

$$P(\text{at least 1 defective}) = 1 - 0.9702695 = \mathbf{0.0297305}$$

5 Conditional Probability

Conditional probability is a foundational concept in data analysis. It denotes the probability that event B occurs, given the prior condition that event A has already occurred. This conditional probability is written $\boldsymbol{P(B \mid A)}$.

Formal Definition

$$\boldsymbol{P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{P(B \cap A)}{P(A)}}, \quad \text{valid only when } P(A) > 0$$

The notation $P_A(B)$ shown in the diagram is an equivalent shorthand for $P(B \mid A)$.

Product Rule Derived from Definition

Rearranging the conditional probability formula yields the product rule for joint probability:

$$\boldsymbol{P(A \cap B) = P(A)\, P(B \mid A) = P(B)\, P(A \mid B) = P(B \cap A)}$$

Generalized Product Rule for n Events $A_1, A_2, \ldots, A_n$

$$\begin{aligned} & P(A_1)\, P(A_2 \mid A_1)\, P(A_3 \mid A_1 \cap A_2) \cdots P(A_n \mid A_1 \cap A_2 \cap \cdots \cap A_{n-1}) \\ &= P(A_1) \cdot \frac{P(A_1 \cap A_2)}{P(A_1)} \cdot \frac{P(A_1 \cap A_2 \cap A_3)}{P(A_1 \cap A_2)} \cdots \frac{P(A_1 \cap A_2 \cap \cdots \cap A_n)}{P(A_1 \cap A_2 \cap \cdots \cap A_{n-1})} \\ &= \boldsymbol{P(A_1 \cap A_2 \cap \cdots \cap A_n)} \end{aligned}$$

Complement Partition Identity for Event A

For any event A, the disjoint partition holds:

$$A = (A \cap B) \cup (A \cap B^c), \quad (A \cap B) \cap (A \cap B^c) = \emptyset$$

By additivity of probability for disjoint sets:

$$P(A) = P(A \cap B) + P(A \cap B^c)$$

Rearranged:

$$\boldsymbol{P(A \cap B) = P(A) - P(A \cap B^c)}$$

Example 7

Given values:

$$P(A) = \frac{21}{25}, \quad P(A \cap B^c) = \frac{1}{5}$$

Goal: Compute $P(B \mid A)$

1: Substitute partition identity into conditional formula

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A) - P(A \cap B^c)}{P(A)}$$

2: Substitute numerical fractions step-by-step First convert $\frac{1}{5}$ to denominator 25 for uniform subtraction:

$$\frac{1}{5} = \frac{1 \times 5}{5 \times 5} = \frac{5}{25}$$

Plug values into numerator:

$$P(A) - P(A \cap B^c) = \frac{21}{25} - \frac{5}{25} = \frac{21 - 5}{25} = \frac{16}{25}$$

3: Divide numerator by $P(A)$

$$P(B \mid A) = \frac{\frac{16}{25}}{\frac{21}{25}}$$

The common denominator 25 cancels out:

$$P(B \mid A) = \mathbf{\frac{16}{21}}$$

6 Bayes' Theorem

Bayes' theorem calculates the probability of an event by incorporating prior knowledge of related conditional conditions. This tool is vital for mathematical decision-making under uncertainty, and it is widely used to value invisible, intangible assets such as information.

Key Terminology Definitions

- **Prior Probability**: $P(A)$ — probability of event A calculated before observing new data or evidence.
- **Posterior Probability**: $P(A \mid B)$ — revised, updated probability of event A after new evidence/event B is observed to occur. In $P(A \mid B)$, event B is the observed condition, and $P(A \mid B)$ updates our belief about A after seeing B.

Bayes' Theorem derives posterior probabilities by combining known prior probabilities and conditional likelihoods from observed events.

Partition of Sample Space and Law of Total Probability

Let sample space S be partitioned into disjoint events $A_1, A_2, \dots, A_n$:

$$A_i \cap A_j = \emptyset \quad (i \neq j), \qquad S = A_1 \cup A_2 \cup \cdots \cup A_n$$

For any arbitrary event $B \subseteq S$:

$$B = S \cap B = \big(A_1 \cup A_2 \cup \cdots \cup A_n\big) \cap B = (A_1 \cap B) \cup (A_2 \cap B) \cup \cdots \cup (A_n \cap B)$$

All sets $A_i \cap B$ are mutually exclusive, so by finite additivity of probability:

$$P(B) = P(A_1 \cap B) + P(A_2 \cap B) + \cdots + P(A_n \cap B)$$

Apply the probability product rule $P(A_i \cap B) = P(A_i)\, P(B \mid A_i)$ to get the **Law of Total Probability**:

$$\boldsymbol{P(B) = \sum_{i=1}^{n} P(A_i)\, P(B \mid A_i)} = P(A_1)P(B \mid A_1) + P(A_2)P(B \mid A_2) + \cdots + P(A_n)P(B \mid A_n)$$

Bayes' Theorem Formula

From conditional probability definition:

$$P(A_j \mid B) = \frac{P(A_j \cap B)}{P(B)}$$

Substitute product rule $P(A_j \cap B) = P(A_j)P(B \mid A_j)$ and total probability for $P(B)$:

$$\boldsymbol{P(A_j \mid B) = \frac{P(A_j)\, P(B \mid A_j)}{\sum_{i=1}^{n} P(A_i)\, P(B \mid A_i)}}$$

- $P(A_j)$: prior probability of partition event A_j

- $P(A_j \mid B)$: posterior probability of A_j given observed event B

6.1 Example 8

Problem Setup:

Three factory machines A, B, C produce 50%, 30%, 20% of total output respectively. Defect rates: $P(\text{Defect} \mid A) = 0.04$, $P(\text{Defect} \mid B) = 0.03$, $P(\text{Defect} \mid C) = 0.02$.

1. Find the total probability that a randomly selected product is defective ($P(X)$, where $X =$ defective item).

2. Find the posterior probability that a defective product was made by machine C ($P(C \mid X)$).

1: Define Prior Probabilities

$$P(A) = 0.5, \quad P(B) = 0.3, \quad P(C) = 0.2$$

Conditional defect likelihoods:

$$P(X \mid A) = 0.04, \quad P(X \mid B) = 0.03, \quad P(X \mid C) = 0.02$$

2: Law of Total Probability for $P(X)$

$$\begin{aligned} P(X) &= P(A)P(X \mid A) + P(B)P(X \mid B) + P(C)P(X \mid C) \\ &= (0.5 \times 0.04) + (0.3 \times 0.03) + (0.2 \times 0.02) \\ &= 0.02 + 0.009 + 0.004 \\ &= \mathbf{0.033} \end{aligned}$$

3: Bayes' Theorem for Posterior $P(C \mid X)$

$$\begin{aligned} P(C \mid X) &= \frac{P(C)\,P(X \mid C)}{P(A)P(X \mid A) + P(B)P(X \mid B) + P(C)P(X \mid C)} \\ &= \frac{0.2 \times 0.02}{0.033} \\ &= \frac{0.004}{0.033} = \mathbf{\frac{4}{33}} \approx 0.1212 \end{aligned}$$

7 Random Variables

Consider the experiment of flipping two fair coins. The sample space of all distinct outcomes is:

$$S = \{\,(\text{Head}, \text{Head}),\ (\text{Head}, \text{Tail}),\ (\text{Tail}, \text{Head}),\ (\text{Tail}, \text{Tail})\,\}$$

For a fair coin flip, every outcome in this sample space is equally likely, so the probability of each single outcome equals $\mathbf{\frac{1}{4}}$.

Define X as the **count of tails observed in the two coin tosses**. The random variable X maps each sample outcome to a real numerical value:

- Outcome $(\text{Head}, \text{Head}) \mapsto X = 0$ tails
- Outcomes $(\text{Head}, \text{Tail}),\ (\text{Tail}, \text{Head}) \mapsto X = 1$ tail
- Outcome $(\text{Tail}, \text{Tail}) \mapsto X = 2$ tails

This mapping function X is defined as a **random variable**.

Definition of a Random Variable

A random variable behaves analogously to a variable in computer programming, except its realized value is determined by probabilistic chance rather than fixed assignment. Formally, a random variable is a function that maps every outcome (sample point) from the sample space S to a real number on the real line $\mathbb{R}$.

Using random variables converts qualitative probabilistic events into quantitative numerical values, which simplifies calculation, comparison, and statistical analysis of random experiments. Notation convention:

- Uppercase letters (e.g., X, Y, Z): denote the random variable function itself
- Lowercase letters (e.g., x, y, z): denote specific numerical values the random variable may take

Step-by-Step Probability Mass Calculation for X (Number of Tails in 2 Coin Flips)

Total equally-likely sample outcomes: $n(S) = 4$, each with probability $\frac{1}{4}$.

1. Value $x = 0$ (zero tails): Only 1 favorable outcome (H,H)

$$P(X = 0) = \frac{\text{Number of outcomes with } X = 0}{n(S)} = \frac{1}{4} = 0.25$$

2. Value $x = 1$ (one tail): Two favorable outcomes (H,T), (T,H)

$$P(X = 1) = \frac{2}{4} = \frac{1}{2} = 0.50$$

3. Value $x = 2$ (two tails): Only 1 favorable outcome (T,T)

$$P(X = 2) = \frac{1}{4} = 0.25$$

Verification of Total Probability Sum to 1

All mutually exclusive possible values of X cover the full sample space:

$$P(X = 0) + P(X = 1) + P(X = 2) = \frac{1}{4} + \frac{1}{2} + \frac{1}{4} = \frac{1 + 2 + 1}{4} = \frac{4}{4} = \mathbf{1}$$

Sample Space $\rightarrow X \rightarrow$ Real Number Mapping

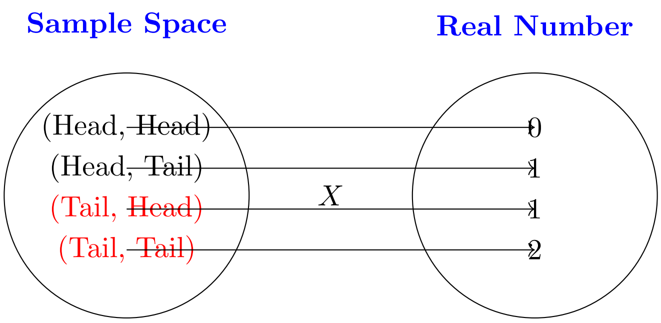

Figure 1: Mapping sample space outcomes to real numbers via random variable X (count of tails)

8 Discrete Probability Distributions

A **discrete random variable** X is a random variable that may only take a *countable set* of distinct real values $x_1, x_2, \ldots, x_n$. A **discrete probability distribution** fully specifies the probability $P(X = x_i)$ for every possible value x_i of the discrete random variable X.

Tabular Format for Discrete Distribution

X	x_1	x_2	$\ldots$	x_n	Sum
Probability	$P(X = x_1)$	$P(X = x_2)$	$\ldots$	$P(X = x_n)$	1

8.1 Example 9: Two Coin Tosses (Count Tails X)

Experiment: Flip two fair coins simultaneously. X = number of tails observed.

Sample space: $S = \{(H,H),(H,T),(T,H),(T,T)\}$, each outcome equally likely with probability $\frac{1}{4}$.

Step-by-step mapping and probability calculation:

1. $X = 0$: Outcome (H,H) only
$$P(X=0) = \frac{1}{4}$$

2. $X = 1$: Outcomes $(H,T),(T,H)$ (2 cases)
$$P(X=1) = \frac{2}{4} = \frac{1}{2}$$

3. $X = 2$: Outcome (T,T) only
$$P(X=2) = \frac{1}{4}$$

Sample Space → X → Probability Mapping

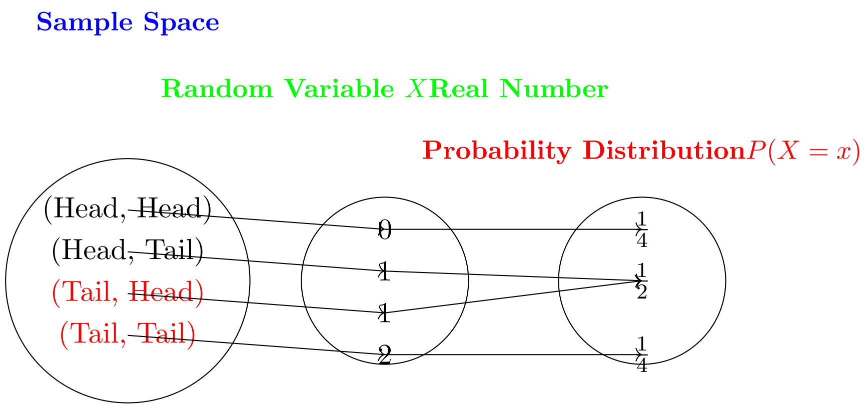

Figure 2: Mapping sample outcomes to X values then to probability masses

Probability Mass Function (PMF) Definition

The function $f(x)$ that assigns the probability value $P(X = x_i)$ to each possible discrete value x_i of X is called the **probability mass function (pmf)** of X:

$$f(x) = \begin{cases} P(X = x_i) & \text{for } x = x_1,\, x_2,\, \ldots,\, x_n \\ 0 & \text{for all other real } x \end{cases}$$

Three Fundamental PMF Properties

1. Bounded non-negativity:
$$\boldsymbol{0 \le f(x_i) \le 1 \quad (i = 1, 2, \ldots, n)}$$

2. Total mass sums to 1 (full sample space coverage):

$$\boldsymbol{f(x_1) + f(x_2) + \cdots + f(x_n) = \sum_{i=1}^{n} f(x_i) = 1}$$

3. Interval probability: For bounds $a \le b$, the probability $P(a \le X \le b)$ equals the sum of pmf values over all x between a and b inclusive:

$$\boldsymbol{P(a \le X \le b) = \sum_{a \le x \le b} f(x)}$$

8.2 Example 10: Verification for Two-Coin Example PMF

PMF values: $f(0) = \frac{1}{4},\ f(1) = \frac{1}{2},\ f(2) = \frac{1}{4}$

1. Bounds check: $0 < \frac{1}{4} < 1,\ 0 < \frac{1}{2} < 1,\ 0 < \frac{1}{4} < 1$ (satisfies property 1)
2. Total sum:

$$f(0) + f(1) + f(2) = \frac{1}{4} + \frac{1}{2} + \frac{1}{4} = \frac{1+2+1}{4} = \frac{4}{4} = \mathbf{1}$$

3. Example interval calculation: $P(0 \le X \le 1)$

$$P(0 \le X \le 1) = f(0) + f(1) = \frac{1}{4} + \frac{1}{2} = \frac{3}{4}$$

9 Continuous Probability Distributions

A **continuous random variable** X is a random variable that can take an uncountably infinite set of real values. For any continuous random variable, the probability that X equals a single exact point value x is always zero:

$$\boldsymbol{P(X = x) = 0}$$

Because point probabilities are zero, the discrete *probability mass function (pmf)* cannot describe a continuous distribution. Instead we introduce the **probability density function (pdf)**, denoted $f(x)$, to model continuous probability behavior.

The pdf $f(x)$ is the continuous analog of the discrete pmf. A function $f(x)$ qualifies as a valid probability density function for continuous random variable X if it satisfies three core properties:

1. Non-negativity for all real inputs:

$$\boldsymbol{f(x) \ge 0 \quad \forall x \in \mathbb{R}}$$

2. Total integral over the entire real line equals 1 (total probability mass):

$$\boldsymbol{\int_{-\infty}^{\infty} f(x)\, dx = 1}$$

3. Interval probability equals the definite integral of $f(x)$ over the interval bounds $[a, b]$. For continuous variables, inclusion or exclusion of endpoints does not change probability (since $P(X = a) = P(X = b) = 0$):

$$\boldsymbol{P(a \leq X \leq b)} = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b) = \boldsymbol{\int_a^b f(x)\,dx}$$

Graphical Interpretation of Interval Probability

Probability $P(a \leq X \leq b)$ corresponds geometrically to the area under the pdf curve $y = f(x)$ between vertical lines $x = a$ and $x = b$.

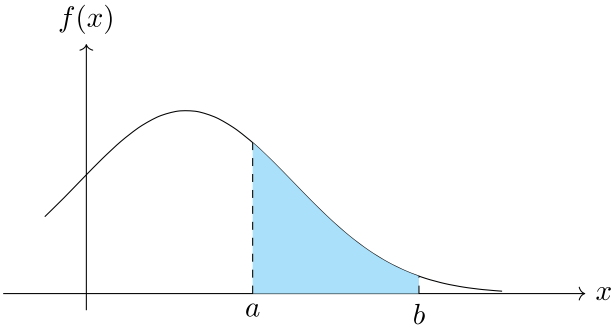

Figure 3: Shaded area $= P(a \leq X \leq b) = \int_a^b f(x)\,dx$

Note: Etymology of "Density" in PDF

We interpret probability as mass, and interval length as a one-dimensional volume:

1. Ratio: $\dfrac{\text{Probability}}{\text{Interval Length}} = \dfrac{\text{Mass}}{\text{Volume}}$
2. The ratio $\frac{\text{mass}}{\text{volume}}$ is the standard definition of physical density.
3. Rearranged relation:

$$\left(\frac{\text{Probability}}{\text{Interval Length}}\right) \times (\text{Interval Length}) = \text{Probability}$$

   Multiplying density (probability per unit length) by interval length yields total probability for that segment. This relationship gives the name *probability density function*.

9.1 Example 12: (Uniform Continuous Distribution Test)

Take uniform pdf over $[0, 2]$:

$$f(x) = \begin{cases} \frac{1}{2} & 0 \leq x \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

1. Property 1 check: $f(x) \geq 0$ for all real x ($\frac{1}{2} > 0$ on support, 0 outside)

2. Property 2 total integral:

$$\int_{-\infty}^{\infty} f(x)dx = \int_{0}^{2} \frac{1}{2}dx = \frac{1}{2}x\Big|_0^2 = \frac{1}{2}(2-0) = \mathbf{1}$$

3. Interval probability example: $P(0.5 \leq X \leq 1.5)$

$$P(0.5 \leq X \leq 1.5) = \int_{0.5}^{1.5} \frac{1}{2}dx = \frac{1}{2}(1.5-0.5) = \mathbf{0.5}$$

Also confirm $P(X=1) = 0$ (single point integral $\int_{1}^{1} \frac{1}{2}dx = 0$)