# Lecture 1: Probability and Counting

Ziyu Shao

School of Information Science and Technology
ShanghaiTech University

September 24, 2024

# Outline

# Outline

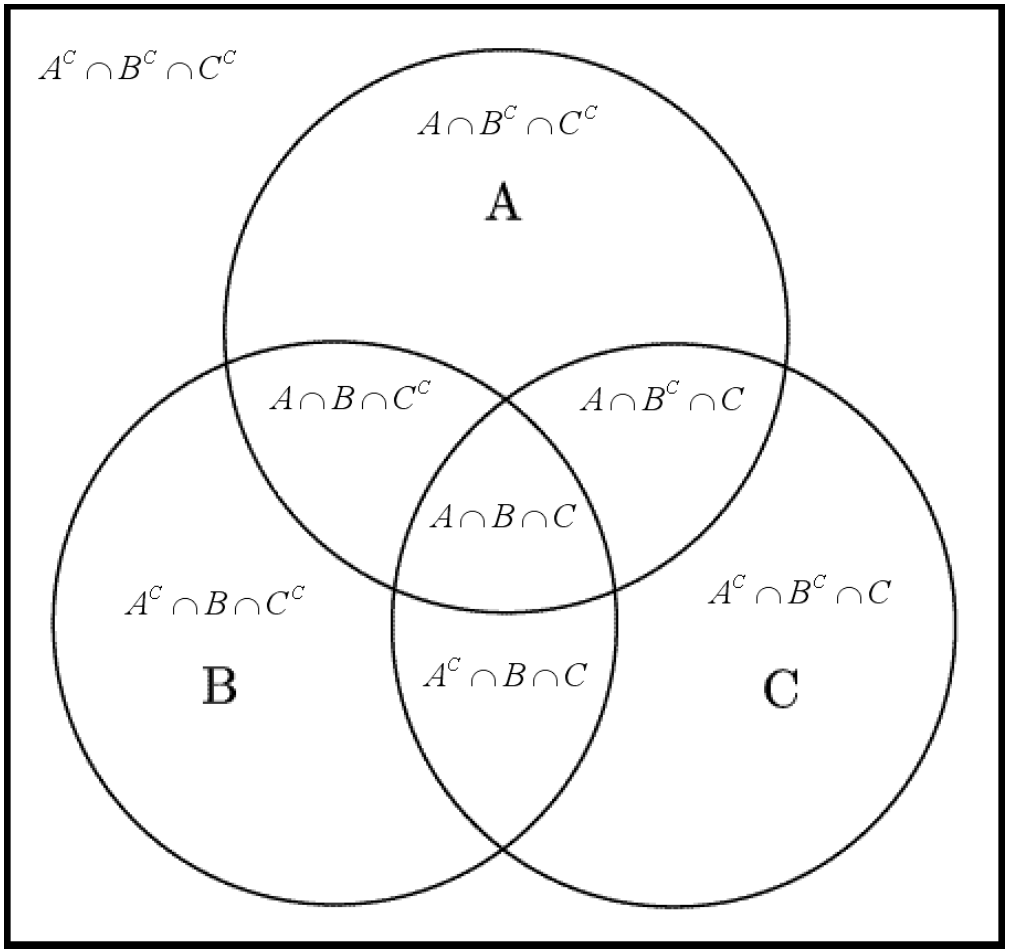
# Set

A *set* is a collection of objects. Given two sets $A, B$, key concepts include

- *empty set*: $\emptyset$
- A is a *subset* of B: $A \subseteq B$
- *union* of $A$ and $B$: $A \cup B$
- *intersection* of $A$ and $B$: $A \cap B$
- *complement* of $A$: $A^c$
- *De Morgan's laws*:

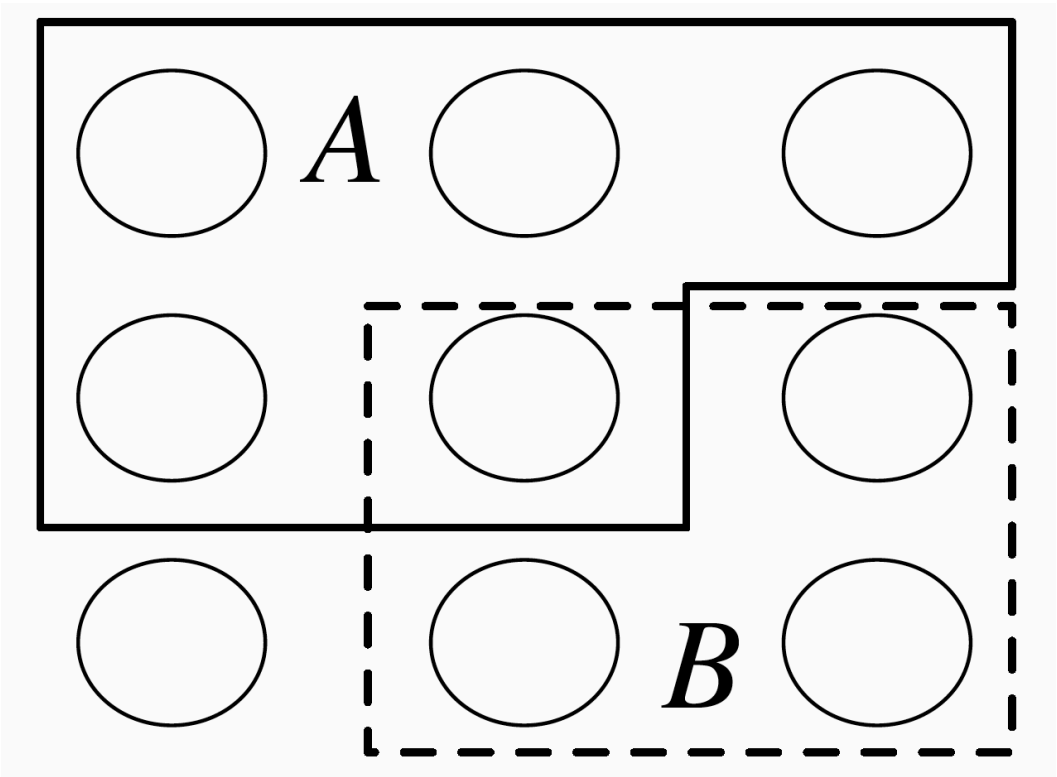
$$(A \cup B)^c = A^c \cap B^c$$
$$(A \cap B)^c = A^c \cup B^c$$

# Venn Diagram



# Sample Space & Event

- The *sample space* $S$ of an experiment: the set of all possible outcomes of the experiment.
- An *event* $A$ is a subset of the sample space $S$.
- *$A$ occurred* if the actual outcome is in $A$.



# Example: Coin flips

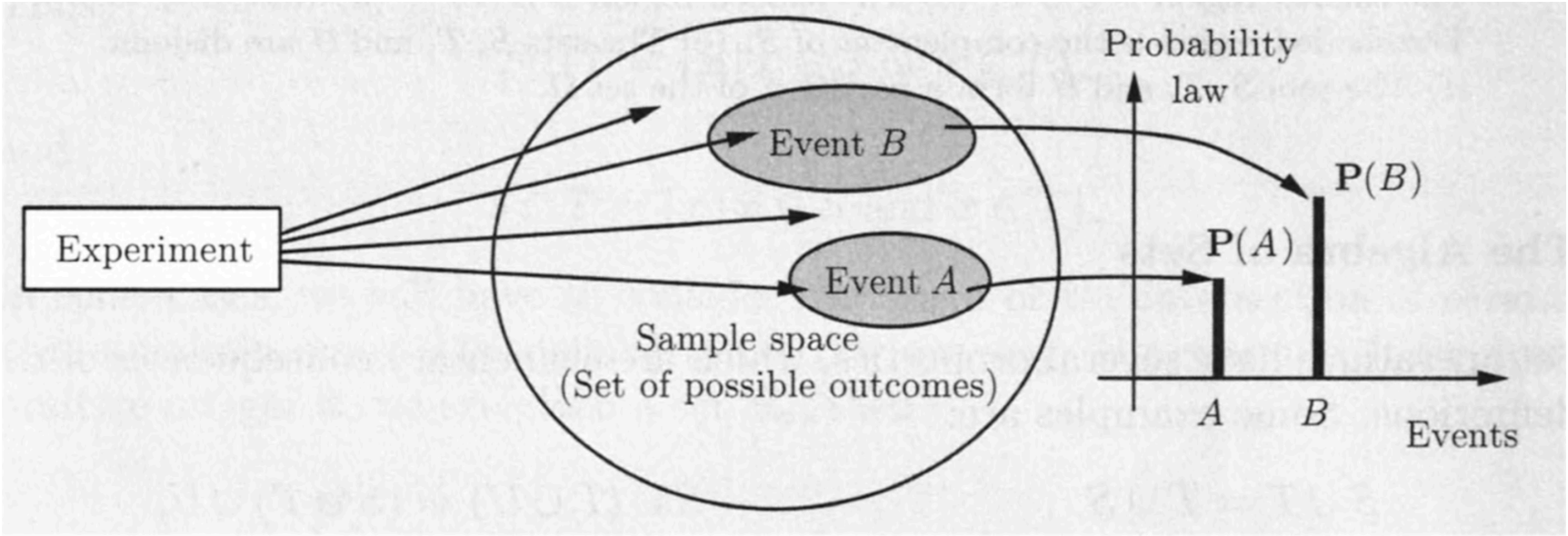
A coin is flipped 10 times. Writing Heads as 1 and Tails as 0. Then

- An outcome is a sequence $(s_1, s_2, \ldots, s_{10})$ with $s_j \in \{0, 1\}$.
- The sample space: the set of all such sequences.
- $A_j$: the event that the $j$th flip is Head.
- $B$: the event that at least one flip was Head. $(B = \bigcup_{j=1}^{10} A_j)$
- $C$: the event that all the flips were Heads. $(C = \cap_{j=1}^{10} A_j)$
- $D$: the event that there were at least two consecutive Heads. $(D = \bigcup_{j=1}^{9} (A_j \cap A_{j+1}))$

# Translation Between English & Sets

| **English** | **Sets** |
|---|---|
| *Events and occurrences* | |
| sample space | $S$ |
| $s$ is a possible outcome | $s \in S$ |
| $A$ is an event | $A \subseteq S$ |
| $A$ occurred | $s_{\text{actual}} \in A$ |
| something must happen | $s_{\text{actual}} \in S$ |
| *New events from old events* | |
| $A$ or $B$ (inclusive) | $A \cup B$ |
| $A$ and $B$ | $A \cap B$ |
| not $A$ | $A^c$ |
| $A$ or $B$, but not both | $(A \cap B^c) \cup (A^c \cap B)$ |
| at least one of $A_1, \ldots, A_n$ | $A_1 \cup \cdots \cup A_n$ |
| all of $A_1, \ldots, A_n$ | $A_1 \cap \cdots \cap A_n$ |
| *Relationships between events* | |
| $A$ implies $B$ | $A \subseteq B$ |
| $A$ and $B$ are mutually exclusive | $A \cap B = \emptyset$ |
| $A_1, \ldots, A_n$ are a partition of $S$ | $A_1 \cup \cdots \cup A_n = S, A_i \cap A_j = \emptyset$ for $i \neq j$ |

# Probabilistic Model



# Outline

# Naive Definition of Probability

- Assumption 1: finite sample space
- Assumption 2: all outcomes occur equally likely

## Definition

Let A be an event for an experiment with a finite sample space S. The naive probability of A is

$$P_{naive}(A) = \frac{|A|}{|S|} = \frac{\text{number of outcomes favorable to } A}{\text{total number of outcomes in } S}.$$

# Pascal-Fermat Correspondence: Unfinished Game

Alice and Bob play a game with a pot of 40 \$, where the one wins three tosses of a fair coin will get the whole pot. On each round, Alice chooses heads, Bob chooses tails. But for some reason they have to abandon the game after three rounds, with Alice ahead, 2 to 1. How do they divide the pot?

# Solution

# Basic Counting

- **Sampling**: sampling from a set means choosing an element (draw a sample) or multiple elements (draw samples) from that set.
- **With Replacement** & **Without Replacement**: put each element(object) back or not after each draw. Or "repetition is allowed or not".
- **Ordered** & **Unordered**: ordering matters or not.

# Basic Counting

- Ordered Sampling with Replacement
- Ordered Sampling without Replacement
- Unordered Sampling without Replacement
- Unordered Sampling with Replacement

# Basic Counting

Choose $k$ elements from a set with $n$ elements (choose $k$ objects out of $n$ distinguishable objects), the number of possible ways:

| | Order Matters | Order Not Matter |
|---|---|---|
| with replacement | | |
| without replacement | | |

# Basic Counting

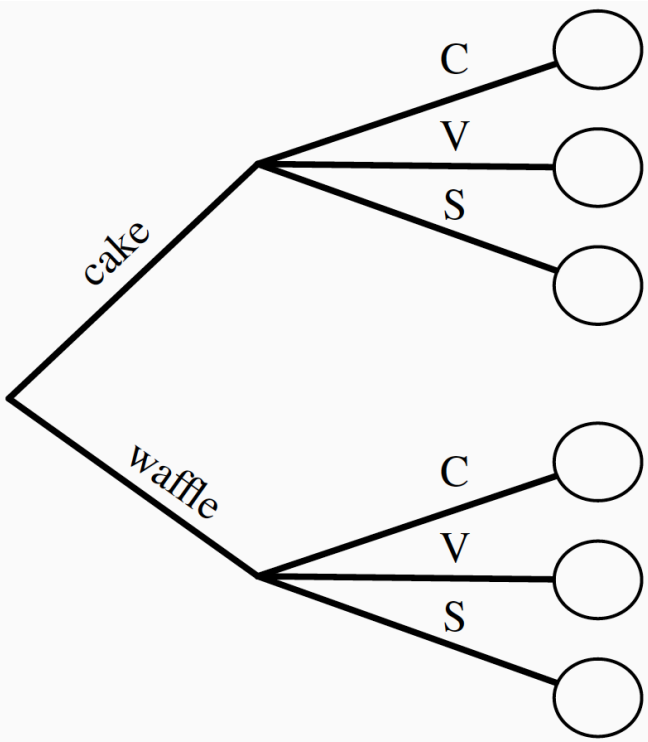
Choose $k$ elements from a set with $n$ elements (choose $k$ objects out of $n$ distinguishable objects), the number of possible ways:

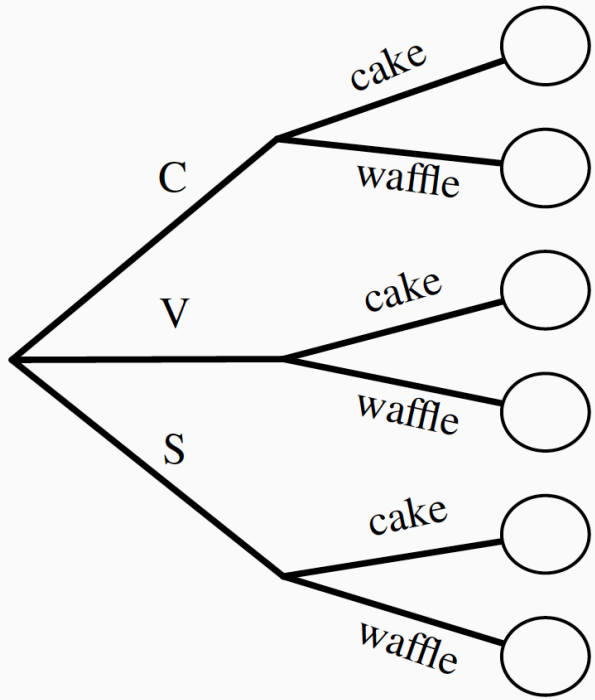
| | Order Matters | Order Not Matter |
|---|---|---|
| with replacement | ? | |
| without replacement | ? | |

# Multiplication Rule

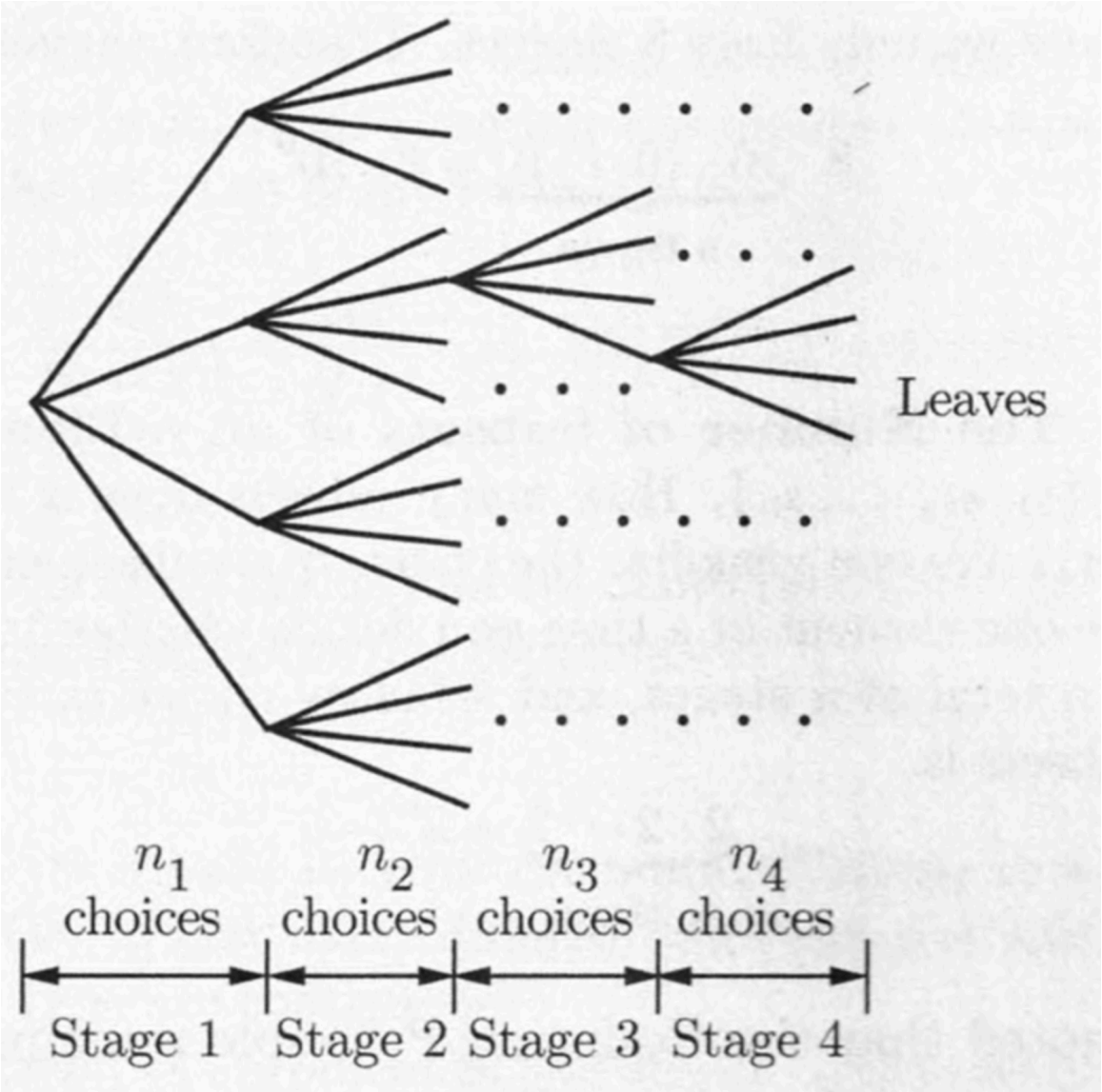
You buy an ice cream cone with several choices:

- cone: cake or waffle
- flavor: chocolate, vanilla, or strawberry



# Multiplication Rule in General



# Ordered Sampling With Replacement

## Theorem

*Consider n objects in a set and making k choices from them, one at a time with replacement (i.e., choosing a certain object does not preclude it from being chosen again). Then there are $n^k$ possible outcomes.*

# Ordered Sampling Without Replacement

## Theorem

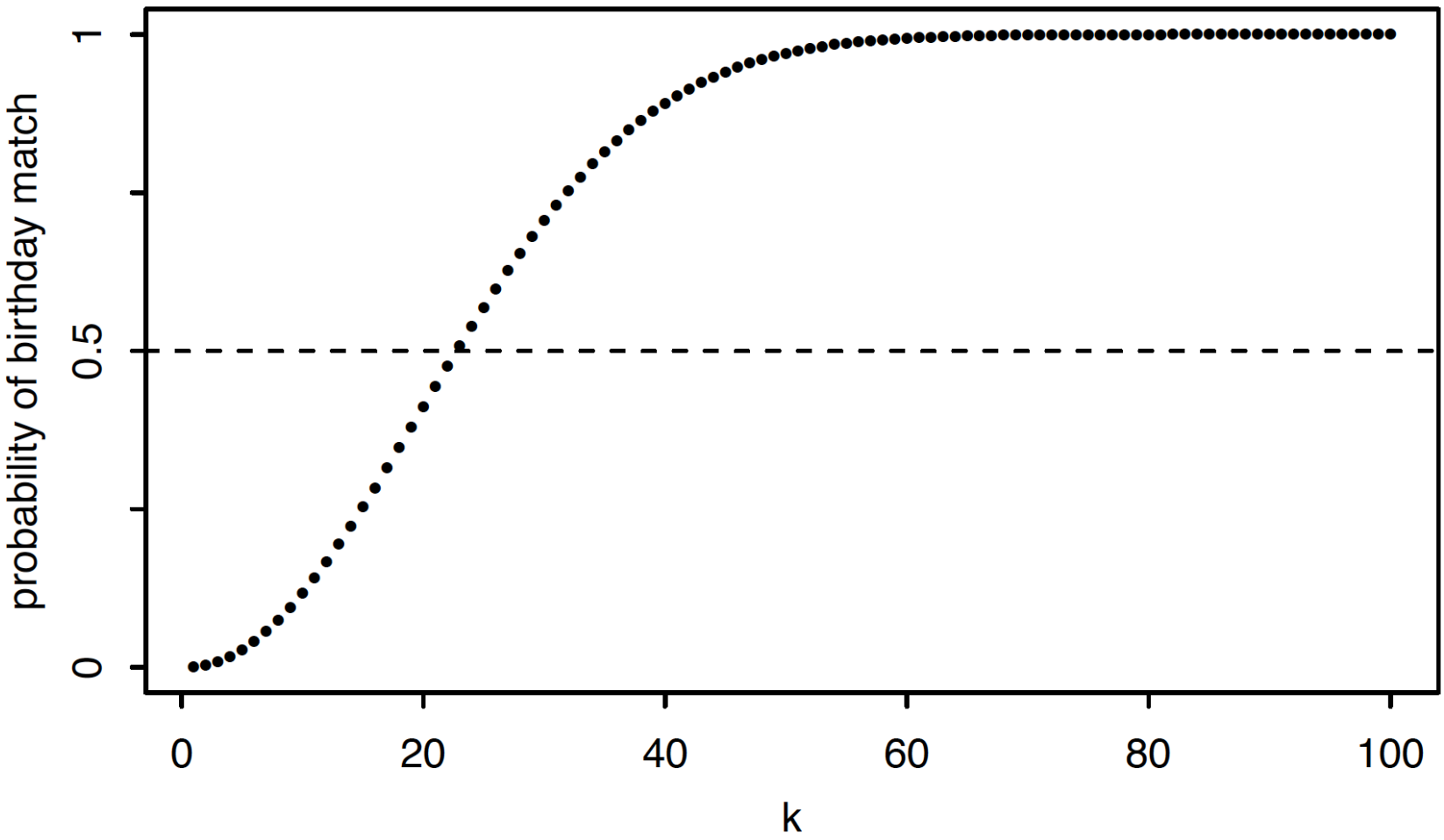
*Consider $n$ objects in a set and making $k$ choices from them, one at a time without replacement (i.e., choosing a certain object preclude it from being chosen again). Then there are $n(n-1)\cdots(n-k+1)$ possible outcomes for $k \leq n$ (and 0 possibilities for $k > n$). When $k = n$, there are $n!$ possible outcomes, each outcome is called a "permutation" of such $n$ objects.*

# Example: Birthday Problem

There are $k$ people in a room. Assume each person's birthday is equally likely to be any of the 365 days of the year (we exclude February 29), and that people's birthdays are independent (we assume there are no twins in the room). What is the probability that two or more people in the group have the same birthday?

# Solution of Birthday Problem

# Solution of Birthday Problem



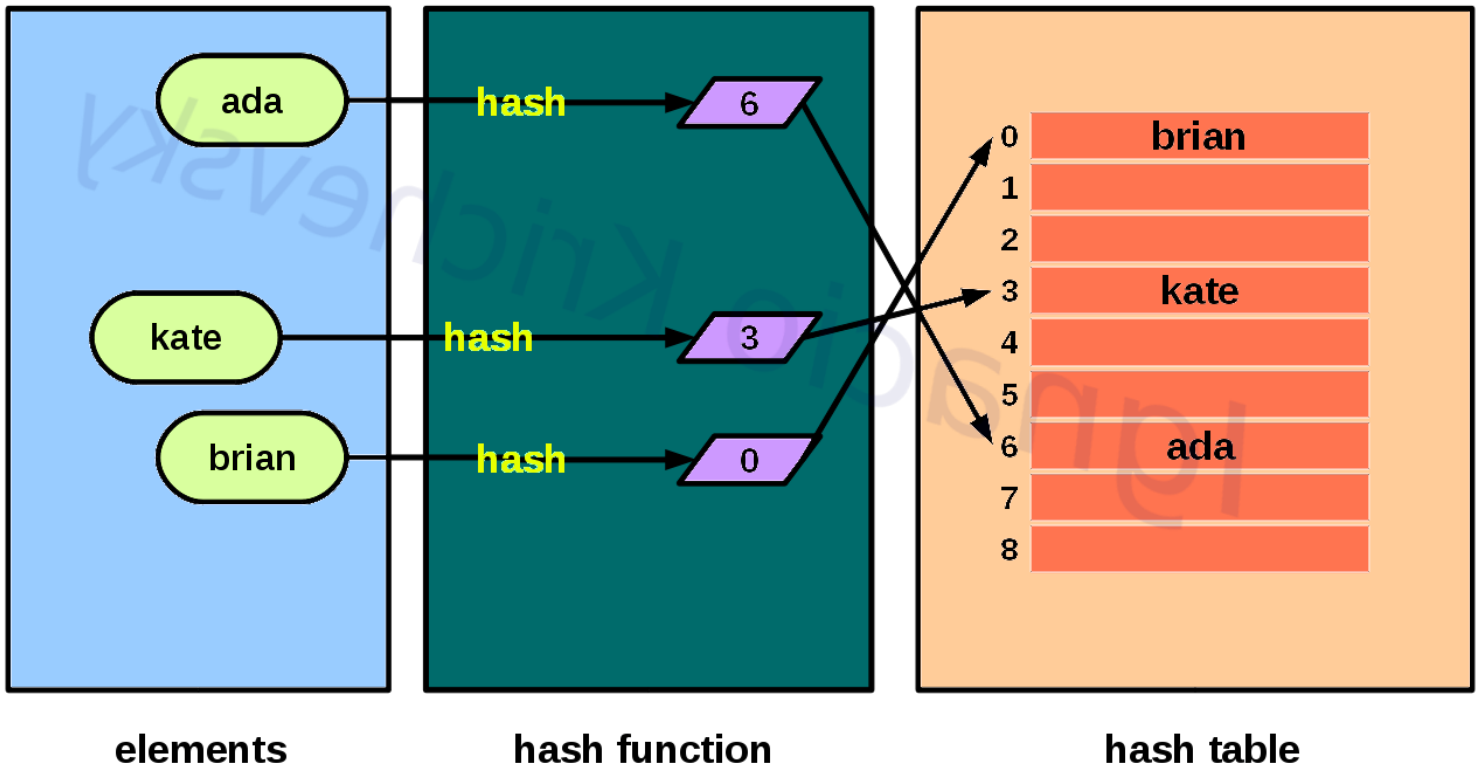
# Generalized Birthday Problem

- Each of $k$ people has a random number ("birthday") drawn from $n$ values ("days").
- If the probability that at least two people have the same number is 50%, then $k \approx 1.18\sqrt{n}$.

# Generalized Birthday Problem

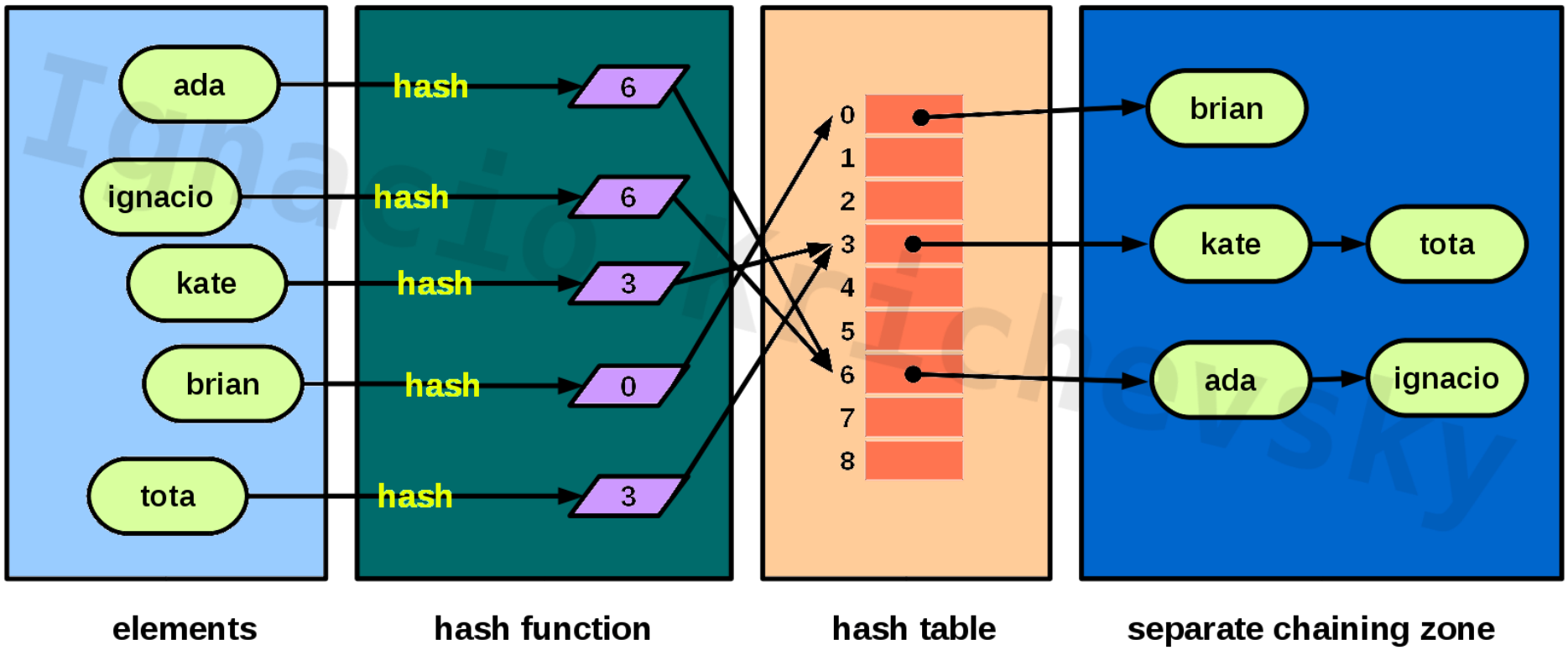
# Application: Hash Table

- A commonly used data structure for fast information retrieval
- Example: store people's name. For each people $x$, a hash function $h$ is computed.
- $h(x)$: the location that will be used to store $x$'s name.



# Hash Collision

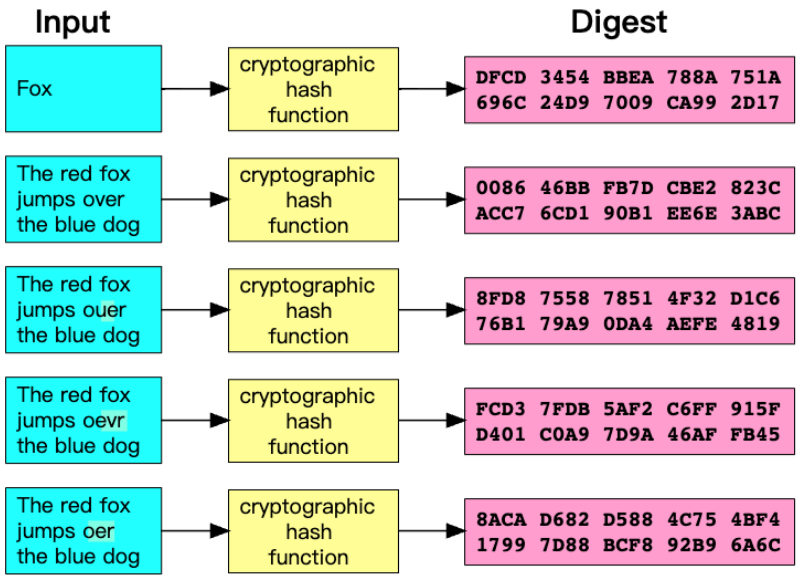
- Collision: $x \neq y$, but $h(x) = h(y)$ ($\geq 1$ locations has $\geq 2$ names stored there)
- Given $k$ people (different names) and $n$ locations, what is the probability of occurrence of hash collision?



# Solution of Hash Collision

# Application: Cryptographic Hash Function

- A map which "scrambles" long strings (message) into $m$-bit "hashes" (digest).
- Example 1: MD (Message-Digest Algorithm) 5 (bittorrent) with $m = 128$.
- Example 2: SHA(Secure Hash Algorithm)-1 (SSL,PGP) with $m = 160$, usually rendered as a hexadecimal number which is 40 digits long.



# Cryptographic Hash Function

- A good cryptographic hash function $f$ has two properties:
  - Given the hash $f(M)$ of a message string $M$, it's computationally infeasible to recover $M$.
  - It's computationally infeasible to find a "collision", meaning a pair of distinct messages $M1 \neq M2$ such that $f(M1) = f(M2)$.

# The Birthday Attack

- Suppose we try to "break" a hash function by finding a collision (forged digital signature).
- One method: take a huge number of messages $M$, hash them all, and hope to find two with the same hash value.
- Now how many messages would you have to try before there was at least a 50% chance of finding two with the same hash?
- $n = 2^m$ and $k \approx \sqrt{n} = 2^{m/2}$. For SHA-1, $k \approx 2^{80}$.

# The Birthday Attack

- The birthday attack: trying to find collisions by testing many random messages.
- A cryptographic hash function is broken: when there is a way of finding collisions much faster than the Birthday Attack method.
- SHA-1 is now broken: Xiaoyun Wang find a SHA-1 collision with $2^{69}$ tests in 2005, then $2^{63}$ tests later (compared to $2^{80}$ tests by the birthday attack).
- Now SHA-2 (m=256) and SHA-3 (m=512) are on the way.

# Summary of Counting

Choose $k$ elements from a set with $n$ elements (choose $k$ objects out of $n$ distinguishable objects), the number of possible ways:

| | Order Matters | Order Not Matter |
|---|---|---|
| with replacement | $n^k$ | |
| without replacement | $n(n-1)\cdots(n-k+1)$ | ? |

# Unordered Sampling without Replacement

- also called Combination
- $k$-combination: choose a $k$-element subset of a set with $n$ elements

# Binomial Coefficient

## Definition

For any nonnegative integers $k$ and $n$, the binomial coefficient $\binom{n}{k}$, read as "$n$ choose $k$", is the number of subsets of size $k$ for a set of size $n$.

## Theorem

*For $k \leq n$, we have*

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!} = \frac{n!}{k!(n-k)!}$$

# Binomial Theorem

## Theorem

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}.$$

# Multinomial Theorem

# Multinomial Theorem

## Theorem

$$(x_1 + x_2 + \cdots + x_r)^n = \sum_{n_1, n_2, \cdots, n_r \geq 0} \frac{n!}{n_1! n_2! \cdots n_r!} x_1^{n_1} x_2^{n_2} \cdots x_r^{n_r}$$

*where* $n_1 + n_2 + \cdots + n_r = n$.

# Story Proof: The Team Captain

For any positive integers $n$ and $k$ with $k \leq n$,

$$n\binom{n-1}{k-1} = k\binom{n}{k}$$

# Story Proof: Vandermonde's Identity

A famous relationship between binomial coefficients, called *Vandermonde's identity*, says that

$$\binom{m+n}{k} = \sum_{j=0}^{k} \binom{m}{j}\binom{n}{k-j}$$

# Summary of Counting

Choose $k$ elements from a set with $n$ elements (choose $k$ objects out of $n$ distinguishable objects), the number of possible ways:

| | Order Matters | Order Not Matter |
|---|---|---|
| with replacement | $n^k$ | ? |
| without replacement | $n(n-1)\cdots(n-k+1)$ | $\binom{n}{k}$ |

# Unordered Sampling with Replacement

- How many ways are there to choose $k$ times from a set of $n$ objects with replacement, if order doesn't matter (we only care about how many times each object was chosen, not the order in which they were chosen)?
- also called "Bose-Einstein Counting"

# Equivalent Problem

# Equivalent Problem

## Theorem

*There are* $\binom{r-1}{n-1}$ *distinct* **positive** *integer-valued vectors* $(x_1, x_2, \ldots, x_n)$ *satisfying the equation*

$$x_1 + x_2 + \cdots + x_n = r, x_i > 0, i = 1, 2, \ldots, n.$$

# Bose-Einstein Counting

## Theorem

*There are* $\binom{n+k-1}{n-1}$ *distinct* **nonnegative** *integer-valued vectors* $(x_1, x_2, \ldots, x_n)$ *satisfying the equation*

$$x_1 + x_2 + \cdots + x_n = k, x_i \geq 0, i = 1, 2, \ldots, n.$$

# Example

How many distinct **positive** integer-valued vectors $(x_1, x_2, x_3, x_4)$ satisfying the equation

$$x_1 + x_2 + x_3 + x_4 = 88, \text{where } x_1 \geq 3, x_2 \geq 5, x_3 \geq 8, x_4 \geq 10.$$

# Solution

# Example: Multinomial Expansion

How many terms are there in the multinomial expansion of $(x_1 + x_2 + \ldots + x_r)^n$?

# Example: Multinomial Expansion

# Summary of Counting

Choose $k$ objects out of $n$ objects, the number of possible ways:

| | Order Matters | Order Not Matter |
|---|---|---|
| with replacement | $n^k$ | $\binom{n+k-1}{k}$ |
| without replacement | $n(n-1)\cdots(n-k+1)$ | $\binom{n}{k}$ |

# Outline

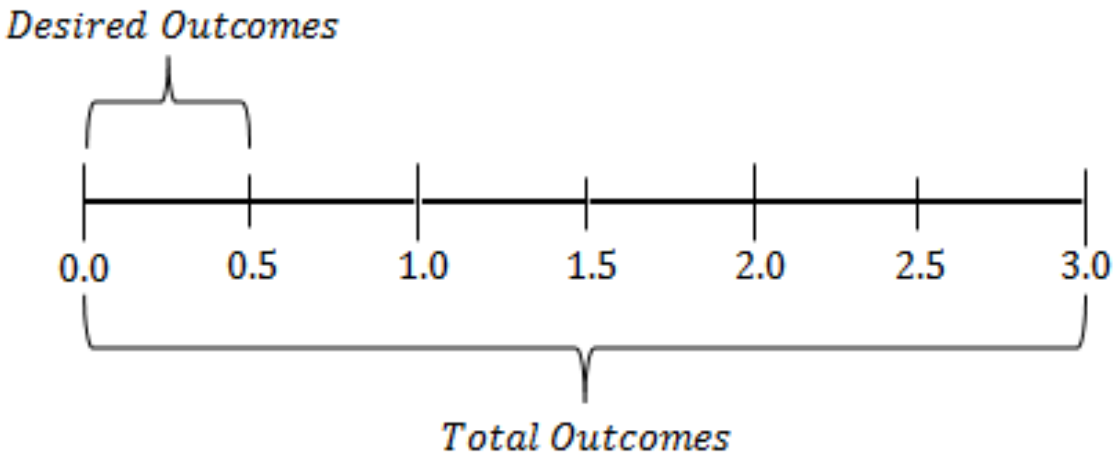
# Geometric Probability: Infinite Sample Space

- Geometric probability is a tool to deal with the problem of infinite outcomes by measuring the number of outcomes geometrically, in terms of geometric measure such as length, area, or volume.
- Equally likely means the probability of falling into some geometric region is proportional to the measure of such region including length, area, or volume.
- Given a sample space $S$, the probability of event $A$ occurring is $P(A) = \frac{M(A)}{M(S)}$, where $M(\cdot)$ is the measure of geometric region.

# Example: 1-dimensional Geometric Probability

$X$ is a random real number between 0 and 3. What is the probability that $X$ is closer to 0 than it is to 1?

# Solution



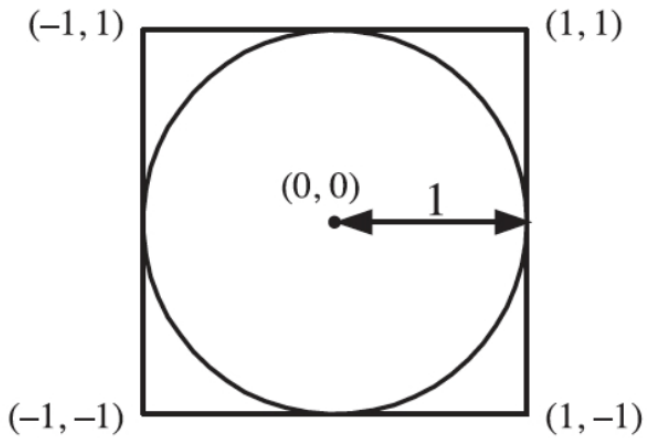
# Example: 1-dimensional Geometric Probability

$X$ is a random real number between 1 and 2. What is the probability that $X$ is equal to 1.5?

# Solution

# Example: 2-dimensional Geometric Probability

A point is chosen uniformly at random in the square. What is the probability that it will land within the circle?



# Solution

# Probability: A Statistical Defintion

- *The frequentist view*: probability represents a long-run frequency over a large number of repetitions of an experiment.
- if we say a coin has probability $1/2$ of Heads, that means the coin would land Heads 50% of the time if we tossed it over and over and over.
- However, the frequency may not exist in general.
- Now as the intuition behind the Monte Carlo Computing Method.

# Probability: A Subjective Defintion

- Probability represents a degree of belief about the event in question.
- So we can assign probabilities to hypotheses like "candidate A will win the election" or "the defendant is guilty" even if it isn't possible to repeat the same election or the same crime over and over again.
- Related to Logic, Philosophy, and Psychology.

# Outline

# Axioms for Events

## Definition

Given a sample space $S$, the class of subsets of $S$ that constitute the set of events satisfies the following axioms:

1. $S$ is an event.
2. For every event $A$, the complement $A^c$ is an event.
3. For every sequence of events $A_1, A_2, \ldots$, the union $\bigcup_{j=1}^{\infty} A_j$ is an event.

# General Definition of Probability

## Definition

A *probability space* consists of a *sample space* $S$ and a *probability function* $P$ which takes an event $A \subseteq S$ as input and returns $P(A)$, a real number between 0 and 1, as output. The function P must satisfy the following axioms:

1. $P(\emptyset) = 0$, $P(S) = 1$.
2. If $A_1, A_2, \ldots$ are disjoint events, then

$$P(\bigcup_{j=1}^{\infty} A_j) = \sum_{j=1}^{\infty} P(A_j)$$

(Saying that these events are disjoint means that they are mutually exclusive: $A_i \cap A_j = \emptyset$ for $i \neq j$.

# Properties of Probability

Probability has the following properties, for any events A and B:

1. $P(A^c) = 1 - P(A)$.
2. If $A \subseteq B$, then $P(A) \leq P(B)$.
3. $P(A \bigcup B) = P(A) + P(B) - P(A \cap B)$.

# Example: Bonferroni's Inequality

## Theorem

*For any $n$ events $A_1, \ldots, A_n$, we have*

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) \geq P(A_1) + P(A_2) + \cdots + P(A_n) - (n-1).$$

# Proof

# Inclusion-Exclusion Formula

For any events $A_1, \ldots, A_n$:

$$P(\bigcup_{i=1}^{n} A_i) = \sum_i P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k)$$
$$+ \ldots + (-1)^{n+1} P(A_1 \cap \cdots \cap A_n).$$

# Example: De Montmort's Matching Problem

Consider a well-shuffled deck of $n$ cards, labeled 1 through $n$. You flip over the cards one by one, saying the numbers 1 through $n$ as you do so. You win the game if, at some point, the number you say aloud is the same as the number on the card being flipped over (for example, if the 7th card in the deck has the label 7). What is the probability of winning?

# Solution

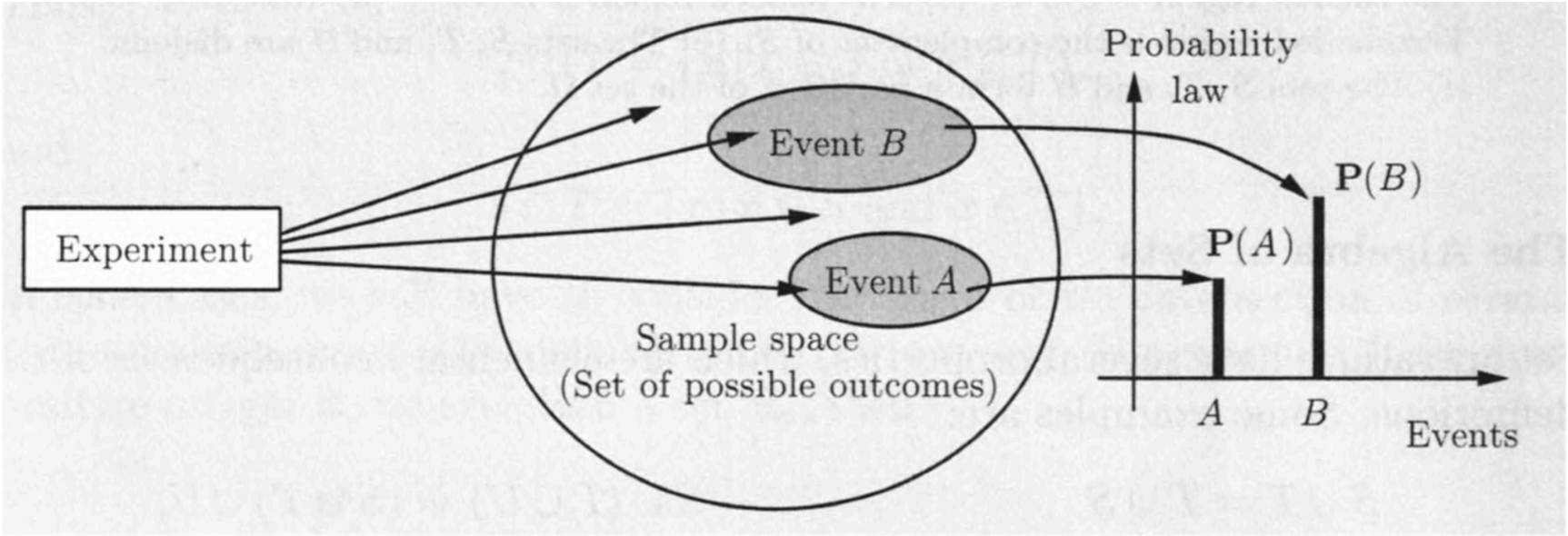
# Solution

# Summary 1: Events & Numbers



| events | | numbers |
|---|---|---|
| $A$ | | $P(A)$ |
| not $A$ | | $P(A^c) = 1 - P(A)$ |
| $A$ and $B$ | $P$ → | $P(A \cap B)$ |
| $A$ or $B$ | | $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ |
| something happened | | $P(S) = 1$ |

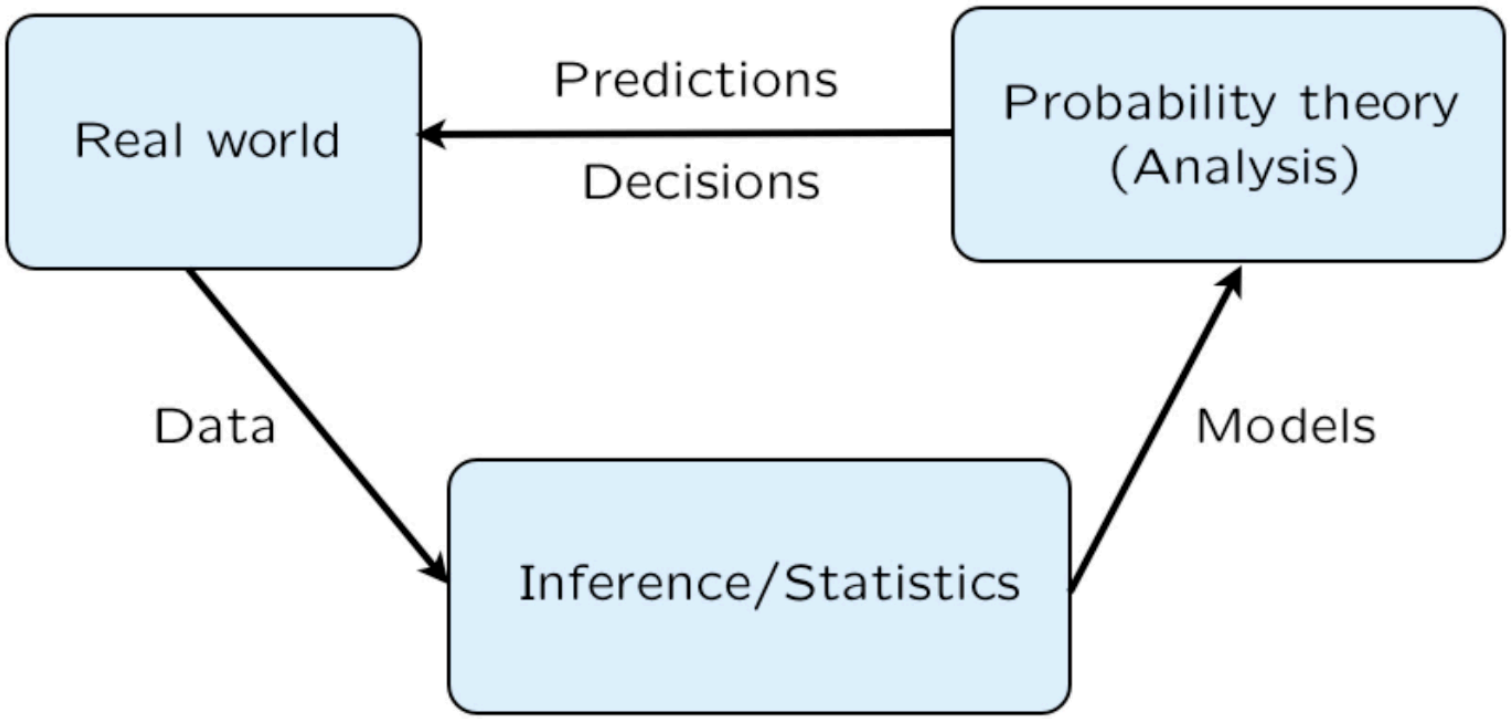
# Summary 2: Probability Space



# Summary 3: The Role of Probability & Statistics

A framework for analyzing phenomena with uncertain outcomes:

- Rules for consistent reasoning
- Used for predictions and decisions



# References

- Chapter 1 of **BH**
- Chapter 1 of **BT**