The Birthday “Paradox”

Suppose there are \( m \) students in a class. What is the probability that at least two students in the class have the same birthday?

*In a class of 70 students, this probability is about 99.87%*
CS145: Lecture 1 Outline

- Sample spaces: Sets of possible outcomes
- Probability: Counting and the Discrete Uniform Law
- Example: The birthday paradox
1.2 PROBABILISTIC MODELS

A probabilistic model is a mathematical description of an uncertain situation. It must be in accordance with a fundamental framework that we discuss in this section. Its two main ingredients are listed below and are visualized in Fig. 1.2.

**Elements of a Probabilistic Model**

- The **sample space** \( \Omega \), which is the set of all possible outcomes of an experiment.

- The **probability law**, which assigns to a set \( A \) of possible outcomes (also called an *event*) a nonnegative number \( P(A) \) (called the *probability* of \( A \)) that encodes our knowledge or belief about the collective “likelihood” of the elements of \( A \). The probability law must satisfy certain properties to be introduced shortly.

**Experiment**

**Sample Space** \( \Omega \) (Set of Outcomes)

**Event A**

**Event B**

**Sample Spaces and Events**

Every probabilistic model involves an underlying process, called the *experiment*, that will produce exactly one out of several possible outcomes. The set of all possible outcomes is called the *sample space* of the experiment, and is denoted by \( \Omega \). A subset of the sample space, that is, a collection of possible outcomes, is called an *event*. For example, flip a coin, roll a die, receive an email, take a picture, …
A set is a collection of objects, which are elements of the set.

A set can be finite, \( S = \{1, 2, \ldots, n\} \). Cardinality (size): \( |S| = n \)

A set can be countably infinite:

\[ S = \{ x \mid x = 2k + 1 \text{ or } x = -2k + 1, \ k \text{ integer} \} \]
\[ = \{ 1, -1, 3, -3, 5, -5, \ldots \} \]

A set can be uncountable, \( S = \{ x \mid x \in [0, 1] \} \).

A set can be empty \( S = \emptyset \).
Sets: Elements & Relationships

• $x \in S$ - the element $x$ is a member of the set $S$
• $x \notin S$ - the element $x$ is not a member of the set $S$
• $\exists x$ - there exists $x$...
• $\forall$ - for all elements $x$ ...
• $T \subseteq S$ - $\forall x \in T, x \in S$
• $T \subset S$ - $\forall x \in T, x \in S$ AND $\exists x \in S$ such that $x \notin T$. 
Sets: Combination & Manipulation

- A base set $\Omega$, all sets are subsets of $\Omega$
- Basic operations: for $S, T \subseteq \Omega$,
  - $S \cup T = \{x \mid x \in S \text{ or } x \in T\}$  
    ➤ union
  - $S \cap T = \{x \mid x \in S \text{ and } x \in T\}$  
    ➤ intersection
  - $\bar{S} = S^c = \{x \mid x \notin S\}$  
    ➤ complement
- De Morgan's laws:
  - $(S \cup T)^c = \bar{S} \cap \bar{T}$
  - $(S \cap T)^c = \bar{S} \cup \bar{T}$

[Diagram of Venn Diagram with $S$, $T$, and $\Omega$]
Visualizing Sets: Venn Diagrams

- **Basic operations**: for $S, T \subseteq \Omega$,
  - $S \cup T = \{ x \mid x \in S \text{ or } x \in T \}$
  - $S \cap T = \{ x \mid x \in S \text{ and } x \in T \}$
  - $\bar{S} = S^c = \{ x \mid x \notin S \}$

- **De Morgan’s laws**:
  - $(S \cup T)^c = \bar{S} \cap \bar{T}$
  - $(S \cap T)^c = \bar{S} \cup \bar{T}$
  - $(\bigcup_{i \in I} S_i)^c = \bigcap_{i \in I} \bar{S}_i$
  - $(\bigcap_{i \in I} S_i)^c = \bigcup_{i \in I} \bar{S}_i$
A set $B$ is **partitioned** into $n$ subsets if:

$$B_1 \cup B_2 \cup \cdots \cup B_n = B$$

$$B_i \cap B_j = \emptyset \text{ for any } i \neq j$$

*mutually disjoint*
The Sample Space

\[ \Omega \] a set, or unordered “list”, of possible outcomes from some random (not deterministic) experiment

“Omega”

The list defining the sample space must be:

- **Mutually exclusive:** Each experiment has a unique outcome.
- **Collectively exhaustive:** No matter what happens in the experiment, the outcome is an element of the sample space.
- **An art:** Choosing the “right” granularity, to capture the phenomenon of interest as simply as possible.

*Modeling in science and engineering involves tradeoffs between accuracy, simplicity, & tractability.*
A Finite Sample Space

You roll a tetrahedral (4-sided) die 2 times.

\[ \Omega = \{ (x, y) \mid x \in \{1, 2, 3, 4\}, y \in \{1, 2, 3, 4\} \} \]

- Formally, sample space is a set of \(4^2 = 16\) discrete outcomes
- Can also model outcome via tree-based sequential description
A Finite Sample Space

You toss a (2-sided) coin 10 times.

Two possible sample spaces for this experiment:
1. You record the number of times the coin comes up heads:
   \[ \Omega = \{0, 1, 2, \ldots, 9, 10\} \]
2. You record the full sequence of head-tail outcomes:
   \[ \Omega = \{H, T\}^{10} \text{ all } 2^{10} \text{ possible H-T sequences} \]

Which is better? It depends on what you want to model:

**Game 1:** Receive $1 each time a head appears.

**Game 2:** Receive $1 per coin toss, up to and including the first head. Then receive $2 per coin toss, up to and including the second head. Then receive $4 per coin toss, up to and including the third head. More generally, dollar amount per toss is doubled with each head…
CS145: Lecture 1 Outline

- Sample spaces: Sets of possible outcomes
- Probability: Counting and the Discrete Uniform Law
- Example: The birthday paradox
The Discrete Uniform Law

Formalizes the idea of “completely random” sampling.

- Let all outcomes be equally likely
- Then,
  \[
  P(A) = \frac{\text{number of elements of } A}{\text{total number of sample points}}
  \]
- Computing probabilities \(\equiv\) counting
- Defines fair coins, fair dice, well-shuffled decks
Uniform Law for a Finite Sample Space

You roll a tetrahedral (4-sided) die 2 times.

- Let every possible outcome have probability $1/16$
  - $P((X,Y) \text{ is } (1,1) \text{ or } (1,2)) =$
  - $P\{X = 1\} =$
  - $P(X + Y \text{ is odd}) =$
  - $P(\min(X, Y) = 2) =$
## Events and Sets

<table>
<thead>
<tr>
<th>Event language</th>
<th>Set language</th>
<th>Set notation</th>
<th>Venn diagram</th>
</tr>
</thead>
<tbody>
<tr>
<td>outcome space</td>
<td>universal set</td>
<td>( \Omega )</td>
<td><img src="image1" alt="Diagram" /></td>
</tr>
<tr>
<td>event</td>
<td>subset of ( \Omega )</td>
<td>( A, B, C, \text{ etc.} )</td>
<td><img src="image2" alt="Diagram" /></td>
</tr>
<tr>
<td>impossible event</td>
<td>empty set</td>
<td>( \emptyset )</td>
<td><img src="image3" alt="Diagram" /></td>
</tr>
<tr>
<td>not ( A ), opposite of ( A )</td>
<td>complement of ( A )</td>
<td>( A^c )</td>
<td><img src="image4" alt="Diagram" /></td>
</tr>
<tr>
<td>either ( A ) or ( B ) or both</td>
<td>union of ( A ) and ( B )</td>
<td>( A \cup B )</td>
<td><img src="image5" alt="Diagram" /></td>
</tr>
<tr>
<td>both ( A ) and ( B )</td>
<td>intersection of ( A ) and ( B )</td>
<td>( A \cap B )</td>
<td><img src="image6" alt="Diagram" /></td>
</tr>
<tr>
<td>( A ) and ( B ) are mutually exclusive</td>
<td>( A ) and ( B ) are disjoint</td>
<td>( A \cap B = \emptyset )</td>
<td><img src="image7" alt="Diagram" /></td>
</tr>
<tr>
<td>if ( A ) then ( B )</td>
<td>( A ) is a subset of ( B )</td>
<td>( A \subseteq B )</td>
<td><img src="image8" alt="Diagram" /></td>
</tr>
</tbody>
</table>
Consider a process that consists of $r$ stages. Suppose that:

(a) There are $n_1$ possible results for the first stage.

(b) For every possible result of the first stage, there are $n_2$ possible results at the second stage.

(c) More generally, for all possible results of the first $i - 1$ stages, there are $n_i$ possible results at the $i$th stage.

Then, the total number of possible results of the $r$-stage process is

$$n_1 \cdot n_2 \cdots n_r.$$ 

**Simple examples:**

- Number of license plates with 3 letters and 4 digits $=$

- … if repetition is prohibited $=$

---

The set of choices at each stage can depend on previous choices, as long as the number of choices at each stage is constant.
Permutations and Subsets

- **Permutations**: Number of ways of ordering $n$ elements is:
  \[ n(n-1)(n-2)\cdots 1 = \prod_{i=1}^{n} i = n! \]

- Number of subsets of \{1,\ldots,n\} =
  \[ 2 \cdot 2 \cdot \cdots 2 = 2^n \]
  \[ n \text{ times} \]

Illustration of the basic counting principle. The counting is carried out in
4 stages through the use of a tree. For example, consider an experiment that
consists of two consecutive stages. The possible results of the two-stage experiment
are all possible outcomes of such an experiment. The number of possible results at
any stage is a power of 2.

- Binomial probabilities: We saw there that the probability of each distinct
  outcome of such an experiment is equal to
  \[ \binom{n}{k} \frac{1}{2^n} \]

- Probability that six rolls of a six-sided die result in heads is easily obtained, but
  the calculation of the probability that six rolls result in tails is
  complicated. This probability is
  \[ \binom{6}{3} \frac{1}{2^6} \]

- Combinations and permutations: We can see that the number of possible
  ways of choosing a subset of \{1,\ldots,n\} is
  \[ \binom{n}{r} \]
  \[ r \text{ elements} \]

- Permutations: We can see that the number of possible
  ways of choosing a subset of \{1,\ldots,n\} is
  \[ \binom{n}{r} \]
  \[ r \text{ elements} \]

- Number of subsets of \{1,\ldots,n\} =
  \[ 2 \cdot 2 \cdot \cdots 2 = 2^n \]
  \[ n \text{ times} \]

- Probability that six rolls of a six-sided die result in heads is easily obtained, but
  the calculation of the probability that six rolls result in tails is
  complicated. This probability is
  \[ \binom{6}{3} \frac{1}{2^6} \]

- Number of subsets of \{1,\ldots,n\} =
  \[ 2 \cdot 2 \cdot \cdots 2 = 2^n \]
  \[ n \text{ times} \]
Combinations

- \( \binom{n}{k} \): number of \( k \)-element subsets of a given \( n \)-element set

- Two ways of constructing an ordered sequence of \( k \) distinct items:
  - Choose the \( k \) items one at a time:
    \[
    n(n-1) \cdots (n-k+1) = \frac{n!}{(n-k)!} \text{ choices}
    \]
  - Choose \( k \) items, then order them (\( k! \) possible orders)

- Hence:
  \[
  \binom{n}{k} \cdot k! = \frac{n!}{(n-k)!} \]
  \[
  \binom{n}{k} = \frac{n!}{k!(n-k)!}
  \]

- The number of total subsets:
  \[
  \sum_{k=0}^{n} \binom{n}{k} = 2^n
  \]
Pascal’s Triangle

\[ \binom{n}{k} = \frac{n!}{k!(n-k)!} \]

\[ \sum_{k=0}^{n} \binom{n}{k} = 2^n \]
Pascal’s Triangle

http://www.mathwarehouse.com/
Binomial Probabilities

If I toss a coin $n$ times, what is the probability that I see $k$ heads?

- $n$ independent coin tosses
  \[ P(H) = p = \frac{1}{2} \]

- $P(HTTHHH) =$

- $P(\text{sequence}) = p^\# \text{ heads} (1 - p)^\# \text{ tails}$

\[
P(k \text{ heads}) = \sum_{k\text{-head seq.}} P(\text{seq.})
\]

\[= (\text{# of } k\text{-head seqs.}) \cdot p^k(1 - p)^{n-k}\]

\[= \binom{n}{k} p^k(1 - p)^{n-k}\]
CS145: Lecture 1 Outline

- Sample spaces: Sets of possible outcomes
- Probability: Counting and the Discrete Uniform Law
- Example: The birthday paradox
The Birthday “Paradox”

Suppose there are \( m \) students in a class. What is the probability that at least two students in the class have the same birthday?

In a class of 70 students, this probability is about 99.87%.

Assumptions:
- Birthdays are equally likely to occur on any of \( N=365 \) days.
- No dependence between birthdays of different students.

Not completely true, but fairly accurate approximations.
The Birthday “Paradox”

- Sort students in arbitrary order (say, alphabetical by name)
- Define two events for student \( j \) in the list of \( m \) students:
  - \( R_j \rightarrow \) birthday \( j \) is a repeat of some previous student
  - \( D_j \rightarrow \) all of the first \( j \) birthdays are distinct

- Tree diagram of event probabilities for \( N=365 \) days:
The Birthday “Paradox”

➢ The probability that the first $m$ birthdays are distinct is then:

$$P(D_m) = \left(1 - \frac{1}{N}\right)\left(1 - \frac{2}{N}\right) \cdots \left(1 - \frac{m-1}{N}\right) = \prod_{i=0}^{m-1} \left(1 - \frac{i}{N}\right)$$

$R_j \rightarrow$ birthday $j$ is a repeat of some previous student  
$D_j \rightarrow$ all of the first $j$ birthdays are distinct

➢ Tree diagram of event probabilities for $N=365$ days:
The Birthday “Paradox”

The probability that the first \( m \) birthdays are distinct is then:

\[
P(D_m) = \left(1 - \frac{1}{N}\right) \left(1 - \frac{2}{N}\right) \cdots \left(1 - \frac{m-1}{N}\right) = \prod_{i=0}^{m-1} \left(1 - \frac{i}{N}\right)
\]

We can also compute this probability by counting elements of the sample space of all \( N^m \) possible birthday patterns:

\[
\Omega = \{(b_1, \ldots, b_m) \mid b_i \in \{1, \ldots, N\}\}
\]

The number of birthday patterns with all pairs distinct:

\[
D_m = \{(b_1, \ldots, b_m) \mid b_i \neq b_j \text{ for all } i \neq j\}
\]

\[
|D_m| = N(N - 1)(N - 2) \cdots (N - m + 1) = \frac{N!}{(N - m)!}
\]
The Birthday “Paradox”

The probability that the first \( m \) birthdays are distinct:

\[
P(D_m) = \left(1 - \frac{1}{N}\right)\left(1 - \frac{2}{N}\right) \cdots \left(1 - \frac{m-1}{N}\right) = \prod_{i=0}^{m-1} \left(1 - \frac{i}{N}\right)
\]

\[
P(D_m) = \frac{N!}{(N - m)!N^m} = \prod_{i=0}^{m-1} \left(\frac{N - i}{N}\right)
\]

The number of \textit{birthday patterns with all pairs distinct}:

\[
D_m = \{(b_1, \ldots, b_m) \mid b_i \neq b_j \text{ for all } i \neq j\}
\]

\[
|D_m| = N(N - 1)(N - 2) \cdots (N - m + 1) = \frac{N!}{(N - m)!}
\]
The Birthday “Paradox”

The probability that \( m \) birthdays on \( N \) days are distinct:

\[
P(D_m) = \left( 1 - \frac{1}{N} \right) \left( 1 - \frac{2}{N} \right) \cdots \left( 1 - \frac{m-1}{N} \right) = \prod_{i=0}^{m-1} \left( 1 - \frac{i}{N} \right)
\]

\[
1 - P(D_m) \leq \prod_{i=0}^{m-1} e^{-\frac{i}{N}} = e^{-\sum_{i=0}^{m-1} \frac{i}{N}} = e^{-\frac{m(m-1)}{2N}}
\]

\[1 - P(D_{23}) = 0.506\]
The Birthday “Paradox”

- The probability that $m$ birthdays on $N$ days are distinct:

\[
P(D_m) = \left(1 - \frac{1}{N}\right) \left(1 - \frac{2}{N}\right) \cdots \left(1 - \frac{m-1}{N}\right) = \prod_{i=0}^{m-1} \left(1 - \frac{i}{N}\right)
\]

Probability that first repeated birthday is found with student $j$.