Some pictures are taken from the internet. Credits belong to original authors.
Basic language and notation

Experiment $\rightarrow$ outcome $\rightarrow$ number

Example 1: Roll a die

Sample space $\Omega$: The set of all the outcomes (or sample points, elements).
Basic language and notation

**Experiment** $\rightarrow$ **outcome** $\rightarrow$ **number**

**Example 1**: Roll a die

**Sample space** $\Omega$: The set of all the outcomes.

**Event** $A$:
1. A **statement** about the outcome, e.g., bigger than 4.
2. A **subset** of sample space, e.g., $\{5, 6\}$. 
Basic language and notation

**Experiment** → **outcome** → **number**

**Example 1:** Roll a die

![Diagram of a dice with sample points and sample space]

**Probability:** defined on event:

\[
P(A) = \frac{|A|}{|\Omega|} = \frac{2}{6} = \frac{1}{3}.
\]

Assume the die is fair so that all the outcomes are **equally likely**. 

|A| counts the size of A, i.e., the number of elements in A.
**Basic language and notation**

**Experiment → outcome → number**

**Example 1:** Roll a die

Random variable: Let $X$ be the number:

$$P(X > 4) = \frac{1}{3}.$$

An event is a **math statement** about the random variable.
Basic language and notation

**Experiment → outcome → number**

**Example 1**: Roll a die

![Diagram of a die with sample points and sample space]

**Conditional probability**: Let $B$ be the event that the number is 6. Given that $A$ happens, what is the probability of $B$?

$$P(B|A) = \frac{1}{2}.$$

As if we randomly sample a number from $A$.

As if $A$ is the sample space.
Basic language and notation

**Experiment → outcome → number**

**Example 1:** Roll a die

Random variable

\[ P(X = 6 | X > 4) = \frac{1}{2}. \]
Basic language and notation

**Example 1**: Roll a die

![Diagram of rolling dice](image)

**Complement**

Statement: Not $A$

Subset: $A^c = \{1, 2, 3, 4\}$. 
Example 1: Roll a die

A = \{1,2,3\}
B = \{3,4,5\}
A \cup B = \{1,2,3,4,5\}

Union
Statement: \( A \) or \( B \).
Subset: \( A \cup B \).
**Example 1**: Roll a die

\[
A = \{1,2,3,4\} \\
B = \{3,4,5,6\} \\
A \cap B = \{3,4\}
\]

**Intersection**

Statement: \( A \) and \( B \).
Subset: \( A \cap B \).
**Population proportion**

**Experiment** → outcome → number

**Example 2**: Sample a random person from a population of 100 people, 50 males and 50 females. 30 males are taller than 6 ft, 10 females are taller than 6 ft. The sample space $\Omega$ is the whole population.

<table>
<thead>
<tr>
<th></th>
<th>male</th>
<th>female</th>
</tr>
</thead>
<tbody>
<tr>
<td>taller than 6 ft</td>
<td>30</td>
<td>10</td>
</tr>
<tr>
<td>shorter than 6 ft</td>
<td>50</td>
<td>50</td>
</tr>
</tbody>
</table>
**Population proportion**

**Experiment** $\rightarrow$ **outcome** $\rightarrow$ **number**

**Example 2**: Let $A$ be the event that the person is male. Let $B$ be the event that the person is taller than 6 feet (or simply the person is tall). $A$ is the sub-population of males, and $B$ is the sup-population of tall people.

<table>
<thead>
<tr>
<th></th>
<th>male</th>
<th>female</th>
</tr>
</thead>
<tbody>
<tr>
<td>taller than 6 ft</td>
<td>30</td>
<td>10</td>
</tr>
<tr>
<td>shorter than 6 ft</td>
<td>50</td>
<td>50</td>
</tr>
</tbody>
</table>
Population proportion

**Experiment → outcome → number**

**Example 2:** A male, B tall.

\[
P(A) = \frac{|A|}{|\Omega|} = \frac{50}{100} = 50\%.
\]

\[
P(B) = \frac{|B|}{|\Omega|} = \frac{30 + 10}{100} = 40\%.
\]

*Probability = population proportion.*
Population proportion

**Experiment → outcome → number**

**Example 2:** A male, B tall.

<table>
<thead>
<tr>
<th></th>
<th>male</th>
<th>female</th>
</tr>
</thead>
<tbody>
<tr>
<td>taller than 6 ft</td>
<td>30</td>
<td>10</td>
</tr>
<tr>
<td>shorter than 6 ft</td>
<td>50</td>
<td>50</td>
</tr>
</tbody>
</table>

\[
P(A|B) = \frac{|A \cap B|}{|B|} = \frac{30}{40} = 75\%.
\]

Among tall people, what is the proportion of males?

\[
P(B|A) = \frac{|A \cap B|}{|A|} = \frac{30}{50} = 60\%.
\]

Among males, what is the proportion of tall people?

**Conditional probability = proportion within sub-population.**
Example 2: A male, B tall.
Let \( \omega \in \Omega \) be a person. Let \( X(\omega) \) be the gender of \( \omega \), so that \( X(\omega) = 1 \) if \( \omega \) is male, and \( X(\omega) = 0 \) if \( \omega \) is female. Let \( Y(\omega) \) be the height of \( \omega \). Then

\[
A = \{ \omega : X(\omega) = 1 \}, \quad B = \{ \omega : Y(\omega) > 6 \}.
\]

\[
P(A) = P(\{ \omega : X(\omega) = 1 \}) = P(X = 1).
\]

\[
P(B) = P(\{ \omega : Y(\omega) > 6 \}) = P(Y > 6).
\]

\[
P(B|A) = P(Y > 6|X = 1), \quad P(A|B) = P(X = 1|Y > 6).
\]

Link between event and random variable.
Equally likely scenario

\[ P(A) = \frac{|A|}{|\Omega|}. \]

Axiom 0.
Can always translate a problem into equally likely setting.
Equally likely scenario

\[ P(A|B) = \frac{|A \cap B|}{|B|} = \frac{|A \cap B|/|\Omega|}{|B|/|\Omega|} = \frac{P(A \cap B)}{P(B)}. \]

As if \( B \) is the sample space.
Axiom 4
Or definition of conditional probability.
Random point

(1) $X$ is uniform random number in $[0, 1]$.
(2) $(X, Y)$ are two independent random numbers in $[0, 1]$.
(3) $(X, Y, Z)$ are three independent random numbers in $[0, 1]$. 
Random point
Example 3: throwing point into region

\(X\) and \(Y\) are independent uniform random numbers in \([0, 1]\). 
\((X, Y)\) is a random point in \(\Omega = [0, 1]^2\).

\[A = \{(x, y) : x^2 + y^2 \leq 1\}\].

\[P(A) = \frac{|A|}{|\Omega|} = \frac{\pi}{4}.

|\(A|\) is the size of \(A\), e.g., area (length, volume).
Example 3: throwing point into region

$X$ and $Y$ are independent uniform random numbers in $[0, 1]$. $(X, Y)$ is a random point in $\Omega = [0, 1]^2$.

$A = \{(x, y) : x^2 + y^2 \leq 1\}$.

$$P(X^2 + Y^2 \leq 1) = \pi/4.$$  

$$P(X^2 + Y^2 = 1) = 0.$$  

Capital letters for random variables.
Inner measure: fill inside by small squares $\rightarrow$ upper limit.  
Outer measure: cover outside by small squares $\rightarrow$ lower limit.  
Measurable: inner measure $=$ outer measure.  
The collection of all measurable sets, $\sigma$-algebra.  
Integral: area under curve.
Axioms

**Probability as measure**

Axiom 1: \( P(\Omega) = 1. \)

Axiom 2: \( P(A) \geq 0. \)

Axiom 3: If \( A \cap B = \phi \) (empty), then

\[
P(A \cup B) = P(A) + P(B).
\]
Example 3: $\pi$

Let $\Omega$ be a unit disk centered at the origin. We want to find the area of a quarter of the circle with radius 1, which is $\frac{\pi}{4}$.

1. **Throw $n$ points into $\Omega$.** $m$ of them fall into $A$.

   $$P(A) \approx \frac{m}{n}.$$ 

2. **Monte Carlo method:**

   $$\hat{\pi} = \frac{4m}{n}.$$ 

   For fixed $n$, $m$ is random.

   As $n \to \infty$, $\frac{m}{n} \to P(A)$ in probability.

   $P(A)$ can be interpreted as long run frequency.
Monte Carlo

Deterministic method

![Grid diagram]

Go over all the \( n = 100 = 10^2 \) square cells, count inner or outer measure, i.e., how many \((m)\) fall into \(A\).

3-dimensional? \( n = 10^3 \) cubic cells.

4-dimensional? \( n = 10^4 \) cells.

10000-dimensional? \( n = 10^{10000} \) cells.

**Monte Carlo**: sample \( n = 1000 \) points in the hyper-cube. Count how many \((m)\) fall into \(A\).
Example 3: $\pi$, buffon needle

Lazzarini threw $n = 3408$ times.

$$P(A) \approx \frac{m}{n}.$$  

Monte Carlo method:

$$\hat{\pi} = \frac{355}{113}$$

Too accurate. $m$ is random.

For fixed $n$, $m$ is random. $m/n$ fluctuates around $P(A)$.

As $n \to \infty$, $\frac{m}{n} \to P(A)$ in probability, law of large number.

$P(A)$ can be interpreted as long run frequency, how often $A$ happens in the long run.
Example 3: throwing point into region

$X$ and $Y$ are independent uniform random numbers in $[0, 1]$. $(X, Y)$ is a random point in $\Omega = [0, 1]^2$.

$A = \{(x, y) : x < 1/2\}$.

$$P(A) = P(X < 1/2) = \frac{|A|}{|\Omega|} = 1/2.$$
Example 3: throwing point into region

$X$ and $Y$ are independent uniform random numbers in $[0, 1]$. 
$(X, Y)$ is a random point in $\Omega = [0, 1]^2$. 
$B = \{(x, y) : x + y < 1\}$.

$$P(B) = P(X + Y < 1) = \frac{|B|}{|\Omega|} = 1/2.$$
Example 3: throwing point into region

Consider throwing a lot of points into $\Omega$. How often $A$ happens? How often $B$ happens? When $B$ happens, how often $A$ happens? Among all the points in $B$, what is the fraction belongs to $A$?

Let $A$ and $B$ be two regions in the plane. Consider the event $A \cap B$ and the event $B$. We want to find the conditional probability $P(A|B)$.

Using the definition of conditional probability,

$$P(A|B) = \frac{|A \cap B|}{|B|} = \frac{1/2 - 1/8}{1/2} = \frac{3}{4}.$$
**Coin flipping**

**Experiment → outcome → number**

**EXAMPLE 4: Coin flipping**

(4.1) Flip a coin → head or tail → 1 or 0
(4.2) Flip a coin twice → (head, head), or (head, tail), or (tail, head) or (tail, tail) → 11 or 10 or 01 or 00

![List and Tree Diagram](image)

The sample space is \{HH, HT, TH, TT\}
Coin flipping

Experiment $\rightarrow$ outcome $\rightarrow$ number

Experiment 4: Coin flipping

(4.3) Flip a coin $n$ times $\rightarrow 2^n$ binary sequences.

Sample space $\Omega$: all $2^n$ sequences.
Each $\omega \in \Omega$ is a sequence.
$Z_i(\omega) = 1$ if $i$-th flip is head; $Z_i(\omega) = 0$ if $i$-th flip is tail.
Experiment 4: Coin flipping

\[ Z_i(\omega) = 1 \text{ if } i\text{-th flip is head}; \ Z_i(\omega) = 0 \text{ if } i\text{-th flip is tail.} \]

\[
\begin{align*}
\text{HHHH, THHH, HHTT, TTHT,} \\
\text{HHHT, HHTT, THHT, THTT,} \\
\text{HHTH, TTHH, HTTH, HTTT,} \\
\text{HTHH, THTH, TTTH, TTTT}
\end{align*}
\]

Flip a fair coin 4 times independently, let \( A \) be the event that there are 2 heads.

\[
P(A) = \frac{|A|}{|\Omega|} = \frac{6}{2^4} = \frac{3}{8}.
\]

\[ A = \{ \omega : Z_1(\omega) + Z_2(\omega) + Z_3(\omega) + Z_4(\omega) = 2 \}. \]
Experiment 4: Coin flipping

$Z_i(\omega) = 1$ if $i$-th flip is head; $Z_i(\omega) = 0$ if $i$-th flip is tail.

Let $X(\omega)$ be the number of heads in the sequence $\omega$.

$$X(\omega) = Z_1(\omega) + Z_2(\omega) + Z_3(\omega) + Z_4(\omega).$$

$$P(A_k) = P(\{\omega : X(\omega) = k\}) = P(X = k) = p_k.$$  

$$(p_k, k = 0, 1, 2, 3, 4) = (1, 4, 6, 4, 1)/16.$$
Experiment 4: Coin flipping

HHHH, THHH, HTHT, THTT, 
HHHT, HHTT, THHT, THTT, 
HHTH, TTHH, HTTH, HTTT, 
HTHH, THTH, TTTH, TTTT

\[ |A_2| = 6. \]
\[ |A_2| = \binom{4}{2} = \frac{4 \times 3}{2}. \]

4 positions, choose 2 of them to be heads, and the rest are tails.
Ordered pair: roll a die twice

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(1,1)</td>
<td>(1,2)</td>
<td>(1,3)</td>
<td>(1,4)</td>
<td>(1,5)</td>
<td>(1,6)</td>
</tr>
<tr>
<td>2</td>
<td>(2,1)</td>
<td>(2,2)</td>
<td>(2,3)</td>
<td>(2,4)</td>
<td>(2,5)</td>
<td>(2,6)</td>
</tr>
<tr>
<td>3</td>
<td>(3,1)</td>
<td>(3,2)</td>
<td>(3,3)</td>
<td>(3,4)</td>
<td>(3,5)</td>
<td>(3,6)</td>
</tr>
<tr>
<td>4</td>
<td>(4,1)</td>
<td>(4,2)</td>
<td>(4,3)</td>
<td>(4,4)</td>
<td>(4,5)</td>
<td>(4,6)</td>
</tr>
<tr>
<td>5</td>
<td>(5,1)</td>
<td>(5,2)</td>
<td>(5,3)</td>
<td>(5,4)</td>
<td>(5,5)</td>
<td>(5,6)</td>
</tr>
<tr>
<td>6</td>
<td>(6,1)</td>
<td>(6,2)</td>
<td>(6,3)</td>
<td>(6,4)</td>
<td>(6,5)</td>
<td>(6,6)</td>
</tr>
</tbody>
</table>

Experiment 1 has $n_1$ outcomes. For each outcome of experiment 1, experiment 2 has $n_2$ outcomes. The number of all possible pairs is $n_1 \times n_2$. 
**Multiplication**

Ordered pair: roll a die twice

![Tree Diagram]

<table>
<thead>
<tr>
<th>Coin</th>
<th>Dice</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head</td>
<td>1</td>
<td>(H, 1)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>(H, 2)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>(H, 3)</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>(H, 4)</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>(H, 5)</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>(H, 6)</td>
</tr>
<tr>
<td>Tail</td>
<td>1</td>
<td>(T, 1)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>(T, 2)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>(T, 3)</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>(T, 4)</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>(T, 5)</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>(T, 6)</td>
</tr>
</tbody>
</table>
Multiplication

Ordered triplet
Permutation

$n$ different cards. Choose $k$ of them. Order matters. Number of different sequences:

$$P_{n,k} = n(n-1)...(n-k+1).$$

$$P_{4,2} = 4 	imes 3 = 12.$$  \[ P_{n,n} = n!. \]

How many different ways to permute things.
Combination

$n$ different balls. Choose $k$ of them. Order does NOT matters. Number of different combinations:

\[
\binom{n}{k} = \frac{P_{n,k}}{k!} = \frac{n(n-1)\ldots(n-k+1)}{k!} = \frac{n!}{k!(n-k)!}.
\]

\[
\binom{4}{2} = \frac{4 \times 3}{2} = 6.
\]
Each combination corresponds to $k!$ permutations.

\[
\binom{n}{k} = \frac{P_{n,k}}{k!} = \frac{n(n-1)...(n-k+1)}{k!} = \frac{n!}{k!(n-k)!}.
\]

\[
\binom{4}{2} = \frac{4 \times 3}{2} = 6.
\]
Experiment 4: Coin flipping

- HHHH, THHH, HTHT, TTHT,
- HHHT, HHTT, THHT, THTT,
- HHTH, TTHH, HTTH, HTTT,
- HTTH, THTH, TTHH, TTTT

\[ |A_2| = 6. \]
\[ |A_2| = \binom{4}{2} = \frac{4 \times 3}{2}. \]

\[ P(A_k) = P(\{\omega : X(\omega) = k\}) = P(X = k) = \binom{n}{k}/2^n. \]

\[ X(\omega) = \sum_{i=1}^{n} Z_i(\omega). \]
Experiment 4: Coin flipping

\[ A = \{ \omega : x(\omega) = 2 \} \]

\[ 1 \quad 2 \quad 3 \quad 4 \]
\[ \text{H} \quad \text{T} \quad \text{T} \quad \text{T} \]

\[ |A_2| = 6. \]
\[ |A_2| = \binom{4}{2} = \frac{4 \times 3}{2}. \]
Experiment 4: Coin flipping

Combinations

\[
\begin{array}{cccc}
0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 \\
2 & 2 & 2 & 2 \\
3 & 3 & 3 & 3 \\
4 & 4 & 4 & 4 \\
5 & 5 & 5 & 5 \\
\end{array}
\]

- H H H H 4 heads
- H H H T 3 heads
- H H T H 3 heads
- H H T T 3 heads
- T H H H 3 heads
- T H H T 2 heads
- T H T H 2 heads
- T H T T 2 heads
- T T H H 2 heads
- T T H T 2 heads
- T T T H 2 heads
- T T T T 1 head
- H T T T 1 head
- T T H T 1 head
- T T T H 1 head
- T T T T 0 head
Experiment 4: Coin flipping

Combinations

Un-ordered sequence in each circle.
Experiment 4: Coin flipping

Pascal triangle

\[ \begin{array}{c|c|c|c|c|c} n & H & H & H & H & 4 \text{ heads} \\
0 & H & H & H & T & 3 \text{ heads} \\
1 & H & T & H & H & 3 \text{ heads} \\
1 & H & H & T & H & 3 \text{ heads} \\
2 & T & H & H & H & 3 \text{ heads} \\
2 & H & T & H & T & 2 \text{ heads} \\
3 & H & T & T & H & 2 \text{ heads} \\
3 & T & H & H & T & 2 \text{ heads} \\
4 & T & H & T & H & 2 \text{ heads} \\
4 & T & T & H & H & 2 \text{ heads} \\
5 & H & T & T & T & 1 \text{ heads} \\
5 & T & H & T & T & 1 \text{ heads} \\
6 & T & T & H & T & 1 \text{ heads} \\
6 & T & T & T & T & 0 \text{ heads} \\
\end{array} \]
Experiment 4: Coin flipping

Binomial

\[(x + y)^0 = 1\]
\[(x + y)^1 = 1x + 1y\]
\[(x + y)^2 = 1x^2 + 2x^1y^1 + 1y^2\]
\[(x + y)^3 = 1x^3 + 3x^2y^1 + 3x^1y^2 + 1y^3\]
\[(x + y)^4 = 1x^4 + 4x^3y^1 + 6x^2y^2 + 4x^1y^3 + 1y^4\]
\[(x + y)^5 = 1x^5 + 5x^4y^1 + 10x^3y^2 + 10x^2y^3 + 5x^1y^4 + 1y^5\]

\( (x + y)(x + y)(x + y)(x + y)\): each \((x + y)\) contributes either \(x\) or \(y\), \(2^n\) products, \(\binom{n}{k}\) of them have \(k\) \(x\)’s and \(n - k\) \(y\)’s.
Experiment 4: Coin flipping
Binomial

\[(a+b)^1 = a + b\]

\[(a+b)^2 = a^2 + 2ab + b^2\]

\[(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3\]

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Heads</th>
</tr>
</thead>
<tbody>
<tr>
<td>H H H H</td>
<td>4</td>
</tr>
<tr>
<td>H H H T</td>
<td>3</td>
</tr>
<tr>
<td>H T H H</td>
<td>3</td>
</tr>
<tr>
<td>H H T H</td>
<td>3</td>
</tr>
<tr>
<td>T H H H</td>
<td>3</td>
</tr>
<tr>
<td>H H T T</td>
<td>2</td>
</tr>
<tr>
<td>H T H T</td>
<td>2</td>
</tr>
<tr>
<td>H T T H</td>
<td>2</td>
</tr>
<tr>
<td>T H H T</td>
<td>2</td>
</tr>
<tr>
<td>T H T H</td>
<td>2</td>
</tr>
<tr>
<td>T T H T</td>
<td>2</td>
</tr>
<tr>
<td>H T T T</td>
<td>1</td>
</tr>
<tr>
<td>T H T T</td>
<td>1</td>
</tr>
<tr>
<td>T T H T</td>
<td>1</td>
</tr>
<tr>
<td>T T T H</td>
<td>1</td>
</tr>
<tr>
<td>T T T T</td>
<td>0</td>
</tr>
</tbody>
</table>
Random walk

**Experiment 4: Coin flipping**

Order of sequence in each circle does not matter.
All $2^n$ paths are equally likely.
Counting the number of paths that end up in $k$-th bin.
Experiment 4: Coin flipping

All $2^n$ paths are equally likely.
Number of paths that end up in $k$-th bin = $\binom{n}{k}$.
$X$: destination. $P(X = k) = \binom{n}{k} / 2^n$.
How often the balls fall into $k$-th bin.
Random walk

Either go forward or backward
Random walk

Either go forward or backward

\[ X_t = Z_1 + Z_2 + \ldots + Z_t. \]

\( Z_k = 1 \) or \( -1 \) with probability \( 1/2 \) each.

Number of heads = \( k \), then random walk ends up at

\[ m = k - (t - k) = 2k - t, \ k = (m + t)/2. \]

\[ P(X_t = m) = \left( \frac{t}{(m + t)/2} \right)^t. \]
Either go forward or backward

\[ X_t = Z_1 + Z_2 + \ldots + Z_t. \]

\[ Z_k = 1 \text{ or } -1 \text{ with probability } \frac{1}{2} \text{ each.} \]

\[ X_{t+1} = X_t + Z_{t+1}. \]

\[ P(X_{t+1} = x + 1 | X_t = x) = P(X_{t+1} = x - 1 | X_t = x) = \frac{1}{2}. \]
Example 5: Random walk over three states

With probability 1/2, stay. With probability 1/4, go to either states.

\[ K_{ij} = P(X_{t+1} = j | X_t = i). \]

Markov property: past history before \( X_t \) does not matter.
Example 5: Random walk over three states

With probability 1/2, stay. With probability 1/4, go to either of the other two states.

\[ K_{ij} = P(X_{t+1} = j | X_t = i). \]

Imagine 1 million people migrating. At each step, for each state, half of the people stay, 1/4 go to each of the other two states.
Example 5: Random walk over three states

With probability $1/2$, stay. With probability $1/4$, go to either of the other two states.

\[
K_{ij} = P(X_{t+1} = j | X_t = i).
\]

\[
K = \begin{bmatrix}
1/2 & 1/4 & 1/4 \\
1/4 & 1/2 & 1/4 \\
1/4 & 1/4 & 1/2
\end{bmatrix}
\]
Example 5: Random walk over three states

With probability 1/2, stay. With probability 1/4, go to either of the other two states.

\[ K_{ij} = P(X_{t+1} = j | X_t = i). \]

\[ p_i^{(t)} = P(X_t = i). \]

Imagine 1 million people migrating. \( p_i^{(t)} \) is the number of people (in million) in state \( i \) at time \( t \).

\[ p^{(t)} = (p_1^{(t)}, p_2^{(t)}, p_3^{(t)}). \]
Example 5: Random walk over three states

Imagine 1 million people migrating. \( p_i^{(t)} \) is the number of people (in million) in state \( i \) at time \( t \).

\[
p_i^{(t)} = P(X_t = i).
\]

\[
p^{(t)} = (p_1^{(t)}, p_2^{(t)}, p_3^{(t)}).
\]
Population migration

**Example 5: Random walk over three states**

\[
K_{ij} = P(X_{t+1} = j \mid X_t = i).
\]

\[
p_i^{(t)} = P(X_t = i).
\]

\[
p_j^{(t+1)} = \sum_i p_i^{(t)} K_{ij}.
\]

Number of people in state \( j \) at time \( t + 1 \) = sum number of people in state \( i \) at time \( t \) \( \times \) fraction of those in \( i \) who go to \( j \).
Example 5: Random walk over three states

\[ p^{(t+1)} = \sum_i p^{(t)} K_{ij}. \]

\[ p^{(t)} \rightarrow \pi_i. \]

\[ \pi_j = \sum_i \pi_i K_{ij}. \]

Stationary distribution.
Example 5: Random walk over three states

\[ p_j^{(t+1)} = \sum_i p_i^{(t)} K_{ij}. \]

\[ p^{(t+1)} = p^{(t)} K. \]

\[ p^{(t)} = p^{(0)} K^t \to \pi. \]
Example 5: Random walk over three states

\[ p_j^{(t+1)} = \sum_i p_i^{(t)} K_{ij}. \]

\[ p_i^{(t)} \rightarrow \pi_i. \]

\[ \pi_j = \sum_i \pi_i K_{ij}. \]

\( \pi_i \): proportion of people who are in page \( i \).

Popularity of \( i \) depends on the popularities of pages linked to \( i \).
Chain rule and rule of total probability

Example 5: Random walk over three states

\[ P(A|B) = \frac{P(A \cap B)}{P(B)}. \]

\[ P(A \cap B) = P(B)P(A|B). \]

\[ P(X_{t+1} = j \cap X_t = i) = P(X_t = i)P(X_{t+1} = j|X_t = i) = p_{ij}^{(t)} K_{ij}. \]

\[ P(X_{t+1} = j) = \sum_i P(X_{t+1} = j \cap X_t = i). \]

\[ p_{ij}^{(t+1)} = \sum_i p_{ij}^{(t)} K_{ij}. \]

Add up probabilities of alternative chains of events.
Independence

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

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

**Independence**

$$P(A|B) = P(A).$$

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

$A$ and $B$ have nothing to do with each other.
Independence

\[ P(A | B) = P(A). \]

\[ P(A \cap B) = P(A)P(B). \]
Conditional independence

**Markov chain:**  \( C \rightarrow B \rightarrow A, \)

\[
\]

\[
P(X_{t+1} = j|X_t = i, X_{t-1}, \ldots, X_0) = P(X_{t+1} = j|X_t = i).
\]

Future is independent of the past given present.
Immediate cause (parent), remote cause (grandparent).

**Shared cause:**  \( C \leftarrow B \rightarrow A, \)

\[
P(A \cap C|B) = P(A|B)P(C|B).
\]

Children given parent.
Example 6: Rare disease example
1% of population has a rare disease.
A random person goes through a test.
If the person has disease, 90% chance test positive.
If the person does not have disease, 90% chance test negative.
If tested positive, what is the chance he or she has disease?

\[ P(D) = 1\% . \]

\[ P(+|D) = 90\% , \quad P(-|N) = 90\% . \]

\[ P(D|+) = ? \]
Example 6: Rare disease example

\[ P(D) = 1\% . \]
\[ P(+)|D) = 90\% , \ P(−|N) = 90\% . \]
\[ P(D|+) = ? \]

\[ P(+)|D) = \frac{9}{9+99} = \frac{1}{12} . \]

\[ P(alarm|fire) \ vs \ P(fire|alarm) . \]
**Example 6: Rare disease example**

\[ P(D) = 1\% \]
\[ P(+) | D) = 90\%, \ P(− | N) = 90\% . \]

\[ P(D \cap +) = P(D)P(+) | D) = 1\% \times 90\%. \]
\[ P(N \cap +) = P(N)P(+) | N) = 99\% \times 10\%. \]
\[ P(+) = P(D \cap +) + P(N \cap +) = 1\% \times 90\% + 99\% \times 10\%. \]
\[ P(D | +) = \frac{P(D \cap +)}{P(+)} = \frac{9}{9+99} = \frac{1}{12}. \]
General formula

$m$ causes: $C_1, ..., C_i, ..., C_m$.
$n$ effects: $E_1, ..., E_j, ..., E_n$.

Given:
Prior: $P(C_i), i = 1, ..., m$.
Conditional: $P(E_j|C_i), j = 1, ..., n$. 

Diagram:
- Causes: $C_1, C_2, C_3, C_4$
- Effects: $E_1, E_2, E_3$
- Inference/Reasoning
  - Causal direction:
    - Cause
    - Effect
Chain rule, rule of total probability, Bayes rule

Prior: \( P(C_i), i = 1, ..., m. \)
Conditional: \( P(E_j|C_i), j = 1, ..., n. \)

\[
P(C_i | E) = \frac{P(C_i \cap E)}{P(E)} = \frac{P(C_i) P(E|C_i)}{\sum_{i'} P(C_{i'}) P(E|C_{i'})}.
\]

\[
P(C_i) = \frac{P(C_i \cap E)}{P(E)} = \frac{P(C_i) P(E|C_i)}{\sum_{i'} P(C_{i'}) P(E|C_{i'})}.
\]

\[
P(E|C_i) = \frac{P(E \cap C_i)}{P(C_i)} = \frac{P(C_i) P(E|C_i)}{\sum_{i'} P(C_{i'}) P(E|C_{i'})}.
\]
Chain rule, rule of total probability, Bayes rule

\[ X = \text{cause} \in \{1, \ldots, i, \ldots, m\}. \]
\[ Y = \text{effect} \in \{1, \ldots, j, \ldots, n\}. \]

\[
P(X = i | Y = j) = \frac{P(X = i \cap Y = j)}{P(Y = j)}
\]
\[
= \frac{P(X = i)}{\sum_{i'} P(X = i') P(Y = j | X = i')}
\]
Chain rule, rule of total probability, Bayes rule

\[ P(X = i|Y = j) = \frac{P(X = i \cap Y = j)}{P(Y = j)} \]

\[ = \frac{P(X = i)P(Y = j|X = i)}{\sum_{i'} P(X = i')P(Y = j|X = i')} \]

\[ p(x|y) = \frac{p(x, y)}{p(y)} = \frac{p(x)p(y|x)}{\sum_{x'} p(x')p(y|x')} \]
**Bayes network**, directed acyclic graph, graphic model

```
Been to Asia
   ↓
Tuberculosis
   ↓
Short of breath
```
```
Smoking
   ↓
Lung cancer
```
```
Bronchitis
   ↓
X-ray
```

**Conditional independence:**
Sibling nodes are independent given parent node. Child node is independent of grandparents given parent.
Take home message

As long as you can count
Count the number of people
Count the number of points or repetitions
Count the number of sequences
Two things
Intuition and visualization (and motivation)
Precise notation and formula
Accomplished
Most of the important concepts via intuitive examples
Next
Systematic and more in-depth treatments