3. Independence and Random Variables
Independence of two events

Let $E_1$ be “first coin comes up $\text{H}$”

$E_2$ be “second coin comes up $\text{H}$”

Then $P(E_2 \mid E_1) = P(E_2)$

$P(E_2 \cap E_1) = P(E_2)P(E_1)$

Events $A$ and $B$ are independent if

$P(A \cap B) = P(A)P(B)$
Examples of (in)dependence

Let $E_1$ be “first die is a 4”
$S_6$ be “sum of dice is a 6”
$S_7$ be “sum of dice is a 7”

$E_1$ and $S_6$?
$\frac{1}{36} \neq \frac{1}{6} \cdot \frac{1}{6}$ \text{ NOT IND.}

$E_1$ and $S_7$?
$\frac{1}{36} = \frac{1}{6} \cdot \frac{1}{6}$ \text{ IND.}

$S_6$ and $S_7$?
$0 \neq 0.00$
**ER:** “East Rail Line is working” \[ P(ER) = 70\% \]

**MS:** “Ma On Shan Line is working” \[ P(MS) = 98\% \]

\[ P(ER \cap MS) = P(ER)P(MS) = 0.686 \]
Algebra of independent events

If $A$ and $B$ are independent, then $A$ and $B^c$ are also independent.

Proof:

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

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

$$= P(A) \cdot P(B^c)$$

$$= P(A \setminus B^c)$$
TW: “Tsuen Wan Line is operational”

TC: “Tung Chung Line is operational”

\[ P(TW) = 80\% \]

\[ P(TC) = 85\% \]

\[
P(IN\text{ UTC}) = 1 - P(TW^c \cap TC^c)
= 1 - P(TW^c) P(TC^c)
= 1 - (1 - P(TW))(1 - P(TC))
= 1 - 20\% \cdot 15\%
= 0.97
\]
Events $A$, $B$, and $C$ are independent if

\[ P(A \cap B) = P(A) \, P(B) \]
\[ P(B \cap C) = P(B) \, P(C) \]
\[ P(A \cap C) = P(A) \, P(C) \]

and \[ P(A \cap B \cap C) = P(A) \, P(B) \, P(C). \]
(In)dependence of three events

Let $E_1$ be “first die is a 4”
$E_2$ be “second die is a 3”
$S_7$ be “sum of dice is a 7”

$E_1, E_2\,?$
\[
\frac{1}{6} \times \frac{1}{6} = \frac{1}{36} \quad \checkmark
\]

$E_1, S_7\,?$
\[
\frac{1}{6} \times \frac{1}{6} = \frac{1}{36} \quad \checkmark
\]

$E_2, S_7\,?$
\[
\frac{1}{6} \times \frac{1}{6} = \frac{1}{36} \quad \checkmark
\]

$E_1, E_2, S_7\,?$
\[
P(E_1 \cap E_2 \cap S_7) = \frac{1}{36} \times
\]
(In)dependence of three events

Let $A$ be “first roll is 1, 2, or 3 ” \( \frac{1}{2} \)

$B$ be “first roll is 3, 4, or 5” \( \frac{1}{2} \)

$C$ be “sum of rolls is 9” \( \frac{1}{9} \)

$A, B$? \( \frac{1}{2} \times \frac{1}{2} = \frac{1}{6} \)

$A, C$? \( \frac{1}{2} \times \frac{1}{9} \neq \frac{1}{12} \)

$B, C$? \( \frac{1}{2} \times \frac{1}{9} \neq \frac{1}{12} \)

$A, B, C$? \( \frac{1}{2} \times \frac{1}{2} \times \frac{1}{9} = \frac{1}{36} \)
Independence of many events

Events $A_1, A_2, \ldots$ are independent if for every subset of the events, the probability of the intersection is the product of their probabilities.

Algebra of independent events

Independence is preserved if we replace some event(s) by their complements, intersections, unions.
Multiple components

\[ P(ER) = 70\% \]

\[ P(WR) = 75\% \]

\[ P(TW) = 85\% \]

\[ P(KT) = 95\% \]

\[ E = ER \cap (WR \cup (KT \cap TW)) \]
Multiple components

\[ P(ER) = 70\% \]
\[ P(WR) = 75\% \]
\[ P(KT) = 95\% \]
\[ P(TW) = 85\% \]

\[
P(E) = P(ER \cap (WR \cup (KT \cap TW)))
\]

\[
= P(ER) \cdot P(WR \cup (KT \cap TW))
\]

\[
P(WR \cup (KT \cap TW)) = 1 - (1 - P(WR)) \cdot \frac{25\%}{70\%} \cdot \frac{80\%}{80\%}
\]

\[
\approx 95\%
\]

\[
P(E) \approx 70\% \cdot 95\% \approx 67\%
\]
Alice wins 60% of her ping pong matches against Bob. They meet for a 3 match playoff. What are the chances that Alice will win the playoff?

**Probability model**

Let $A_i$ be the event Alice wins match $i$

Assume $P(A_1) = P(A_2) = P(A_3) = 0.6$

Also assume $A_1, A_2, A_3$ are independent
## Playoffs

<table>
<thead>
<tr>
<th>outcome</th>
<th>probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>$0.6^3$</td>
</tr>
<tr>
<td>AAB</td>
<td>$0.6^2 \cdot 0.4$</td>
</tr>
<tr>
<td>ABA</td>
<td>$0.6^2 \cdot 0.4$</td>
</tr>
<tr>
<td>BAA</td>
<td>$0.6^2 \cdot 0.4$</td>
</tr>
</tbody>
</table>

$$P(A) = 0.6^3 + 3 \cdot 0.6^2 \cdot 0.4 = 0.648$$
Bernoulli trials

$n$ trials, each succeeds independently with probability $p$

The probability at least $k$ out of $n$ succeed is

\[
\binom{n}{k} p^k (1-p)^{n-k} + \binom{n}{k+1} p^{k+1} (1-p)^{n-k-1} + \ldots + \binom{n}{n} p^n
\]
The probability that Alice wins an $n$ game tournament
The Lakers and the Celtics meet for a 7-game playoff. They play until one team wins four games.

Suppose the Lakers win 60% of the time. What is the probability that all 7 games are played?
ALL 7 PLAYED

FIRST 6 HAVE 3 LAKERS WINS
3 LAKERS LOSSES

\[ P(E) = \binom{6}{3} 0.6^3 \cdot 0.4^3 \]
A and B are independent conditioned on F if

\[ P(A \cap B \mid F) = P(A \mid F) \cdot P(B \mid F) \]

Alternative definition:

\[ P(A \mid B \cap F) = P(A \mid F) \]
<table>
<thead>
<tr>
<th>today</th>
<th>tomorrow</th>
</tr>
</thead>
<tbody>
<tr>
<td>☀</td>
<td>80% ☀, 20% ⛈</td>
</tr>
<tr>
<td>⛈</td>
<td>40% ☀, 60% ⛈</td>
</tr>
</tbody>
</table>

It is ☀ on Monday. Will it ⛈ on Wednesday?

\[
P(W|\text{MT}) = P(W|T)
\]

\[
P(T) = P(T|\text{M})P(M) + P(T|\text{MC})P(M^c) = 0.8
\]

\[
P(W) = P(W|T)P(T) + P(W|T^c)P(T^c) = 0.72
\]
Conditioning does not preserve independence

Let $E_1$ be “first die is a 4”
$E_2$ be “second die is a 3”
$S_7$ be “sum of dice is a 7”

$E_1, E_2$ independent but

$P(E_1 \cap E_2 \mid S_7) \neq P(E_1 \mid S_7)P(E_2 \mid S_7)$

$\frac{1}{6} \quad \frac{1}{6} \quad \frac{1}{6}$
‘Crazy Rich Asians’ Has Soared, but It May Not Fly in China
Conditioning may destroy dependence

Probability model

A, B independent given C but

\[
P(A) = \frac{1}{2} \cdot 99\% + \frac{1}{2} \cdot 1\% = \frac{1}{2}
\]

\[
P(B) = \frac{1}{2} \cdot 99\% + \frac{1}{2} \cdot 1\% = \frac{1}{2}
\]

\[
P(A \mid B) = \frac{1}{2} \cdot 0.99^2 + \frac{1}{2} \cdot 0.01^2 
\approx 49\%.
\]
Random variable

A **discrete random variable** assigns a discrete value to every outcome in the sample space.

**Example**

\[ N = \text{number of } Hs \]
Probability mass function

The probability mass function (p.m.f.) of discrete random variable $X$ is the function

$$p(x) = P(X = x)$$

Example

$N =$ number of $\text{Hs}$

$\begin{align*}
\text{ } & \quad \{ \text{HH, HT, TH, TT} \} \\
\text{ } & \quad \begin{array}{cccc}
\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4}
\end{array}
\end{align*}$

$p(0) = P(N = 0) = P(\{TT\}) = 1/4$

$p(1) = P(N = 1) = P(\{HT, TH\}) = 1/2$

$p(2) = P(N = 2) = P(\{HH\}) = 1/4$
Probability mass function

We can describe the p.m.f. by a table or by a chart.

<table>
<thead>
<tr>
<th>x</th>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>p(x)</td>
<td>$\frac{1}{4}$</td>
<td>$\frac{1}{2}$</td>
<td>$\frac{1}{4}$</td>
</tr>
</tbody>
</table>
Two six-sided dice are tossed. Calculate the p.m.f. of the difference $D$ of the rolls.

What is the probability that $D > 1$? $D$ is odd?
<p>| | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>32</td>
<td>33</td>
<td>34</td>
<td>35</td>
<td>36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>42</td>
<td>43</td>
<td>44</td>
<td>45</td>
<td>46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>52</td>
<td>53</td>
<td>54</td>
<td>55</td>
<td>56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>61</td>
<td>62</td>
<td>63</td>
<td>64</td>
<td>65</td>
<td>66</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[ d \quad P(D=d) \]
\[
\begin{align*}
-5 & : \frac{1}{36} \\
-4 & : \frac{2}{36} \\
-3 & : \frac{3}{36} \\
-2 & : \frac{4}{36} \\
-1 & : \frac{5}{36} \\
0 & : \frac{6}{36} \\
1 & : \frac{5}{36} \\
2 & : \frac{4}{36} \\
3 & : \frac{3}{36} \\
4 & : \frac{2}{36} \\
5 & : \frac{1}{36}
\end{align*}
\]
The binomial random variable

Binomial\((n, p)\): Perform \(n\) independent trials, each of which succeeds with probability \(p\).

\(X = \text{number of successes}\)

Examples

Toss \(n\) coins. “number of heads” is Binomial\((n, \frac{1}{2})\).

Toss \(n\) dice. “Number of \(\bullet\)s” is Binomial\((n, 1/6)\).
A less obvious example

Toss $n$ coins. Let $C$ be the number of consecutive changes ($HT$ or $TH$).

Examples:

<table>
<thead>
<tr>
<th>$\omega$</th>
<th>$C(\omega)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>HHHHHHHH</td>
<td>0</td>
</tr>
<tr>
<td>THHHHHT</td>
<td>2</td>
</tr>
<tr>
<td>HTHHHHT</td>
<td>3</td>
</tr>
</tbody>
</table>

Then $C$ is $\text{Binomial}(n-1, \frac{1}{2})$. 
A non-example

Draw a 10-card hand from a 52-card deck.
Let $N =$ number of aces among the drawn cards.

Is $N$ a Binomial$(10, 1/13)$ random variable?

No! Trial outcomes are not independent.
Probability mass function

If $X$ is $\text{Binomial}(n, p)$, its p.m.f. is

$$p(k) = P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$
Geometric random variable

Let $X_1, X_2, \ldots$ be independent trials with success $p$.

A Geometric($p$) random variable $N$ is the time of the first success among $X_1, X_2, \ldots$:

$$N = \text{first (smallest) } n \text{ such that } X_n = 1.$$ 

So $P(N = n) = P(X_1 = 0, \ldots, X_{n-1} = 0, X_n = 1)$

$$= (1 - p)^{n-1}p.$$ 

This is the p.m.f. of $N$. 
About 10% of the apples on your farm are rotten.

You sell 10 apples. How many are rotten?

**Probability model**

Number of rotten apples you sold is Binomial($n = 10, p = 1/10$).
Apples

You improve productivity; now only 5% apples rot.

You can now sell 20 apples.

$N$ is now $\text{Binomial}(20, 1/20)$. 
Binomial(10, 1/10)

- 0.349
- 0.194
- 0.001
- 10^{-10}

Binomial(20, 1/20)

- 0.354
- 0.189
- 0.002
- 10^{-8}
- 10^{-26}

Poisson(1)

- 0.367
- 0.183
- 0.003
- 10^{-7}
- 10^{-19}
The Poisson random variable

A Poisson(\(\lambda\)) random variable has this p.m.f.:

\[
p(k) = e^{-\lambda} \frac{\lambda^k}{k!} \quad k = 0, 1, 2, 3, \ldots
\]

Poisson random variables do not occur “naturally” in the sample spaces we have seen.

They approximate Binomial\((n, p)\) random variables when \(\lambda = np\) is fixed and \(n\) is large (so \(p\) is small)

\[
p_{\text{Poisson}(\lambda)}(k) = \lim_{n \to \infty} p_{\text{Binomial}(n, \lambda/n)}(k)
\]
Functions of random variables

If $X$ is a random variable with p.m.f. $p_X$, then $Y = f(X)$ is a random variable with p.m.f.

$$p_Y(y) = \sum_{x : f(x) = y} p_X(x).$$
Two six-sided dice are tossed. $D$ is the difference of rolls. Calculate the p.m.f. of $|D|$.

<table>
<thead>
<tr>
<th>$d$</th>
<th>$P(D=d)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-5</td>
<td>$\frac{1}{36}$</td>
</tr>
<tr>
<td>-4</td>
<td>$\frac{2}{36}$</td>
</tr>
<tr>
<td>-3</td>
<td>$\frac{5}{36}$</td>
</tr>
<tr>
<td>-2</td>
<td>$\frac{6}{36}$</td>
</tr>
<tr>
<td>-1</td>
<td>$\frac{5}{36}$</td>
</tr>
<tr>
<td>0</td>
<td>$\frac{6}{36}$</td>
</tr>
<tr>
<td>1</td>
<td>$\frac{1}{36}$</td>
</tr>
<tr>
<td>2</td>
<td>$\frac{10}{36}$</td>
</tr>
<tr>
<td>3</td>
<td>$\frac{8}{36}$</td>
</tr>
<tr>
<td>4</td>
<td>$\frac{6}{36}$</td>
</tr>
<tr>
<td>5</td>
<td>$\frac{4}{36}$</td>
</tr>
<tr>
<td>6</td>
<td>$\frac{2}{36}$</td>
</tr>
</tbody>
</table>

For $|D|$: $P(|D|=0)=\frac{6}{36}$, $P(|D|=1)=\frac{10}{36}$, $P(|D|=2)=\frac{8}{36}$, $P(|D|=3)=\frac{6}{36}$, $P(|D|=4)=\frac{4}{36}$, $P(|D|=5)=\frac{2}{36}$.