S&DS 241 Lecture 6

Functions of random variables. LOTUS rule.
Independence of random variables. Conditional independence.

B-H: 2.5, 3.7, 3.8, 4.5
Functions of random variables

Let $X$ be a discrete random variable. Let $g : \mathbb{R} \rightarrow \mathbb{R}$ be a function. Then $Y = g(X)$ is also a discrete random variable,
Let $X$ be a discrete random variable. Let $g : \mathbb{R} \to \mathbb{R}$ be a function. Then $Y = g(X)$ is also a discrete random variable, because

- $Y : \Omega \to \mathbb{R}$ is the composition of $X : \Omega \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$, which maps $\omega$ to $Y(\omega) = g(X(\omega))$

- the number of possible values of $Y$ is at most that of $X$
Find PMF of $Y = g(X)$

If we know the PMF of $X$, we can obtain the PMF of $Y$:

**Step 1** Find the set of values that $Y$ takes

**Step 2** Find the corresponding probabilities

$$p_Y(y) = P(g(X) = y) = \sum_{x:g(x)=y} p_X(x)$$
Example

Given the PMF of $X$

<table>
<thead>
<tr>
<th>$x$</th>
<th>$-2$</th>
<th>$-1$</th>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_X(x)$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
</tr>
</tbody>
</table>

Find the PMF of $Y = 2X + 1$: 
Example

Given the PMF of $X$

<table>
<thead>
<tr>
<th>$x$</th>
<th>$-2$</th>
<th>$-1$</th>
<th>$0$</th>
<th>$1$</th>
<th>$2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_X(x)$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
</tr>
</tbody>
</table>

Find the PMF of $Y = 2X + 1$:

**Step 1** $Y$ takes values $-3, -1, 1, 3, 5$

**Step 2** $P(Y = -3) = P(2X + 1 = -3) = P(X = -2) = 1/5$, and so on and so forth
Example

Given the PMF of $X$

\[
\begin{array}{c|c|c|c|c|c}
  x & -2 & -1 & 0 & 1 & 2 \\
p_X(x) & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \\
\end{array}
\]

Find the PMF of $Y = 2X + 1$:

Step 1 $Y$ takes values $-3, -1, 1, 3, 5$

Step 2 $P(Y = -3) = P(2X + 1 = -3) = P(X = -2) = \frac{1}{5}$, and so on and so forth

\[
\begin{array}{c|c|c|c|c|c}
y & -3 & -1 & 1 & 3 & 5 \\
p_Y(y) & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \\
\end{array}
\]
Example

Given the PMF of $X$

<table>
<thead>
<tr>
<th>$x$</th>
<th>$-2$</th>
<th>$-1$</th>
<th>$0$</th>
<th>$1$</th>
<th>$2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_X(x)$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
</tr>
</tbody>
</table>

Find the PMF of $Z = X^2$: 

\[ Z \text{ takes values } 0, 1, 4 \]

\[ P(Z = 4) = P(X^2 = 4) = P(X = 2) + P(X = -2) = \frac{2}{5} \]
Example

Given the PMF of $X$

<table>
<thead>
<tr>
<th>$x$</th>
<th>$-2$</th>
<th>$-1$</th>
<th>$0$</th>
<th>$1$</th>
<th>$2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_X(x)$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
</tr>
</tbody>
</table>

Find the PMF of $Z = X^2$:

Step 1  $Z$ takes values 0, 1, 4

Step 2  $P(Z = 4) = P(X^2 = 4) = P(X = 2) + P(X = -2) = \frac{2}{5}$
Example

Given the PMF of $X$

<table>
<thead>
<tr>
<th>$x$</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_X(x)$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
</tr>
</tbody>
</table>

Find the PMF of $Z = X^2$:

**Step 1**  $Z$ takes values 0, 1, 4

**Step 2**  $P(Z = 4) = P(X^2 = 4) = P(X = 2) + P(X = -2) = \frac{2}{5}$

<table>
<thead>
<tr>
<th>$z$</th>
<th>0</th>
<th>1</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_Z(z)$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{2}{5}$</td>
<td>$\frac{2}{5}$</td>
</tr>
</tbody>
</table>
Example

Given the PMF of $X$

<table>
<thead>
<tr>
<th>$x$</th>
<th>$-2$</th>
<th>$-1$</th>
<th>$0$</th>
<th>$1$</th>
<th>$2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_X(x)$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
</tr>
</tbody>
</table>

Find the PMF of $W = -X$: 

Note: $X$ and $W$ have the same distribution (PMF), but these two random variables are NOT the same! In fact, $P(X = -W) = 1$. 
Example

Given the PMF of $X$

<table>
<thead>
<tr>
<th>$x$</th>
<th>$-2$</th>
<th>$-1$</th>
<th>$0$</th>
<th>$1$</th>
<th>$2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_X(x)$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
</tr>
</tbody>
</table>

Find the PMF of $W = -X$:

<table>
<thead>
<tr>
<th>$w$</th>
<th>$-2$</th>
<th>$-1$</th>
<th>$0$</th>
<th>$1$</th>
<th>$2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_W(w)$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
</tr>
</tbody>
</table>
Example

Given the PMF of $X$

<table>
<thead>
<tr>
<th>$x$</th>
<th>$-2$</th>
<th>$-1$</th>
<th>$0$</th>
<th>$1$</th>
<th>$2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_X(x)$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
</tr>
</tbody>
</table>

Find the PMF of $W = -X$:

<table>
<thead>
<tr>
<th>$w$</th>
<th>$-2$</th>
<th>$-1$</th>
<th>$0$</th>
<th>$1$</th>
<th>$2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_W(w)$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
</tr>
</tbody>
</table>

Note: $X$ and $W$ have the same distribution (PMF), but these two random variables are NOT the same! In fact, $P(X = -W) = 1$. 
Compute $E(g(X))$
Compute $E(g(X))$

Example:

- PMF of $X$:

<table>
<thead>
<tr>
<th>$x$</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_X(x)$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
</tr>
</tbody>
</table>

- PMF of $Z = X^2$:

<table>
<thead>
<tr>
<th>$z$</th>
<th>0</th>
<th>1</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_Z(z)$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{2}{5}$</td>
<td>$\frac{2}{5}$</td>
</tr>
</tbody>
</table>
Compute $E(g(X))$

Example:

- **PMF of $X$:**

  \[ x \quad | \quad -2 \quad | \quad -1 \quad | \quad 0 \quad | \quad 1 \quad | \quad 2 \]
  \[ p_X(x) \quad | \quad \frac{1}{5} \quad | \quad \frac{1}{5} \quad | \quad \frac{1}{5} \quad | \quad \frac{1}{5} \quad | \quad \frac{1}{5} \]

- **PMF of $Z = X^2$:**

  \[ z \quad | \quad 0 \quad | \quad 1 \quad | \quad 4 \]
  \[ p_Z(z) \quad | \quad \frac{1}{5} \quad | \quad \frac{2}{5} \quad | \quad \frac{2}{5} \]

- **Expectation:**

  \[
  E(Z) = \frac{1}{5} \times 0 + \frac{2}{5} \times 1 + \frac{2}{5} \times 4 = 2
  \]
Compute $E(g(X))$

Example:

- PMF of $X$:
  
<table>
<thead>
<tr>
<th>$x$</th>
<th>$-2$</th>
<th>$-1$</th>
<th>$0$</th>
<th>$1$</th>
<th>$2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_X(x)$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
<td>$\frac{1}{5}$</td>
</tr>
</tbody>
</table>

- PMF of $Z = X^2$:
  
  | $z$ | $0$ | $1$ | $4$ |
  |-----|-----|-----|
  | $p_Z(z)$ | $\frac{1}{5}$ | $\frac{2}{5}$ | $\frac{2}{5}$ |

- Expectation:
  
  $$E(Z) = \frac{1}{5} \times 0 + \frac{2}{5} \times 1 + \frac{2}{5} \times 4 = 2$$

- Can we find $E(Z)$ without finding $p_Z$?
  
  $$E(X^2) = \frac{1}{5} \times (-2)^2 + \frac{1}{5} \times (-1)^2 + \frac{1}{5} \times 0^2 + \frac{1}{5} \times 1^2 + \frac{1}{5} \times 2^2 = 2$$

  Compare:
  
  $$E(X) = \frac{1}{5} \times (-2) + \frac{1}{5} \times (-1) + \frac{1}{5} \times 0 + \frac{1}{5} \times 1 + \frac{1}{5} \times 2 = 0.$$
Law Of The Unconscious Statistician (LOTUS)

Random variable $X$ takes values in $\mathcal{X}$. Then for any function $g$:

$$E(g(X)) = \sum_{x \in \mathcal{X}} p_X(x)g(x),$$

that is, average the function values weighted by the PMF of $X$. 

**Law Of The Unconscious Statistician (LOTUS)**

Random variable $X$ takes values in $\mathcal{X}$. Then for any function $g$:

$$E(g(X)) = \sum_{x \in \mathcal{X}} p_X(x)g(x),$$

that is, average the function values weighted by the PMF of $X$.

**Proof.**

Group the sum according to the value of $X$:

$$E(g(X)) = \sum_{\omega \in \Omega} P(\{\omega\}) g(X(\omega))$$

$$= \sum_{x \in \mathcal{X}} \sum_{\omega \in \Omega: X(\omega) = x} P(\{\omega\}) g(x)$$

$$= \sum_{x \in \mathcal{X}} \left( \sum_{\omega \in \Omega: X(\omega) = x} P(\{\omega\}) \right) g(x)$$

$$= \sum_{x \in \mathcal{X}} \left( \sum_{\omega \in \Omega: X(\omega) = x} \frac{P(\{\omega\})}{P(X=x)} \right) g(x)$$

$$= \sum_{x \in \mathcal{X}} p_X(x)g(x)$$
Summary

Utility of PMF: Find

- probability of events: $P(X \in I)$
- conditional probabilities: $P(X \in I | X \in J)$
- PMF of $g(X)$
- Expectation of $g(X)$
- ...

Independence of random variables
Independence of two random variables

Recall independence of two events:

\[ P(A \cap B) = P(A)P(B) \]
Independence of two random variables

Recall independence of two events:

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

**Definition**

Discrete random variables \( X \) and \( Y \) are **independent** if for any \( x, y \in \mathbb{R} \), \( \{ X = x \} \) and \( \{ Y = y \} \) are independent events, i.e.,

\[ P(X = x, Y = y) = P(X = x)P(Y = y). \]
Independence of two random variables

Recall independence of two events:

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

Definition

Discrete random variables \( X \) and \( Y \) are independent if for any \( x, y \in \mathbb{R} \), \( \{X = x\} \) and \( \{Y = y\} \) are independent events, i.e.,

\[ P(X = x, Y = y) = P(X = x)P(Y = y). \]

Equivalently: conditioning does not change PMF

\[ P(X = x|Y = y) = P(X = x) \]

provided that \( P(Y = y) > 0 \).
Consequences

$X$ and $Y$ independent $\implies$ any event involving $X$ is independent of that involving $Y$: for any $I, J \subset \mathbb{R}$,

$$P(X \in I, Y \in J) = P(X \in I)P(Y \in J)$$
Consequences

$X$ and $Y$ independent $\implies$ any event involving $X$ is independent of that involving $Y$: for any $I, J \subset \mathbb{R}$,

$$P(X \in I, Y \in J) = P(X \in I)P(Y \in J)$$

Proof.

$$P(X \in I, Y \in J) = \sum_{x \in I} \sum_{y \in J} P(X = x, Y = y)$$

$$= \sum_{x \in I} \sum_{y \in J} P(X = x)P(Y = y) \quad \text{independence}$$

$$= \left( \sum_{x \in I} P(X = x) \right) \left( \sum_{y \in J} P(Y = y) \right)$$

$$= P(X \in I) \cdot P(Y \in J)$$
Consequences

If $X$ and $Y$ are independent, then

$$E(XY) = E(X)E(Y)$$
Consequences

If $X$ and $Y$ are independent, then

$$E(XY) = E(X)E(Y)$$

Proof.

$$E(XY) = \sum_{x} \sum_{y} xyP(X = x, Y = y) = \sum_{x} \sum_{y} xyP(X = x)P(Y = y)$$

LOTUS independence

$$= \left( \sum_{x} xP(X = x) \right) \left( \sum_{y} yP(Y = y) \right)$$

$E(X)$ $E(Y)$
Summary

- For any $X$ and $Y$, $E(X + Y) = E(X) + E(Y)$

- For independent $X$ and $Y$, $E(XY) = E(X)E(Y)$
Consequences

If $X$ and $Y$ are independent, then

- $f(X)$ and $g(Y)$ are also independent
- Hence

$$E(f(X)g(Y)) = E(f(X))E(g(Y))$$
Example: fair die

Let $X$ be the result of a fair die. Let

$$Y = X \mod 2$$
$$Z = X \mod 3$$

Then

- $p_Y(0) = p_Y(1) = 1/2$
Example: fair die

Let $X$ be the result of a fair die. Let

\[ Y = X \mod 2 \]
\[ Z = X \mod 3 \]

Then

- $p_Y(0) = p_Y(1) = 1/2$
- $p_Z(0) = p_Z(1) = p_Z(2) = 1/3$
Example: fair die

Let $X$ be the result of a fair die. Let

$$Y = X \mod 2$$
$$Z = X \mod 3$$

Then

- $p_Y(0) = p_Y(1) = 1/2$
- $p_Z(0) = p_Z(1) = p_Z(2) = 1/3$
- Are $Y$ and $Z$ independent?
Example: fair die

Let $X$ be the result of a fair die. Let

$$Y = X \mod 2$$

$$Z = X \mod 3$$

Then

- $p_Y(0) = p_Y(1) = 1/2$
- $p_Z(0) = p_Z(1) = p_Z(2) = 1/3$
- Are $Y$ and $Z$ independent? Yes!
Example: fair die

Let $X$ be the result of a fair die. Let

$$Y = X \mod 2$$
$$Z = X \mod 3$$

Then

- $p_Y(0) = p_Y(1) = 1/2$
- $p_Z(0) = p_Z(1) = p_Z(2) = 1/3$
- Are $Y$ and $Z$ independent? Yes!

$$P(Y = 0, Z = 0) = P(X = 6) = 1/6$$
Example: fair die

Let $X$ be the result of a fair die. Let

$$Y = X \mod 2$$
$$Z = X \mod 3$$

Then

- $p_Y(0) = p_Y(1) = 1/2$
- $p_Z(0) = p_Z(1) = p_Z(2) = 1/3$
- Are $Y$ and $Z$ independent? Yes!

$$P(Y = 0, Z = 0) = P(X = 6) = 1/6$$
$$P(Y = 0, Z = 1) = P(X = 4) = 1/6$$
$$\ldots\ldots$$
$$P(Y = 1, Z = 2) = P(X = 5) = 1/6$$
Example

Are $X$ and $g(X)$ independent?
Example

Are $X$ and $g(X)$ independent? In general no, e.g.,

- $X = \pm 1$ equally likely, $Y = -X$
- $P(X = 1, Y = 1) = 0 \neq P(X = 1)P(Y = 1) = \frac{1}{4}$
Example

Are $X$ and $g(X)$ independent? In general no, e.g.,

- $X = \pm 1$ equally likely, $Y = -X$
- $P(X = 1, Y = 1) = 0 \neq P(X = 1)P(Y = 1) = \frac{1}{4}$

unless in trivial cases (e.g. $X$ is a constant or $g$ is a constant function)
Independence of multiple random variables

**Definition**

Discrete random variables $X_1, \ldots, X_n$ are independent if for any $x_1, \ldots, x_n \in \mathbb{R}$,

$$P(X_1 = x_1, \ldots, X_n = x_n) = P(X_1 = x_1) \times \cdots \times P(X_n = x_n).$$
Independence of multiple random variables

**Definition**

Discrete random variables $X_1, \ldots, X_n$ are **independent** if for any $x_1, \ldots, x_n \in \mathbb{R}$,

$$P(X_1 = x_1, \ldots, X_n = x_n) = P(X_1 = x_1) \times \cdots \times P(X_n = x_n).$$

**Consequences:**

- $X_1$ and $X_2 + X_3$ are independent
- $X_4^2$ and $X_5X_6 - X_7$ are independent
- ...


Conditional independence

Independence (of events and random variables) naturally extends to conditional scenarios (B-H: Def 2.5.7 and 3.8.10):

**Definition**

Let $C$ be an event with $P(C) > 0$.

- **Events** $A$ and $B$ are *conditionally independent* given an event $C$ if
  \[ P(A \cap B|C) = P(A|C)P(B|C) \]
Conditional independence

Independence (of events and random variables) naturally extends to conditional scenarios (B-H: Def 2.5.7 and 3.8.10):

**Definition**

Let $C$ be an event with $P(C) > 0$.

- **Events** $A$ and $B$ are conditionally independent given an event $C$ if
  \[ P(A \cap B|C) = P(A|C)P(B|C) \]

- **Random variables** $X$ and $Y$ are conditionally independent given an event $C$ if
  \[ P(X = x, Y = y|C) = P(X = x|C)P(Y = y|C) \]
  for all $x, y$
Conditional independence

Independence (of events and random variables) naturally extends to conditional scenarios (B-H: Def 2.5.7 and 3.8.10):

**Definition**

Let $C$ be an event with $P(C) > 0$.

- **Events** $A$ and $B$ are conditionally independent given an event $C$ if
  \[ P(A \cap B|C) = P(A|C)P(B|C) \]
- **Random variables** $X$ and $Y$ are conditionally independent given an event $C$ if
  \[ P(X = x, Y = y|C) = P(X = x|C)P(Y = y|C) \]
  for all $x, y$
- **Random variables** $X$ and $Y$ are conditionally independent given a random variable $Z$ if
  \[ P(X = x, Y = y|Z = z) = P(X = x|Z = z)P(Y = y|Z = z) \]
  for all $x, y$ and for all $z$ such that $P(Z = z) > 0$
Example: fair die (continued)

Let $X$ be the result of a fair die. Let

$$Y = X \mod 2$$
$$Z = X \mod 3$$

Then

- $p_Y(0) = p_Y(1) = 1/2$
- $p_Z(0) = p_Z(1) = p_Z(2) = 1/3$
- We have verified that $Y$ and $Z$ independent.
Example: fair die (continued)

Let $X$ be the result of a fair die. Let

$$Y = X \mod 2$$
$$Z = X \mod 3$$

Then

- $p_Y(0) = p_Y(1) = 1/2$
- $p_Z(0) = p_Z(1) = p_Z(2) = 1/3$
- We have verified that $Y$ and $Z$ independent.
- Questions:
  - Are $Y$ and $Z$ independent conditioned on the event $\{X \leq 3\}$?
Example: fair die (continued)

Let $X$ be the result of a fair die. Let

$$Y = X \mod 2$$
$$Z = X \mod 3$$

Then

- $p_Y(0) = p_Y(1) = 1/2$
- $p_Z(0) = p_Z(1) = p_Z(2) = 1/3$
- We have verified that $Y$ and $Z$ independent.
- Questions:
  - Are $Y$ and $Z$ independent conditioned on the event $\{X \leq 3\}$? No.
    - For example, $P(Y = 0, Z = 1|X \leq 3) = 0$, but
    $$P(Y = 0|X \leq 3) = P(Z = 1|X \leq 3) = \frac{1}{3}$$
Example: fair die (continued)

Let $X$ be the result of a fair die. Let

$$Y = X \mod 2$$
$$Z = X \mod 3$$

Then

- $p_Y(0) = p_Y(1) = 1/2$
- $p_Z(0) = p_Z(1) = p_Z(2) = 1/3$
- We have verified that $Y$ and $Z$ independent.

Questions:

- Are $Y$ and $Z$ independent conditioned on the event $\{X \leq 3\}$? No.
  - For example, $P(Y = 0, Z = 1|X \leq 3) = 0$, but $P(Y = 0|X \leq 3) = P(Z = 1|X \leq 3) = \frac{1}{3}$
- Are $Y$ and $Z$ independent conditioned on the event $\{X \text{ is even}\}$? Yes. (Exercise)
Monty Hall problem

- A prize is behind one of the three doors. The other two are empty.
Monty Hall problem

- A prize is behind one of the three doors. The other two are empty.
- The guest randomly chooses a door.
- The host, knowing where the prize is, opens one of the remaining two doors that is empty. Specifically:
  - If guest's door is empty, host opens the other empty door;
  - If guest's door has prize, host randomly chooses an empty door.
- The host offers the guest the option to switch the choice to the other closed door.

**Monty Hall problem**

- A prize is behind one of the three doors. The other two are empty.
- The guest randomly chooses a door.
- The host, knowing where the prize is, opens one of the remaining two doors that is **empty**.

```
1 2
? ?
```
Monty Hall problem

- A prize is behind one of the three doors. The other two are empty.
- The guest randomly chooses a door.
- The host, knowing where the prize is, opens one of the remaining two doors that is empty. Specifically:
  - If guest’s door is empty, host opens the other empty door;
  - If guest’s door has prize, host randomly chooses an empty door.
Monty Hall problem

- A prize is behind one of the three doors. The other two are empty.
- The guest randomly chooses a door.
- The host, knowing where the prize is, opens one of the remaining two doors that is empty. Specifically:
  - If guest’s door is empty, host opens the other empty door;
  - If guest’s door has prize, host randomly chooses an empty door.
- The host offers the guest the option to switch the choice to the other closed door.
Monty Hall problem

• A prize is behind one of the three doors. The other two are empty
• The guest randomly chooses a door
• The host, knowing where the prize is, opens one of the remaining two doors that is empty. Specifically
  ▶ If guest’s door is empty, host opens the other empty door;
  ▶ If guest’s door has prize, host randomly chooses an empty door.
• The host offers the guest the option to switch the choice to the other closed door
• Question: Would you switch? Not switch? Doesn’t matter?
Monty Hall problem

Define the relevant random variables

- $X$: location of prize

...
Monty Hall problem

Define the relevant random variables

- $X$: location of prize
- $G$: choice of guest
- $H$: choice of host
- $S$: the other door

Then

- $X$ and $G$ are independent and equally likely to be 1, 2, 3.

Then $P(X = G) = \frac{1}{3}$ (analogy: for two dice $P(double) = \frac{1}{6}$ in lecture 2).

Per rule of the game, either $X = G$ or $X = S$.

Hence $P(X = S) = \frac{2}{3}$, i.e., switching is twice as likely to win.

Another way for guest to think:

- If my door is empty ($\frac{2}{3}$ chance), switch will win.
- If my door has prize ($\frac{1}{3}$ chance), switch will lose.
Monty Hall problem

Define the relevant random variables

- $X$: location of prize
- $G$: choice of guest
- $H$: choice of host
- $S$: the other door

Then

- $X$ and $G$ are independent and equally likely to be $1, 2, 3$.
- $P(X = G) = \frac{1}{3}$ (analogy: for two dice $P$ (double) = $\frac{1}{6}$ in lecture 2).
- Per rule of the game, either $X = G$ or $X = S$.
- Hence $P(X = S) = \frac{2}{3}$, i.e., switching is twice as likely to win.

Another way for guest to think:

- If my door is empty ($\frac{2}{3}$ chance), switch will win.
- If my door has prize ($\frac{1}{3}$ chance), switch will lose.
Monty Hall problem

Define the relevant random variables

- $X$: location of prize
- $G$: choice of guest
- $H$: choice of host
- $S$: the other door

Then

- $X$ and $G$ are independent and equally likely to be 1, 2, 3
- $P(X = G) = \frac{1}{3}$ (analogy: for two dice $P(\text{double}) = \frac{1}{6}$ in lecture 2)
- Per rule of the game, either $X = G$ or $X = S$
- Hence $P(X = S) = \frac{2}{3}$, i.e., switching is twice as likely to win

Another way for guest to think:

- If my door is empty ($\frac{2}{3}$ chance), switch will win
- If my door has prize ($\frac{1}{3}$ chance), switch will lose
Monty Hall problem

Define the relevant random variables

- \(X\): location of prize
- \(G\): choice of guest
- \(H\): choice of host
- \(S\): the other door

Then

- \(X\) and \(G\) are independent and equally likely to be 1, 2, 3
- \(P(X = G) = \frac{1}{3}\) (analogy: for two dice \(P\) (double) = \(\frac{1}{6}\) in lecture 2)
- Per rule of the game, either \(X = G\) or \(X = S\)
- Hence \(P(X = S) = \frac{2}{3}\), i.e., switching is twice as likely to win

Another way for guest to think:

- If my door is empty (\(\frac{2}{3}\) chance), switch will win
- If my door has prize (\(\frac{1}{3}\) chance), switch will lose
Monty Hall problem

Define the relevant random variables

- $X$: location of prize
- $G$: choice of guest
- $H$: choice of host
- $S$: the other door

Then

- $X$ and $G$ are independent and equally likely to be 1, 2, 3
Monty Hall problem

Define the relevant random variables

- $X$: location of prize
- $G$: choice of guest
- $H$: choice of host
- $S$: the other door

Then

- $X$ and $G$ are independent and equally likely to be 1, 2, 3
- Then $P(X = G) = \frac{1}{3}$
  (analogy: for two dice $P$(double) = $\frac{1}{6}$ in lecture 2)
Monty Hall problem

Define the relevant random variables

- $X$: location of prize
- $G$: choice of guest
- $H$: choice of host
- $S$: the other door

Then

- $X$ and $G$ are independent and equally likely to be 1, 2, 3
- Then $P(X = G) = \frac{1}{3}$
  (analogy: for two dice $P(\text{double}) = \frac{1}{6}$ in lecture 2)
- Per rule of the game, either $X = G$ or $X = S$
Monty Hall problem

Define the relevant random variables

- $X$: location of prize
- $G$: choice of guest
- $H$: choice of host
- $S$: the other door

Then

- $X$ and $G$ are independent and equally likely to be 1, 2, 3
- Then $P(X = G) = \frac{1}{3}$
  (analogy: for two dice $P$(double) = 1/6 in lecture 2)
- Per rule of the game, either $X = G$ or $X = S$
- Hence $P(X = S) = \frac{2}{3}$, i.e., switching is twice as likely to win
Monty Hall problem

Define the relevant random variables

- $X$: location of prize
- $G$: choice of guest
- $H$: choice of host
- $S$: the other door

Then

- $X$ and $G$ are independent and equally likely to be 1, 2, 3
- Then $P(X = G) = \frac{1}{3}$
  (analogy: for two dice $P($double$) = \frac{1}{6}$ in lecture 2)
- Per rule of the game, either $X = G$ or $X = S$
- Hence $P(X = S) = \frac{2}{3}$, i.e., switching is twice as likely to win
- Another way for guest to think:
  ▶ If my door is empty ($\frac{2}{3}$ chance), switch will win
  ▶ If my door has prize ($\frac{1}{3}$ chance), switch will lose
Monty Hall problem

Define the relevant random variables

- $X$: location of prize
- $G$: choice of guest
- $H$: choice of host
- $S$: the other door

Then

- $X$ and $G$ are independent and equally likely to be 1, 2, 3
- Then $P(X = G) = \frac{1}{3}$
  (analogy: for two dice $P$(double) = 1/6 in lecture 2)
- Per rule of the game, either $X = G$ or $X = S$
- Hence $P(X = S) = \frac{2}{3}$, i.e., switching is twice as likely to win
- Another way for guest to think:
  - If my door is empty (2/3 chance), switch will win
Monty Hall problem

Define the relevant random variables

- $X$: location of prize
- $G$: choice of guest
- $H$: choice of host
- $S$: the other door

Then

- $X$ and $G$ are independent and equally likely to be 1, 2, 3
- Then $P(X = G) = \frac{1}{3}$
  (analogy: for two dice $P$ (double) = $1/6$ in lecture 2)
- Per rule of the game, either $X = G$ or $X = S$
- Hence $P(X = S) = \frac{2}{3}$, i.e., switching is twice as likely to win
- Another way for guest to think:
  - If my door is empty (2/3 chance), switch will win
  - If my door has prize (1/3 chance), switch will lose
Monty Hall problem

- Maybe you are still not convinced because the guest knows which door he chooses and observes which door the hosts opens.
Monty Hall problem

- Maybe you are still not convinced because the guest knows which door he chooses and observes which door the hosts opens.
- The guest should evaluate the probability conditioned on these information in order to make an informed decision.
• Maybe you are still not convinced because the guest knows which door he chooses and observes which door the hosts opens
• The guest should evaluate the probability conditioned on these information in order to make an informed decision.
• So let’s compute some conditional probabilities
Monty Hall problem: conditional probability

Suppose guest chooses first door and host opens third door, i.e., $G = 1$ and $H = 3$. Goal: find

$$P(X = 1|G = 1, H = 3) \quad \text{and} \quad P(X = 2|G = 1, H = 3)$$
Monty Hall problem: conditional probability

Suppose guest chooses first door and host opens third door, i.e., $G = 1$ and $H = 3$. Goal: find

$$P(X = 1|G = 1, H = 3) \quad \text{and} \quad P(X = 2|G = 1, H = 3)$$

Use law of total probability and Bayes rule:

$$P(G = 1, H = 3) = P(G = 1, H = 3|X = 1) P(X = 1)$$

$$+ P(G = 1, H = 3|X = 2) P(X = 2)$$

$$+ P(G = 1, H = 3|X = 3) P(X = 3)$$

$$= \frac{1}{6}$$

and hence
Monty Hall problem: conditional probability

Suppose guest chooses first door and host opens third door, i.e., \( G = 1 \) and \( H = 3 \). Goal: find

\[
P(X = 1|G = 1, H = 3) \quad \text{and} \quad P(X = 2|G = 1, H = 3)
\]

Use law of total probability and Bayes rule:

\[
P(G = 1, H = 3) = P(G = 1, H = 3|X = 1) P(X = 1) + P(G = 1, H = 3|X = 2) P(X = 2) + P(G = 1, H = 3|X = 3) P(X = 3)
\]

\[
= \frac{1}{3} \times \frac{1}{2} \quad \frac{1}{3} + \frac{1}{3} \times 1 \quad \frac{1}{3} + \frac{1}{3} \times 0 \quad \frac{1}{3}
\]

\[
= \frac{1}{6}
\]

and hence

\[
P(X = 1|G = 1, H = 3) = \frac{1}{3} \quad \text{and} \quad P(X = 2|G = 1, H = 3) = \frac{2}{3}
\]