06: Random Variables

Lisa Yan
April 17, 2020
# Quick slide reference

<table>
<thead>
<tr>
<th></th>
<th>Topic</th>
<th>Slide Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Conditional Independence</td>
<td>06a_cond_indep</td>
</tr>
<tr>
<td>15</td>
<td>Random Variables</td>
<td>06b_random_variables</td>
</tr>
<tr>
<td>22</td>
<td>PMF/CDF</td>
<td>06c_pmf_cdf</td>
</tr>
<tr>
<td>30</td>
<td>Expectation</td>
<td>06d_expectation</td>
</tr>
<tr>
<td>40</td>
<td>Exercises</td>
<td>LIVE</td>
</tr>
</tbody>
</table>
Conditional Independence
Conditional Paradigm

For any events A, B, and E, you can condition consistently on E, and all formulas still hold:

Axiom 1

\[ 0 \leq P(A|E) \leq 1 \]

Corollary 1 (complement)

\[ P(A|E) = 1 - P(A^c|E) \]

Transitivity

\[ P(AB|E) = P(BA|E) \]

Chain Rule

\[ P(AB|E) = P(B|E)P(A|BE) \]

Bayes’ Theorem

\[ P(A|BE) = \frac{P(B|AE)P(A|E)}{P(B|E)} \]

BAE’s theorem?
Conditional Independence

Conditional Probability

Independence
Conditional Independence

Two events $A$ and $B$ are defined as conditionally independent given $E$ if:

$$P(AB|E) = P(A|E)P(B|E)$$

An equivalent definition:

A. $P(A|B) = P(A)$
B. $P(A|BE) = P(A)$
C. $P(A|BE) = P(A|E)$
Conditional Independence

Two events $A$ and $B$ are defined as conditionally independent given $E$ if:

$$P(AB|E) = P(A|E)P(B|E)$$

An equivalent definition:

A. $P(A|B) = P(A)$
B. $P(A|BE) = P(A)$
C. $P(A|BE) = P(A|E)$

Independent events $E$ and $F$

\[ P(EF) = P(E)P(F) \]
\[ P(E|F) = P(E) \]

\[ P(A|B) = P(A) \]
\[ P(A|BE) = P(A) \]
\[ P(A|BE) = P(A|E) \]
Conditional Independence

Independence relations can change with conditioning.

⚠ A and B independent does NOT always mean A and B independent given E. ⚠

(additional reading in lecture notes)

Conditional Probability Independence
Netflix and Condition

Let $E = \text{a user watches Life is Beautiful}$. 
Let $F = \text{a user watches Amelie}$. 
What is $P(E)$?

$$P(E) \approx \frac{\# \text{ people who have watched movie}}{\# \text{ people on Netflix}} = \frac{10,234,231}{50,923,123} \approx 0.20$$

What is the probability that a user watches Life is Beautiful, given they watched Amelie?

$$P(E|F) = \frac{P(EF)}{P(F)} = \frac{\# \text{ people who have watched both}}{\# \text{ people who have watched Amelie}} \approx 0.42$$
Let $E$ be the event that a user watches the given movie. Let $F$ be the event that the same user watches Amelie.

\[
P(E) = 0.19 \\
P(E) = 0.32 \\
P(E) = 0.20 \\
P(E) = 0.09 \\
P(E) = 0.20
\]

\[
P(E|F) = 0.14 \\
P(E|F) = 0.35 \\
P(E|F) = 0.20 \\
P(E|F) = 0.20 \\
P(E|F) = 0.42
\]

Independent!
What if $E_1, E_2, E_3, E_4$ are not independent? (e.g., all international emotional comedies)

$$P(E_4|E_1E_2E_3) = \frac{P(E_1E_2E_3E_4)}{P(E_1E_2E_3)} = \frac{\# \text{ people who have watched all 4}}{\# \text{ people who have watched those 3}}$$

We need to keep track of an exponential number of movie-watching statistics.
Netflix and Condition (on many movies)

Assume: $E_1, E_2, E_3, E_4$ are conditionally independent given $K$

$$P(E_4 | E_1 E_2 E_3 K) = P(E_4 | K)$$

An easier probability to store and compute!
Conditional independence is a Big Deal

Conditional independence is a practical, real-world way of decomposing hard probability questions.

“Exploiting conditional independence to generate fast probabilistic computations is one of the main contributions CS has made to probability theory.”

– Judea Pearl wins 2011 Turing Award,

“For fundamental contributions to artificial intelligence through the development of a calculus for probabilistic and causal reasoning”
Dependent events can become conditionally independent. And vice versa: Independent events can become conditionally dependent.

Netflix and Condition

Challenge: How do we determine \( K \)? Stay tuned in 6 weeks’ time!
Random Variables
Conditional independence review

Next Episode Playing in 5 seconds

\[ P(X = k) \]

\[ E[X] \]
Random variables are like typed variables

<table>
<thead>
<tr>
<th>type</th>
<th>name</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>int</td>
<td>a</td>
<td>5;</td>
</tr>
<tr>
<td>double</td>
<td>b</td>
<td>4.2;</td>
</tr>
<tr>
<td>bool</td>
<td>c</td>
<td>1;</td>
</tr>
</tbody>
</table>

Random variables are like typed variables (with uncertainty)

$A$ is the number of Pokemon we bring to our future battle.

$A \in \{1, 2, \ldots, 6\}$

$B$ is the amount of money we get after we win a battle.

$B \in \mathbb{R}^+$

$C$ is 1 if we successfully beat the Elite Four. 0 otherwise.

$C \in \{0,1\}$
Random Variable

A **random variable** is a real-valued function defined on a sample space.

Example:

3 coins are flipped. Let \( X = \# \) of heads. \( X \) is a random variable.

1. What is the value of \( X \) for the outcomes:
   - (T,T,T)?
   - (H,H,T)?

2. What is the event (set of outcomes) where \( X = 2 \)?

3. What is \( P(X = 2) \)?
Random Variable

A random variable is a real-valued function defined on a sample space.

Example:

3 coins are flipped. Let $X =$ # of heads. $X$ is a random variable.

1. What is the value of $X$ for the outcomes:
   - (T,T,T)? $X = 0$
   - (H,H,T)? $X = 2$

2. What is the event (set of outcomes) where $X = 2$?
   \[ \{ (H,H,T), (H,T,H), (T,H,H) \} = E \]

3. What is $P(X = 2)$?
   \[ P(E) = \frac{3}{8} \]
Random variables are **NOT** events!

It is confusing that random variables and events use the same notation.

- **Random variables ≠ events.**
- We can define an event to be a particular assignment of a random variable.

**Example:**

3 coins are flipped.
Let \( X = \# \) of heads.
\( X \) is a random variable.

\[
\begin{align*}
X &= 2 \\
P(X = 2) &= \text{probability (number b/t 0 and 1)}
\end{align*}
\]
Random variables are **NOT** events!

It is confusing that random variables and events use the same notation.

- **Random variables ≠ events.**
- We can define an event to be a particular assignment of a random variable.

<table>
<thead>
<tr>
<th>$X = x$</th>
<th>Set of outcomes</th>
<th>$P(X = k)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X = 0$</td>
<td>{$(T,T,T)$}</td>
<td>$1/8$</td>
</tr>
<tr>
<td>$X = 1$</td>
<td>{$(H,T,T)$; $(T,H,T)$, $(T,T,H)$}</td>
<td>$3/8$</td>
</tr>
<tr>
<td>$X = 2$</td>
<td>{$(H,H,T)$; $(H,T,H)$, $(T,H,H)$}</td>
<td>$3/8$</td>
</tr>
<tr>
<td>$X = 3$</td>
<td>{$(H,H,H)$}</td>
<td>$1/8$</td>
</tr>
<tr>
<td>$X \geq 4$</td>
<td>{}</td>
<td>$0$</td>
</tr>
</tbody>
</table>

Example:

3 coins are flipped. Let $X = \#$ of heads. $X$ is a **random variable.**
PMF/CDF
So far

3 coins are flipped. Let $X = \# \text{ of heads.}$ $X$ is a random variable.

<table>
<thead>
<tr>
<th>$X = x$</th>
<th>$P(X = k)$</th>
<th>Set of outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X = 0$</td>
<td>$1/8$</td>
<td>${(T, T, T)}$</td>
</tr>
<tr>
<td>$X = 1$</td>
<td>$3/8$</td>
<td>${(H, T, T), (T, H, T), (T, T, H)}$</td>
</tr>
<tr>
<td>$X = 2$</td>
<td>$3/8$</td>
<td>${(H, H, T), (H, T, H), (T, H, H)}$</td>
</tr>
<tr>
<td>$X = 3$</td>
<td>$1/8$</td>
<td>${(H, H, H)}$</td>
</tr>
<tr>
<td>$X \geq 4$</td>
<td>$0$</td>
<td>${}$</td>
</tr>
</tbody>
</table>

Can we get a “shorthand” for this last step? Seems like it might be useful!
Probability Mass Function

3 coins are flipped. Let $X = \# \text{ of heads}$. $X$ is a random variable.

A function on $k$ with range $[0,1]$ where $k$ is a parameter/input.

$P(X = k)$ is the return value/output number between 0 and 1.

What would be a useful function to define?

The probability of the event that a random variable $X$ takes on the value $k$!

For discrete random variables, this is a probability mass function.
Probability Mass Function

3 coins are flipped. Let $X = \#$ of heads. $X$ is a random variable.

A function on $k$ with range $[0,1]$:

$$P(X = k)$$

parameter/input $k$: a value of $X$

return value/output: probability of the event $X = k$

Example:

$N = 3$
$P = 0.5$

```python
def prob_event_y_equals(k):
    n_ways = scipy.special.binom(N, k)
    p_way = np.power(P, k) * np.power(1 - P, N - k)
    return n_ways * p_way
```

$0.375 = \frac{3}{8}$
Discrete RVs and Probability Mass Functions

A random variable $X$ is **discrete** if it can take on countably many values.

- $X = x$, where $x \in \{x_1, x_2, x_3, \ldots\}$

The **probability mass function** (PMF) of a discrete random variable is

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

- Probabilities must sum to 1:

$$\sum_{i=1}^{\infty} p(x_i) = 1$$

This last point is a good way to verify any PMF you create.
PMF for a single 6-sided die

Let $X$ be a random variable that represents the result of a single dice roll.

- **Support of $X$**: $\{1, 2, 3, 4, 5, 6\}$
- Therefore $X$ is a **discrete** random variable.
- **PMF of $X$**:

$$p(x) = \begin{cases} 
  \frac{1}{6} & x \in \{1, \ldots, 6\} \\
  0 & \text{otherwise} 
\end{cases}$$
Cumulative Distribution Functions

For a random variable $X$, the cumulative distribution function (CDF) is defined as

$$F(a) = F_X(a) = P(X \leq a), \text{ where } -\infty < a < \infty$$

For a discrete RV $X$, the CDF is:

$$F(a) = P(X \leq a) = \sum_{x \leq a} p(x)$$
CDFs as graphs

Let $X$ be a random variable that represents the result of a single dice roll.

\[ F(a) = P(X \leq a) \]

CDF of $X$

- $P(X \leq 0) = 0$
- $P(X \leq 6) = 1$
- Monotonically increasing function

PMF of $X$
Expectation

Challenging, but fun
Discrete random variables

PMF
\[ P(X = x) = p(x) \]

CDF \( F(x) = \Pr(X \leq x) \)

Without performing the experiment:

- The support gives us a ballpark of what values our RV will take on
- Next up: How do we report the “average” value?

Definition

Properties

Experiment outcomes

Support
Expectation

The **expectation** of a discrete random variable $X$ is defined as:

$$E[X] = \sum_{x:\ p(x)>0} p(x) \cdot x$$

- Note: sum over all values of $X = x$ that have non-zero probability.

- Other names: **mean**, expected value, **weighted average**, center of mass, first moment
Expectation of a die roll

What is the expected value of a 6-sided die roll?

1. Define random variables
   
   \[ X = \text{RV for value of roll} \]
   
   \[ P(X = x) = \begin{cases} 
   \frac{1}{6} & x \in \{1, \ldots, 6\} \\
   0 & \text{otherwise} 
   \end{cases} \]

2. Solve
   
   \[ E[X] = \sum_{x=p(x)>0} p(x) \cdot x \]

   \[ E[X] = 1 \left( \frac{1}{6} \right) + 2 \left( \frac{1}{6} \right) + 3 \left( \frac{1}{6} \right) + 4 \left( \frac{1}{6} \right) + 5 \left( \frac{1}{6} \right) + 6 \left( \frac{1}{6} \right) = \frac{7}{2} = 3.5 \]
Important properties of expectation

1. **Linearity:**

   \[ E[aX + b] = aE[X] + b \]

2. **Expectation of a sum = sum of expectation:**

   \[ E[X + Y] = E[X] + E[Y] \]

3. **Unconscious statistician:**

   \[ E[g(X)] = \sum_x g(x)p(x) \]

   - Let \( X \) = 6-sided dice roll, \( Y = 2X - 1 \).
   - \( E[X] = 3.5 \)
   - \( E[Y] = 6 = 2(3.5) - 1 \)

   Sum of two dice rolls:
   - Let \( X \) = roll of die 1
     \( Y = \) roll of die 2
   - \( E[X + Y] = 3.5 + 3.5 = 7 \)

   **(more in lecture)**

   These properties let you avoid defining difficult PMFs.
Proofs (OK to stop here)
Important properties of expectation

1. Linearity:
   \[ E[aX + b] = aE[X] + b \]

2. Expectation of a sum = sum of expectation:
   \[ E[X + Y] = E[X] + E[Y] \]

3. Unconscious statistician:
   \[ E[g(X)] = \sum_x g(x)p(x) \]
Linearity of Expectation proof

\[ E[ax + b] = aE[X] + b \]

Proof:

\[ E[ax + b] = \sum_x (ax + b)p(x) = \sum_x axp(x) + bp(x) \]

\[ = a \sum_x xp(x) + b \sum_x p(x) \]

\[ = a \cdot E[X] + b \cdot 1 \]

\[ = a \cdot E[X] + b \]
Expectation of Sum intuition

\[ E[X + Y] = E[X] + E[Y] \]

(we’ll prove this in two weeks)

**Intuition for now:**

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(X)</td>
<td>(Y)</td>
<td>(X + Y)</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>12</td>
<td>18</td>
</tr>
<tr>
<td>10</td>
<td>20</td>
<td>30</td>
</tr>
<tr>
<td>-1</td>
<td>-2</td>
<td>-3</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>16</td>
<td>24</td>
</tr>
</tbody>
</table>

Average:

\[ \frac{1}{n} \sum_{i=1}^{n} x_i + \frac{1}{n} \sum_{i=1}^{n} y_i = \frac{1}{n} \sum_{i=1}^{n} (x_i + y_i) \]

\[ \frac{1}{7} (28) + \frac{1}{7} (56) = \frac{1}{7} (84) \]
LOTUS proof

Let $Y = g(X)$, where $g$ is a real-valued function.

$$E[g(X)] = E[Y] = \sum_j y_j p(y_j) = \sum_j \sum_{i: g(x_i) = y_j} p(x_i) = \sum_j \sum_{i: g(x_i) = y_j} y_j p(x_i) = \sum_j \sum_{i: g(x_i) = y_j} g(x_i) p(x_i) = \sum_i g(x_i) p(x_i)$$

Expectation of $g(X)$

For you to review so that you can sleep at night
06: Random Variables

Lisa Yan
April 15, 2020
Reminders: Lecture with zoom

• Turn on your camera if you are able, mute your mic in the big room
• Virtual backgrounds are encouraged (classroom-appropriate)

Breakout Rooms for meeting your classmates
  ◦ Just like sitting next to someone new
  ◦ This experience is optional: You should be comfortable leaving the room at any time.

We will use Ed instead of Zoom chat

Today’s discussion thread: https://us.edstem.org/courses/109/discussion/24491
Discrete random variables

Definition

Properties

Experiment outcomes

Discrete Random Variable, $X$

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

$E[X]$

Support

Note: Random Variables also called distributions
A Whole New World with Random Variables

Event-driven probability

• Relate only binary events
  ◦ Either happens ($E$)
  ◦ or doesn’t happen ($E^C$)

• Can only report probability

• Lots of combinatorics

Random Variables

• Link multiple similar events together ($X = 1$, $X = 2$, ..., $X = 6$)

• Can compute statistics: report the “average” outcome

• Once we have the PMF (discrete RVs), we can do regular math
PMF for the sum of two dice

Let $Y$ be a random variable that represents the sum of two independent dice rolls.

Support of $Y$: \{2, 3, ..., 11, 12\}

$$p(y) = \begin{cases} \frac{y-1}{36} & y \in \mathbb{Z}, 2 \leq y \leq 6 \\ \frac{13-y}{36} & y \in \mathbb{Z}, 7 \leq y \leq 12 \\ 0 & \text{otherwise} \end{cases}$$

Sanity check: $$\sum_{y=2}^{12} p(y) = 1$$
Think, then Breakout Rooms

Then check out the question on the next slide (Slide 46). Post any clarifications here!

https://us.edstem.org/courses/109/discussion/27280

Think by yourself: 1 min

Breakout rooms: 5 min. Introduce yourself! (or leave)
Example random variable

Consider 5 flips of a coin which comes up heads with probability $p$. Each coin flip is an independent trial. Let $Y = \# \text{ of heads on 5 flips.}$

1. What is the support of $Y$? In other words, what are the values that $Y$ can take on with non-zero probability?

2. Define the event $Y = 2$. What is $P(Y = 2)$?

3. What is the PMF of $Y$? In other words, what is $P(Y = k)$, for $k$ in the support of $Y$?
Example random variable

Consider 5 flips of a coin which comes up heads with probability $p$. Each coin flip is an independent trial. Let $Y = \# \text{ of heads on 5 flips}$.

1. What is the support of $Y$? In other words, what are the values that $Y$ can take on with non-zero probability? $\{0, 1, 2, 3, 4, 5\}$

2. Define the event $Y = 2$. What is $P(Y = 2)$?

$$P(Y = 2) = \binom{5}{2} p^2 (1-p)^3$$

3. What is the PMF of $Y$? In other words, what is $P(Y = k)$, for $k$ in the support of $Y$?

$$P(Y = k) = \binom{5}{k} p^k (1-p)^{5-k}$$
Expectation

\[ E[X] = \sum_{x:p(x)>0} p(x) \cdot x \]

Expectation: The **average value** of a random variable

Remember that the expectation of a die roll is 3.5.

\[ X = \text{RV for value of roll} \]

\[ E[X] = 1 \left( \frac{1}{6} \right) + 2 \left( \frac{1}{6} \right) + 3 \left( \frac{1}{6} \right) + 4 \left( \frac{1}{6} \right) + 5 \left( \frac{1}{6} \right) + 6 \left( \frac{1}{6} \right) = \frac{7}{2} \]
Lying with statistics

“There are three kinds of lies: lies, damned lies, and statistics” – popularized by Mark Twain, 1906
Lying with statistics

A school has 3 classes with 5, 10, and 150 students. What is the average class size?

1. Interpretation #1
   - Randomly choose a class with equal probability.
   - \( X = \) size of chosen class

   \[
   E[X] = 5 \left( \frac{1}{3} \right) + 10 \left( \frac{1}{3} \right) + 150 \left( \frac{1}{3} \right)
   \]

   \[
   = \frac{165}{3} = 55
   \]

2. Interpretation #2
   - Randomly choose a student with equal probability.
   - \( Y = \) size of chosen class

   \[
   E[Y] = 5 \left( \frac{5}{165} \right) + 10 \left( \frac{10}{165} \right) + 150 \left( \frac{150}{165} \right)
   \]

   \[
   = \frac{22635}{165} \approx 137
   \]

What universities usually report

Average student perception of class size

Key takeaway: these two RVs are different. The problem statement was vague.
Interlude for jokes/announcements
Every quarter, students report **counting** as the hardest topic in this class.

- **We are making psets shorter** to reflect the additional time you spend on lecture. PS2 has just 10 problems.

- **More practice resources:** Textbook’s Self-Test Problems
The pedagogy behind concept checks

- Spaced practice (vs. "practice makes perfect"): better memory retention
- Low-stakes testing: better concept retrieval, actively connect concepts

**It is okay if you don’t understand material off-the-bat.** In fact, learning research suggests that you will learn more in the long run.
Abel Prize in Mathematics Shared by 2 Trailblazers of Probability and Dynamics

nytimes.com

Important properties of expectation

1. **Linearity**:\[ E[aX + b] = aE[X] + b \]


3. **Unconscious statistician**:\[ E[g(X)] = \sum_x g(x)p(x) \]

Roll a die, outcome is $X$. You win $2X - 1$. What are your expected winnings?

Let $X = 6$-sided dice roll.

\[ E[2X - 1] = 2(3.5) - 1 = 6 \]
Important properties of expectation

1. Linearity:
   \[ E[aX + b] = aE[X] + b \]
   Roll a die, outcome is \( X \). You win \( 2X - 1 \). What are your expected winnings?
   Let \( X \) = 6-sided dice roll.
   \[ E[2X - 1] = 2(3.5) - 1 = 6 \]

2. Expectation of a sum = sum of expectation:
   \[ E[X + Y] = E[X] + E[Y] \]
   What is the expectation of the sum of two dice rolls?
   Let \( X \) = roll of die 1, \( Y \) = roll of die 2.
   \[ E[X + Y] = 3.5 + 3.5 = 7 \]

3. Unconscious statistician:
   \[ E[g(X)] = \sum_x g(x)p(x) \]
   Big takeaway: You don’t need to calculate crazy PMFs if you can use these properties of expectation.

Lisa Yan, CS109, 2020
Important properties of expectation

1. Linearity:
   \[ E[aX + b] = aE[X] + b \]

Roll a die, outcome is \( X \). You win \( 2X - 1 \).
What are your expected winnings?
Let \( X = 6 \)-sided dice roll.
\[ E[2X - 1] = 2(3.5) - 1 = 6 \]

2. Expectation of a sum = sum of expectation:
   \[ E[X + Y] = E[X] + E[Y] \]

What is the expectation of the sum of two dice rolls?
Let \( X = \) roll of die 1, \( Y = \) roll of die 2.
\[ E[X + Y] = 3.5 + 3.5 = 7 \]

3. Unconscious statistician:
   \[ E[g(X)] = \sum_{x} g(x)p(x) \]

(next up)
Think, then Breakout Rooms

Then check out the question on the next slide (Slide 58). Post any clarifications here!

https://us.edstem.org/courses/109/discussion/27280

Think by yourself: 1 min

Breakout rooms: 5 min. Introduce yourself!
Being a statistician unconsciously

Let $X$ be a discrete random variable.

- $P(X = x) = \frac{1}{3}$ for $x \in \{-1, 0, 1\}$

Let $Y = |X|$. What is $E[Y]$?

A. $\frac{1}{3} \cdot 1 + \frac{1}{3} \cdot 0 + \frac{1}{3} \cdot -1 = 0$

B. $E[Y] = E[0] = 0$

C. $\frac{1}{3} \cdot 0 + \frac{2}{3} \cdot 1 = \frac{2}{3}$

D. $\frac{1}{3} \cdot |−1| + \frac{1}{3} \cdot |0| + \frac{1}{3} \cdot |1| = \frac{2}{3}$

E. C and D
Being a statistician unconsciously

Let $X$ be a discrete random variable.
- $P(X = x) = \frac{1}{3}$ for $x \in \{-1, 0, 1\}$

Let $Y = |X|$. What is $E[Y]$?

A. $\frac{1}{3} \cdot 1 + \frac{1}{3} \cdot 0 + \frac{1}{3} \cdot (-1) = 0$ $\times E[X]$

B. $E[Y] = E[0] = 0$ $E[E[X]]$

C. $\frac{1}{3} \cdot 0 + \frac{2}{3} \cdot 1 = \frac{2}{3}$

D. $\frac{1}{3} \cdot |-1| + \frac{1}{3} \cdot |0| + \frac{1}{3} \cdot |1| = \frac{2}{3}$

E. C and D

**Recompute PMF**

1. Find PMF of $Y$: $p_Y(0) = \frac{1}{3}, p_Y(1) = \frac{2}{3}$
2. Compute $E[Y]$ Use LOTUS by using PMF of $X$:
   1. $P(X = x) \cdot |x|$
   2. Sum up
I want to play a game
Then check out the question on the next slide (Slide 30). Post any clarifications here!

https://us.edstem.org/courses/109/discussion/27280

Think by yourself: 2 min
St. Petersburg Paradox

• A fair coin (comes up “heads” with $p = 0.5$)
• Define $Y = \text{number of coin flips (“heads”)}$ before first “tails”
• You win $2^Y$

How much would you pay to play? (How much you expect to win?)

A. $10000$
B. $\infty$
C. $1$
D. $0.50$
E. $0$ but let me play
F. I will not play
St. Petersburg Paradox

- A fair coin (comes up “heads” with \( p = 0.5 \))
- Define \( Y \) = number of coin flips (“heads”) before first “tails”
- You win \( 2^Y \)

How much would you pay to play? (How much you expect to win?)

1. Define random variables
   
   For \( i \geq 0 \):
   \[
   P(Y = i) = \left( \frac{1}{2} \right)^{i+1}
   \]

   Let \( W \) = your winnings, \( 2^Y \).

2. Solve
   \[
   E[W] = E[2^Y] = \left( \frac{1}{2} \right)^1 2^0 + \left( \frac{1}{2} \right)^2 2^1 + \left( \frac{1}{2} \right)^3 2^2 + \cdots
   \]
   \[
   = \sum_{i=0}^{\infty} \left( \frac{1}{2} \right)^{i+1} 2^i = \sum_{i=0}^{\infty} \left( \frac{1}{2} \right) = \infty
   \]

\[ E[g(x)] = \sum_x g(x)p(x) \]
St. Petersburg + Reality

What if Lisa has only $65,536?

- Same game
- Define $Y = \#$ heads before first tails
- You win $W = 2^Y$

- If you win over $65,536$, I leave the country

1. Define random variables

For $i \geq 0$: 
$$P(Y = i) = \left(\frac{1}{2}\right)^{i+1}$$

Let 
$$W = \text{your winnings, } 2^Y.$$

2. Solve

$$E[W] = \left(\frac{1}{2}\right)^1 2^0 + \left(\frac{1}{2}\right)^2 2^1 + \left(\frac{1}{2}\right)^3 2^2 + \cdots$$

$k = \log_2(65,536)$

$= 16$ # Heads you can flip

$$= \sum_{i=0}^{16} \left(\frac{1}{2}\right)^{i+1} 2^i = \sum_{i=0}^{16} \left(\frac{1}{2}\right) = 8.5$$