Averages and Expected Values

- Say the scores on a quiz are 9, 10, 9, 6, and 7.
  What is the average? $\frac{1}{5}(9 + 10 + 9 + 6 + 7) = \frac{41}{5} = 8.2$

- We need a probabilistic notion of averages.

- The expected value $\mathbb{E}[X]$ of a discrete random variable X is

$$\boxed{\mathbb{E}[X] = \sum_{x \in R_X} x \, P_X(x)}$$

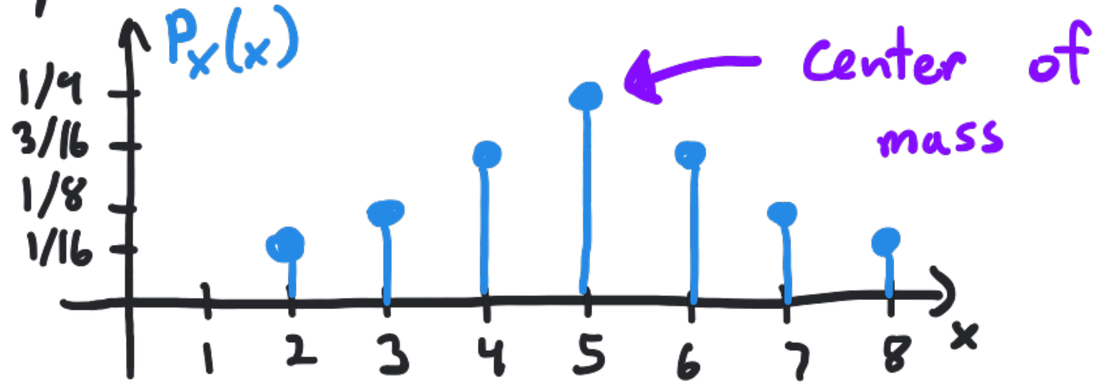
- Other Terminology: average, mean, μ_X

- Intuition: The probability (or PMF value) at x is the "mass" of that point and the expected value is the "center of mass."

- Back to Experiment 3: Roll two 4-sided dice, all outcomes equally likely. X is the sum of the rolls.

(from previous video)

x	2	3	4	5	6	7	8
$P_X(x)$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{3}{16}$	$\frac{1}{4}$	$\frac{3}{16}$	$\frac{1}{8}$	$\frac{1}{16}$



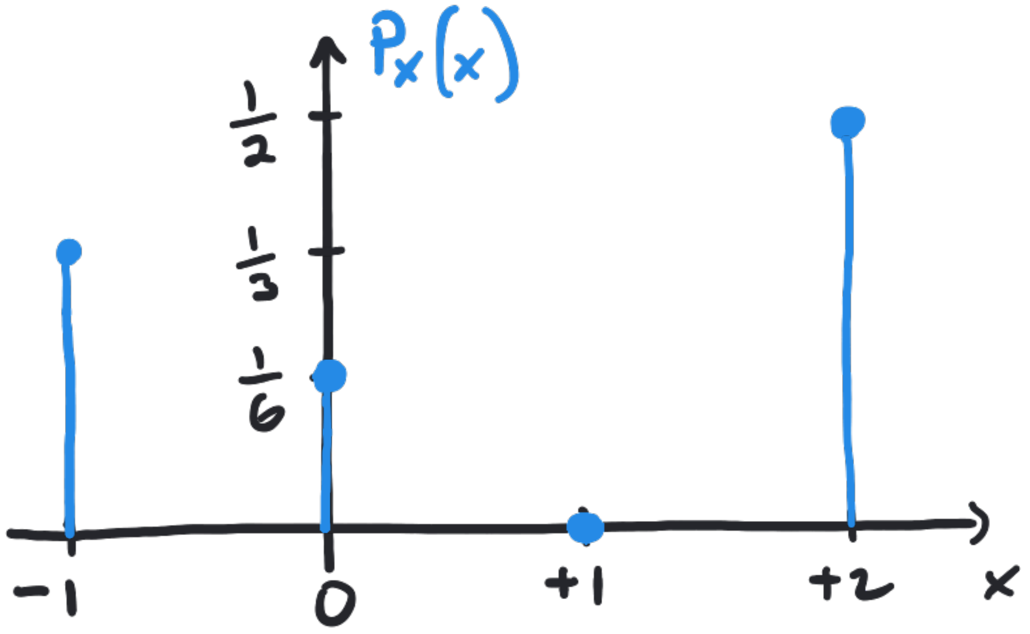
$$\mathbb{E}[X] = \sum_{x \in R_X} x \, P_X(x) = 2 \cdot \frac{1}{16} + 3 \cdot \frac{1}{8} + 4 \cdot \frac{3}{16} + 5 \cdot \frac{1}{4} + 6 \cdot \frac{3}{16} + 7 \cdot \frac{1}{8} + 8 \cdot \frac{1}{16}$$

$$= \frac{2 + 6 + 12 + 20 + 18 + 14 + 8}{16}$$

$$= \frac{80}{16}$$

$$= 5$$ ← corresponds to the center of mass of the plot

- **Example:**



$R_X = \{-1, 0, +2\}$

$$\mathbb{E}[X] = \sum_{x \in R_X} x \, P_X(x) = (-1) \cdot P_X(-1) + 0 \cdot P_X(0) + 2 \cdot P_X(2)$$

$$= (-1) \cdot \frac{1}{3} + 0 \cdot \frac{1}{6} + 2 \cdot \frac{1}{2}$$

$$= \frac{2}{3}$$

→ Note that $\mathbb{E}[X]$ does not need to belong to R_X.

Functions of a Discrete Random Variable

- A function $Y = g(X)$ of a random variable is itself a random variable.
  - $\to Y(\omega) = g(X(\omega))$ Mapping from sample space Ω to real line.
  - $\to$ The range R_Y and PMF $P_Y(y)$ of Y can be determined using the range R_X and PMF $P_X(x)$ of X along with the function $y = g(x)$.

Range: $\boxed{R_Y = \{ g(x) : x \in R_X \}}$

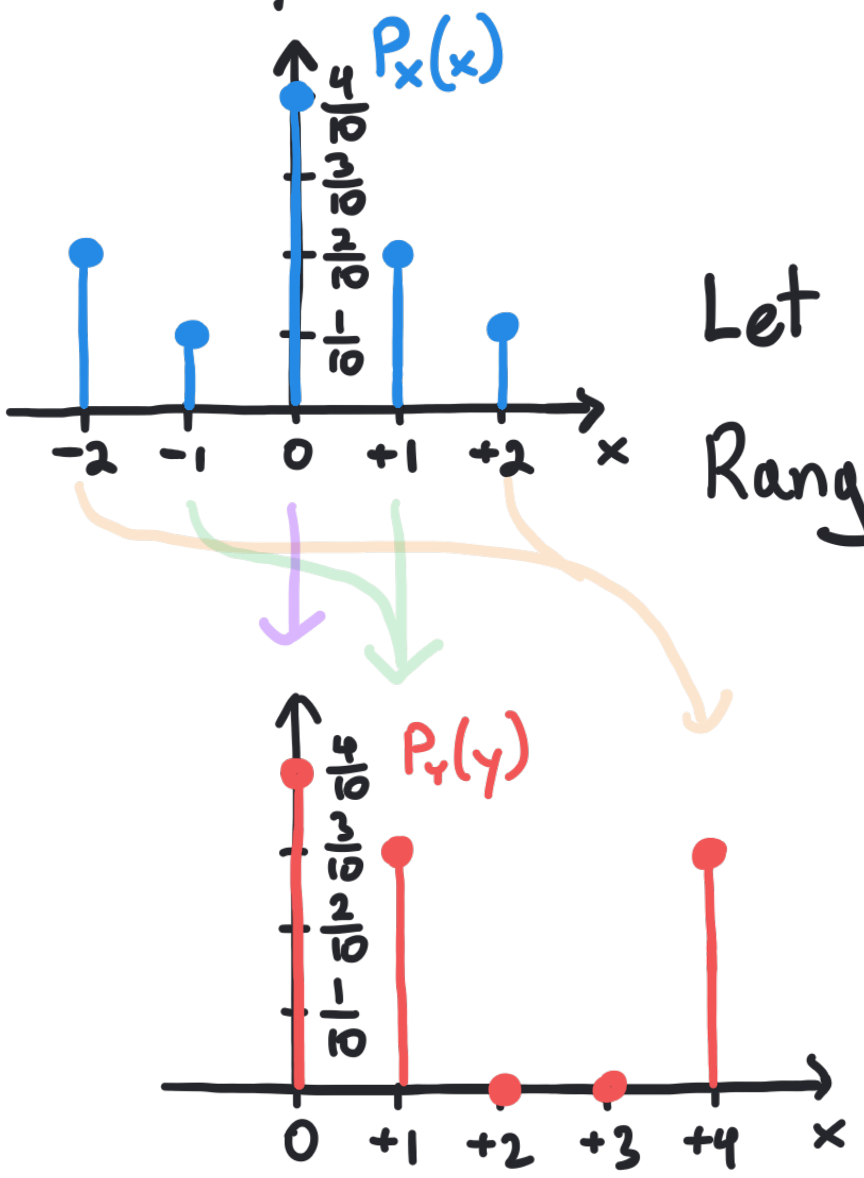
PMF:
$$P_Y(y) = \mathbb{P}[\{\omega \in \Omega : Y(\omega) = y\}]$$
$$= \mathbb{P}[\{\omega \in \Omega : g(X(\omega)) = y\}]$$

$$\Rightarrow \boxed{P_Y(y) = \sum_{x : g(x) = y} P_X(x)}$$

Notation: $\sum_{x : g(x) = y}$ means "sum over $x \in R_X$ such that $g(x) = y$"

• **Example:**

$$P_X(x) = \begin{cases} \frac{4}{10} & x = 0 \\ \frac{2}{10} & x = -2, +1 \\ \frac{1}{10} & x = -1, +2 \end{cases} \qquad R_X = \{-2, -1, 0, +1, +2\}$$



Let $Y = X^2$.

Range of Y: $R_Y = \{x^2 : x \in R_X\}$

$= \{0, +1, +4\}$

PMF of Y: $P_Y(y) = \sum_{x : x^2 = y} P_X(x)$

$$P_Y(0) = \sum_{x : x^2 = 0} P_X(x) = P_X(0) = \frac{4}{10}$$

$$P_Y(+1) = \sum_{x : x^2 = +1} P_X(x) = P_X(-1) + P_X(+1) = \frac{1}{10} + \frac{2}{10} = \frac{3}{10}$$

$$P_Y(+4) = \sum_{x : x^2 = +4} P_X(x) = P_X(-2) + P_X(+2) = \frac{2}{10} + \frac{1}{10} = \frac{3}{10}$$

Expectation of Y: $\mathbb{E}[Y] = \sum_{y \in R_Y} y\, P_Y(y) = 0 \cdot \frac{4}{10} + (+1) \cdot \frac{3}{10} + (+4) \cdot \frac{3}{10}$

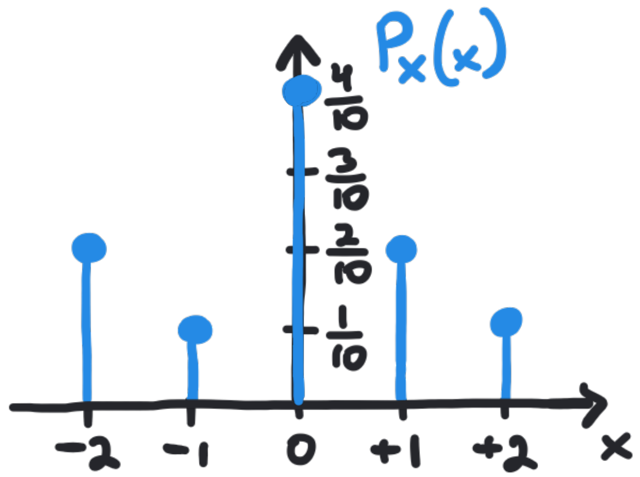
$$= \frac{15}{10} = \frac{3}{2}$$

- We will often only wish to know the expected value of a function $Y = g(X)$ (and not the PMF nor the range).
- In these settings, we can directly calculate $\mathbb{E}[Y]$ using the PMF $P_X(x)$ of X and the function $y = g(x)$.

$$\mathbb{E}[Y] = \sum_{y \in R_Y} y \, P_Y(y) = \sum_{y \in R_Y} y \sum_{x : g(x) = y} P_X(x) = \sum_{y \in R_Y} \sum_{x : g(x) = y} g(x) \, P_X(x)$$

$$\Rightarrow \boxed{\mathbb{E}[Y] = \sum_{x \in R_X} g(x) \, P_X(x)}$$

- Previous Example: Let $Y = X^2$. Calculate $\mathbb{E}[Y]$.



$$\mathbb{E}[Y] = \sum_{x \in R_X} x^2 P_X(x)$$

$$= (-2)^2 \cdot \frac{2}{10} + (-1)^2 \cdot \frac{1}{10} + (0)^2 \cdot \frac{4}{10} + (+1)^2 \cdot \frac{2}{10} + (+2)^2 \cdot \frac{1}{10}$$

$$= \frac{8 + 1 + 0 + 2 + 4}{10} = \frac{15}{10} = \frac{3}{2}$$

Same as before!

- Another important property of expectation is linearity:

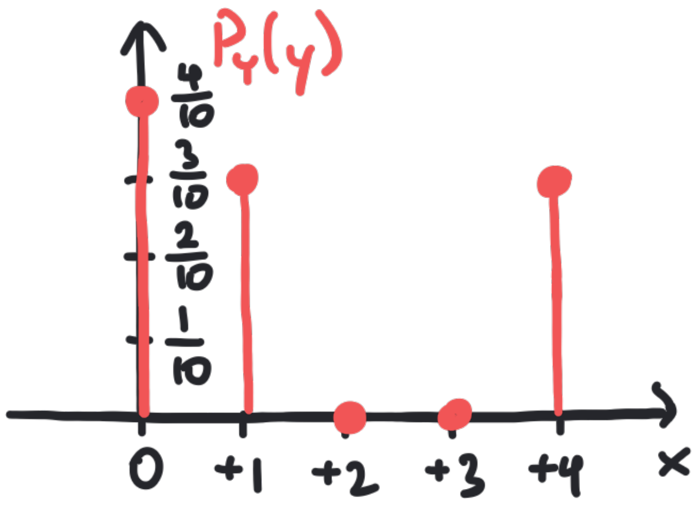
$$\boxed{\text{For any random variable } X \text{ and real numbers } a \text{ and } b, \quad \mathbb{E}[aX+b] = a\,\mathbb{E}[X] + b}$$

→ Why? $\mathbb{E}[aX+b] = \sum_{x \in R_X} (ax+b)\,P_X(x)$

$$= a \sum_{x \in R_X} x\,P_X(x) + b \underbrace{\sum_{x \in R_X} P_X(x)}_{\text{Normalization}}$$

$$= a\,\mathbb{E}[X] + b \cdot 1$$

- Previous Example: Let $Z = -\frac{1}{2}Y + 3$. Calculate $\mathbb{E}[Z]$.



Recall $\mathbb{E}[Y] = \frac{3}{2}$

$$\mathbb{E}[Z] = -\frac{1}{2}\,\mathbb{E}[Y] + 3 = -\frac{1}{2}\cdot\frac{3}{2} + 3$$

$$= -\frac{3}{4} + 3$$

$$= \frac{9}{4}$$