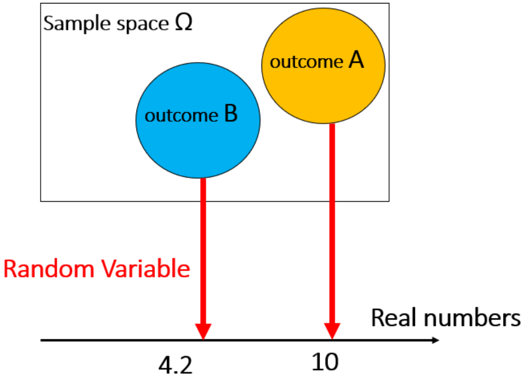
Discrete Random Variables, PMF, Expectations

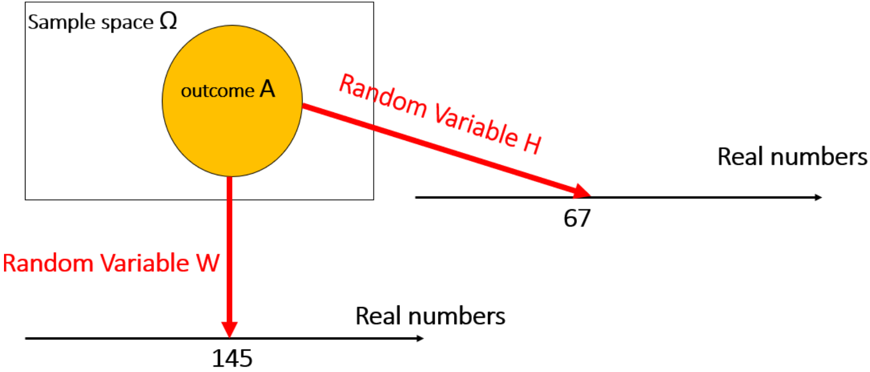
Random Variables

- In an experiment, a **random variable** is an assignment of a value (number) to every possible outcome.
- **Mathematically**: A random variable is a **function** from the **sample space** Ω to the **real numbers**.
- Random variable is not random, is not a variable, it's just a function from sample space to real numbers.



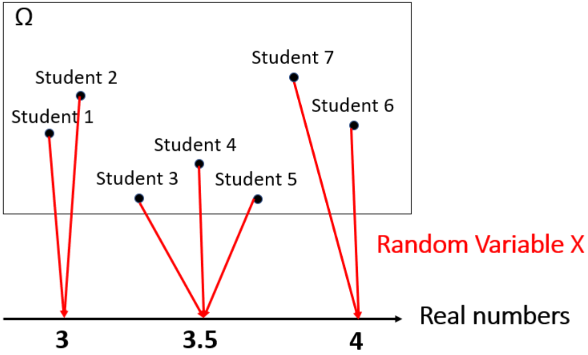
Random Variables

- In a single probability experiment, you can have multiple random variables.
- For example, sample space is set of all students in a class. One random variable can be "Height of the students" and one other random variable can be "Weight of the students".



Random Variables— Example

- In an experiment, sample space is all students of a class (suppose the class has 7 students).
- We define random variable X to be "GPA" of the students.

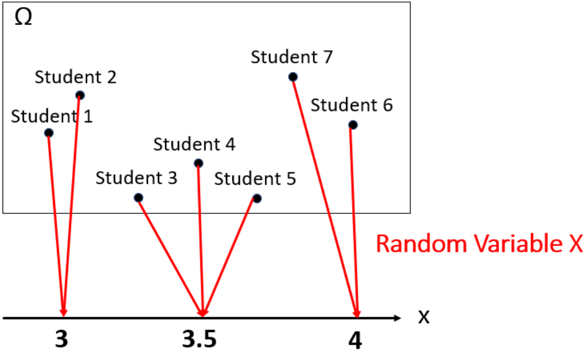


Random Variables—Example

Notations:

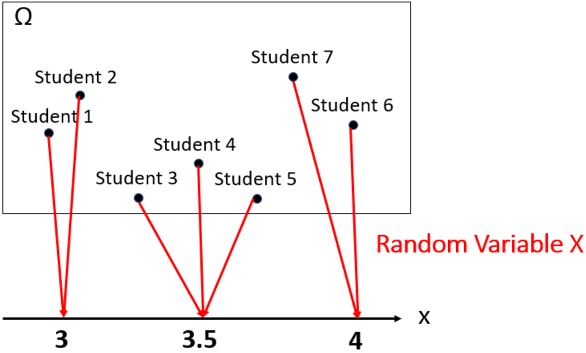
- X: random variable, a function from sample space to real numbers
- x: numerical value, a real number

- Random variable X is mapping 2 outcomes to $x = 3$.
- Random variable X is mapping 3 outcomes to $x = 3.5$.
- Random variable X is mapping 2 outcomes to $x = 4$.



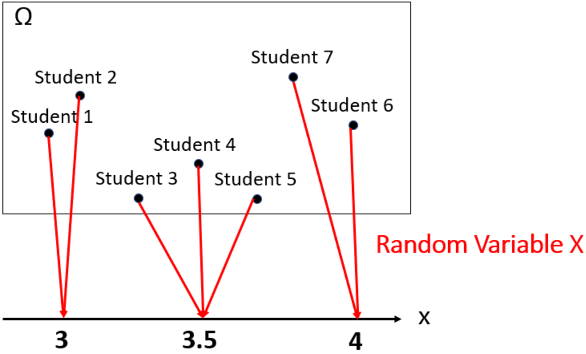
Random Variables and Probabilities

- The experiment: choose a student and ask their GPA.
- Question: how likely is it that we choose a student with GPA 4?
- Out of 7 students, GPA of 2 of them is 4. Therefore, the answer is $\frac{2}{7}$.



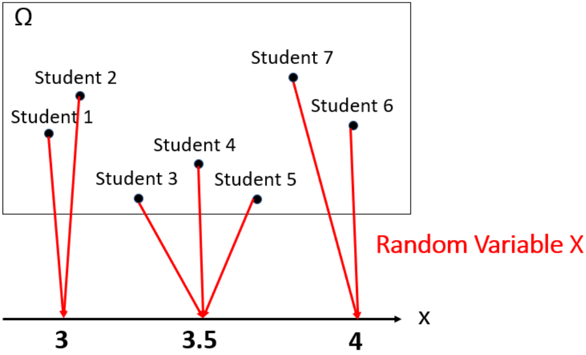
Random Variables and Probabilities

- Some numerical values are more likely to occur than other numerical values and we capture this by assigning probabilities to them.
- How likely is it that we choose a student with GPA 3? $\frac{2}{7}$
- How likely is it that we choose a student with GPA 3.5? $\frac{3}{7}$
- How likely is it that we choose a student with GPA 4? $\frac{2}{7}$



Random Variables and Probabilities

- **Probability of a particular numerical value**: probability of all outcomes that lead to that particular numerical value.
- Question: what is $P(X = 4)$?
- Answer: $P(X = 4) = P(\text{seeing all students with GPA 4}) = \frac{2}{7}$.
- Therefore, in this experiment, $x = 3.5$ is more likely to occur than $x = 4$ or $x = 3$.



Probability Mass Function(PMF)

Notation:

$$p_X(x) = P(X = x) = P(\{\omega \in \Omega \ s.t. \ X(\omega) = x\})$$

is probability of <u>all outcomes that are mapped to x</u> by random variable X.

Properties:

- $p_X(x) \geq 0$: probability is always non negative.
- $\sum_x p_X(x) = 1$: sum of probability of all outcomes is always equal to 1.

Probability Mass Function(PMF)—Example

- Experiment: You have a biased coin, $P(Head) = p$ and $P(Tail) = 1 - p$, and $p > 0$. You toss the coin until you see a "Head" and then you stop. Assume tosses are independent.
- Define random variable X to be "number of coin tosses until first head". Then

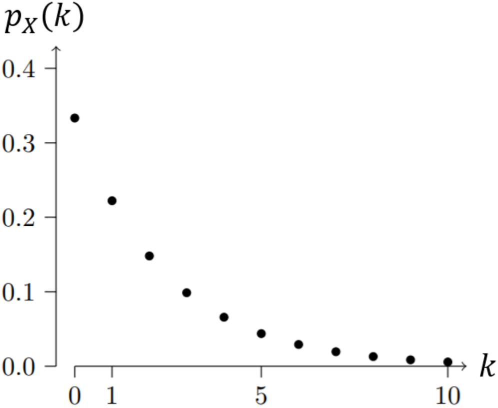
$$p_X(k) = P(X = k) = P(\text{"number of coin tosses until first head''} = k)$$

$$= P(\underbrace{TT \cdots T}_{\text{k-1}} H) = (1 - p)^{k-1} p$$

- If H $\Rightarrow X = 1 \Rightarrow p$
- If TH $\Rightarrow X = 2 \Rightarrow (1 - p)p$
- If TTH $\Rightarrow X = 3 \Rightarrow (1 - p)^2 p$
- If TTTH $\Rightarrow X = 4 \Rightarrow (1 - p)^3 p$

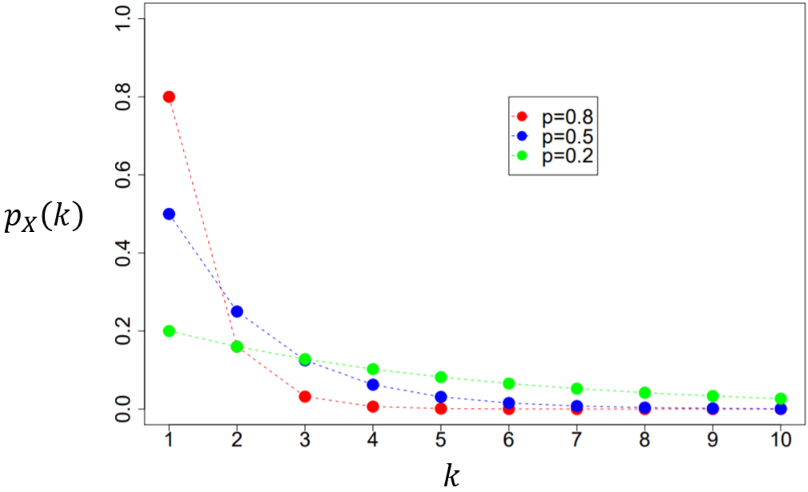
Geometric PMF

- Probability distribution
$$p_X(k) = (1-p)^{k-1}p \text{ for } k = 1, 2, 3 \cdots$$
is called **Geometric Distribution.**
- Parameter of geometric distribution is p.
- Draw geometric distribution for $p = \frac{1}{3}$.



Geometric PMF

- Draw geometric distribution for $p = 0.2, 0.5, 0.8$.



Probability Mass Function

How to compute PMF for $p_X(x)$?

- Collect all possible outcomes for which X is equal to x.
- Add their probabilities.
- Repeat for all x.

Example:

- You have a 4 sided dice and you roll it twice. Random variable X is minimum of first roll and second roll.
- What is $p_X(2) =?$
- Find all possible outcomes such that minimum of first roll and second roll is 2: $\{(2,2),(2,3),(2,4),(3,2),(4,2)\}$
- $p_X(2) = P(X = 2) = P(\{(2,2),(2,3),(2,4),(3,2),(4,2)\}) = \frac{5}{16}$.
- You can check that $p_X(1) = \frac{7}{16}, p_X(3) = \frac{3}{16}, p_X(4) = \frac{1}{16}$.

Probability Mass Function—Example

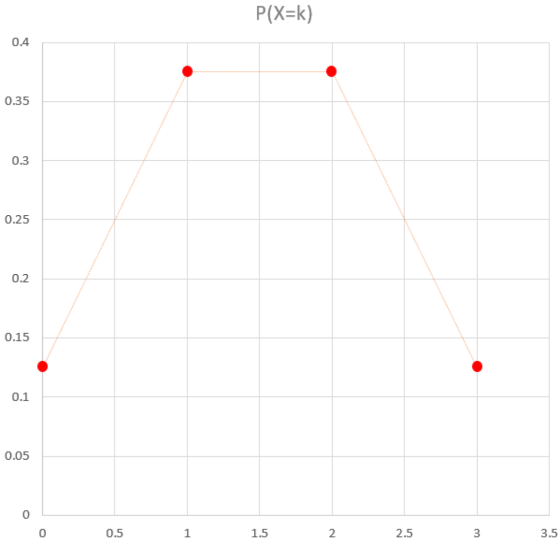
- You toss a coin three times. The coin is biased and $P(Head) = p$. Random variable X is "number of heads in your experiment".
- **Compute PMF**: find $p_X(x)$ for all possible values of x
- $x \in \{0, 1, 2, 3\}$.
- For $x = 0$:
  - Find all possible outcomes for which $X = 0$: $\{TTT\}$
  - Add their probabilities: $P(TTT) = (1 - p)^3$
  - Therefore, $p_X(0) = (1 - p)^3$
- For $x = 1$:
  - Find all possible outcomes for which $X = 1$: $\{TTH, THT, HTT\}$
  - Add their probabilities: $P(TTH) + P(THT) + P(HTT) = 3p(1 - p)^2$
  - Therefore, $p_X(1) = 3p(1 - p)^2$

Probability Mass Function—Example

- For $x = 2$:
  - Find all possible outcomes for which $X = 2$: $\{HTH, THH, HHT\}$
  - Add their probabilities: $P(HTH) + P(THH) + P(HHT) = 3p^2(1-p)$
  - Therefore, $p_X(2) = 3p^2(1-p)$
- For $x = 3$:
  - Find all possible outcomes for which $X = 3$: $\{HHH\}$
  - Add their probabilities: $P(HHH) = p^3$
  - Therefore, $p_X(3) = p^3$
- So we have:
  - $p_X(0) = (1-p)^3$
  - $p_X(1) = 3p(1-p)^2$
  - $p_X(2) = 3p^2(1-p)$
  - $p_X(3) = p^3$
- You can check that: $p_X(0) + p_X(1) + p_X(2) + p_X(3) = 1$

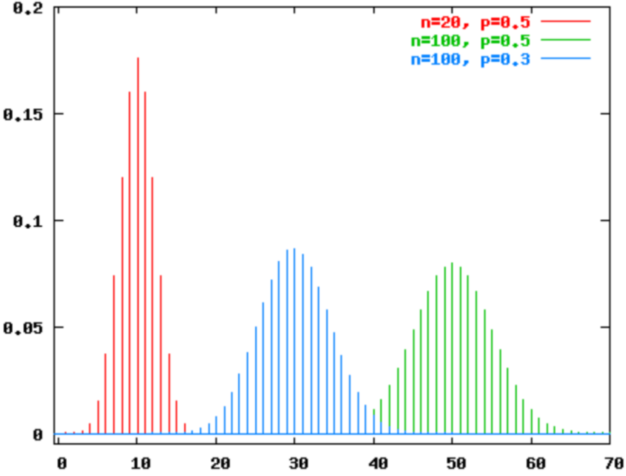
Probability Mass Function—Example

- If $p = \frac{1}{2}$ then: $p_X(0) = (1-p)^3 = \frac{1}{8}$, $p_X(1) = 3p(1-p)^2 = \frac{3}{8}$
  $p_X(2) = 3p^2(1-p) = \frac{3}{8}$, $p_X(3) = p^3 = \frac{1}{8}$



Binomial PMF

- You have a coin and $P(Head) = p$. You toss the coin n times and tosses are independent. Random variable X is number of "heads".
- $p_X(k) = P(X = k) = P(\text{k heads}) = \binom{n}{k} p^k (1-p)^{n-k}$, for $k = 0, 1, 2, \cdots$
- The shape of the binomial distribution depends on the values of n and p.



Expectation

- Average in large number of repetitions of the experiment
- Definition:

$$E[X] = \sum_x x p_X(x)$$

- Example: You are playing a game and you get one dollar reward with probability $\frac{1}{6}$, two dollars reward with probability $\frac{1}{2}$ and four dollars reward with probability $\frac{1}{3}$. If you play this game million times, on "average" how much do you win?
- Random variable $X =$ your reward, $x \in \{1, 2, 4\}$.
- $p_X(1) = P(X = 1) = \frac{1}{6}$
- $p_X(2) = P(X = 2) = \frac{1}{2}$
- $p_X(4) = P(X = 4) = \frac{1}{3}$

Expectation

- Therefore,

$$E[X] = \sum_x x p_X(x) = 1 \cdot p_X(1) + 2 \cdot p_X(2) + 4 \cdot p_X(4)$$

$$\Rightarrow E[X] = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{2} + 4 \cdot \frac{1}{3} = 2.5$$

- **Interpretations**: if you play the games million times, than $\frac{1}{6}$ of the times you win 1 dollar, $\frac{1}{2}$ of the times you win 2 dollar, and $\frac{1}{3}$ of the times you win 4 dollar. Therefore, on average you win 2.5 dollars.

Expectation—Example

- In an experiment, there are $n+1$ equally likely outcomes. Random variable X maps outcomes to $\{0, 1, 2, \cdots, n\}$. What is $E[X] =$?
- Because all outcomes are equally likely, and there are $n+1$ outcomes, then $p_X(i) = P(X = i) = \frac{1}{n+1}$ for all $i = 0, 1, \cdots n$. Therefore,

$$E[X] = \sum_{i=0}^{n} i \cdot \frac{1}{n+1} = \frac{1}{n+1}(1 + 2 + \cdots + n) = \frac{1}{n+1} \frac{n(n+1)}{2} = \frac{n}{2}$$

- **Interpretation**: Expectation of random variable X is "the center of gravity" of PMF of X. In this example, center of gravity is $\frac{n}{2}$ (because of symmetry).