Lecture 7: Conditional Probability - Part IV & Discrete Random Variables

Satyajit Thakor
IIT Mandi

Random Variables

- Usually, we may not be interested in the outcome of an experiment itself but rather in some numerical function of the outcome or an event.
- The idea of random variable enables us to put aside the complex structure of the sample space and its events by assigning a real number to the elements of the sample space.
- Thus we can simply work with numbers rather than complexly defined outcomes and events.
- Given an experiment with sample space S, a random variable (r.v.) is a function from the sample space S to the real numbers $\mathbb{R}$.

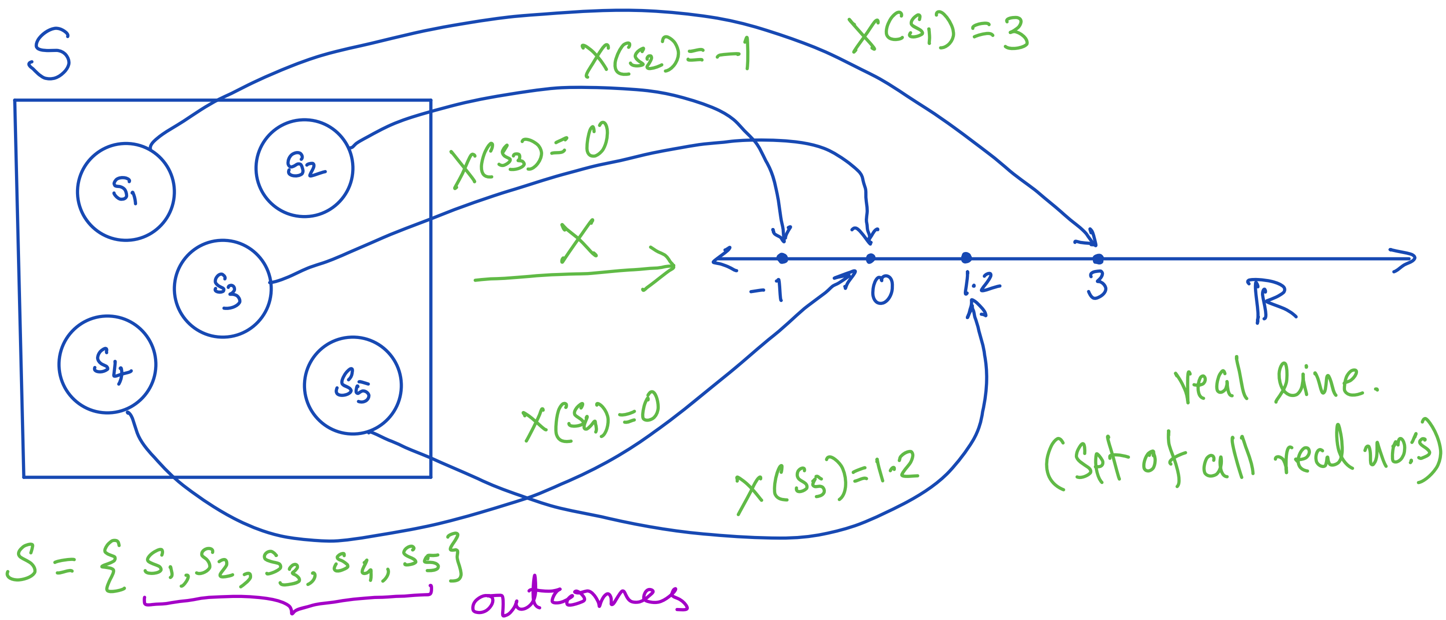
A random variable is a function : $S \longrightarrow \mathbb{R}$.

Random Variables

- It is common to denote random variables by capital letters. For example,

$$X : S \to \mathbb{R}$$

denotes a random variable X and to each $s \in S$ it assigns a numerical value (real number) $X(s)$.



Random Variables

- Example: Consider an experiment where we toss a fair coin twice. The sample space consists of four possible outcomes: $S = \{HH, HT, TH, TT\}$. Here are some random variables on this space. Each r.v. is a numerical summary of some aspect of the experiment.
- Let X be the number of Heads. Then:

$$X(HH) = 2, \quad \underbrace{X(HT) = X(TH) = 1}_{\text{The mapping is not injective in general.}}, \quad X(TT) = 0.$$

- Let Y be the number of Tails.

$$Y(HH) = 0, \; Y(HT) = Y(TH) = 1, \; Y(TT) = 2. \quad \text{i.e., } \underbrace{Y(s) = 2 - X(s)}_{\text{The r.v. } Y \text{ is a function of the r.v. } X.}.$$

- Let I be 1 if the first toss is Heads and 0 otherwise.

$I(HH) = I(HT) = 1,$
$I(TH) = I(TT) = 0.$

Such an r.v. is called an "indicator" r.v. since it indicates whether an event has occurred (1) or not (0).

Random Variables

- Note that, a random variable is neither random nor a variable! It is a function.
- The source of randomness is the choice of an outcome (form S) of the experiment according to the probability function P.
- A random variable X is said to be discrete if there is a finite list of values $a_1, a_2, \ldots, a_n$ or an infinite list of values $a_1, a_2, \ldots$ such that $P(X = a_j \text{ for some } j) = 1$. i.e., $P(X=a_1 \text{ or } X=a_2 \text{ or} \ldots) = 1$
- If X is a discrete r.v., then the finite or countably infinite set of values x such that $P(X = x) > 0$ is called the support of X.
- Example: In the "Two-coin" experiment

  - $P(X=0 \text{ or } X=1 \text{ or } X=2) = 1$.
  - The set $\{0,1,2\}$ is finite. Hence, X is a discrete random variable.

Random Variables

- Given a random variable, we would like to be able to describe its behaviour using the language of probability.
- For example, we might want to answer questions about the probability that the r.v. will fall into a given range: if L is the lifetime earnings of a randomly chosen U.S. college graduate, what is the probability that L exceeds a million dollars?

  i.e., find $P(L > 1000000)$.
- The distribution of a random variable provides the answers to such questions.
- There are several equivalent ways to express the distribution of an r.v., e.g., cumulative distribution function, probability mass function.
- For a discrete r.v., the most natural way to do so is with a probability mass function.

Random Variables

i.e., $p_X : \mathbb{R} \to [0,1]$.

- $\boxed{\text{The probability mass function (PMF) of a discrete r.v. } X \text{ is the function } p_X \text{ given by } p_X(x) = P(X = x)\text{. Note that this is positive if } x \text{ is in the support of } X\text{, and 0 otherwise.}}$
- Note that the event that X takes the value x, i.e.,

The subscript X is used to specify that the PMF is for X (not for some another r.v. Y)

$$\{X = x\}$$

e.g., $\{X=1\}$

is equivalent to the event that "s occurred where $X(s) = x$", i.e.,

$$\{s \in S : X(s) = x\}.$$

e.g., $\{HT, TH\}$ (s_1, s_2) $\left(\because \begin{array}{l} X(HT)=1 \\ X(TH)=1 \end{array}\right)$

- But the notation $\{X = x\}$ is shorter, more convenient and does not involve reference to the sample space.
- Example: In the "Two-coins" example $\{X = 1\}$ refers to $\{HT, TH\}$, i.e., the event that "the outcome is HT or TH".

Random Variables

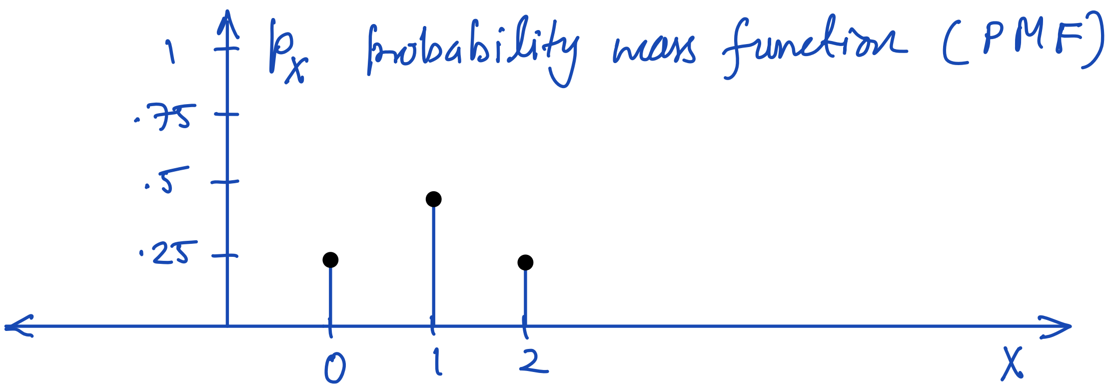
- Example: Find the PMFs of all the random variables in the "Two-coins" example considering fair coins.
- X: the number of Heads.

$p_X(0) = P(X=0) = 1/4,$

$p_X(1) = P(X=1) = 1/2,$

$p_X(2) = P(X=2) = 1/4,$

and $p_X(x) = 0$ for all other $x \in \mathbb{R}$.



Random Variables

- Y: the number of Tails.

$$p_Y(0) = 1/4, \quad p_Y(1) = 1/2, \quad p_Y(2) = 1/4.$$

- I: 1 if the first toss is Heads and 0 otherwise.

$$p_I(0) = 1/2, \quad p_I(1) = 1/2.$$

- Homework: Plot PMFs for Y and I.

Conditional probability

Monty Hall problem

- Example: On a TV game show, hosted by Monty Hall, a contestant chooses one of three closed doors, two of which have a goat behind them and one of which has a car. Monty, who knows where the car is, then opens one of the two remaining doors. The door he opens always has a goat behind it (he never reveals the car!). If he has a choice, then he picks a door at random with equal probabilities. Monty then offers the contestant the option of switching to the other unopened door.
- If the contestant's goal is to get the car, should she switch doors?

- Let W be the event that she wins.
- We want to find the best of the two strategies to win (i.e., the one with the higher prob. of winning)
- Strategies: ① switch, ② No switch.

Conditional probability

- Let C_i be the event that the car is behind door i.
- Assume without loss of generality that she picks the the door 1 (else we can relabel/rename the doors)
- Then:

① $P(W) = \underbrace{P(W|C_1)}_{\text{since switch from door 1 (given that the car is behind door 1)}} \cdot P(C_1) + P(W|C_2) \cdot P(C_2) + P(W|C_3) P(C_3)$

$= 0 \cdot 1/3 + 1 \cdot 1/3 + 1 \cdot 1/3 = 2/3.$

② $P(W) = P(W|C_1) \cdot P(C_1) + P(W|C_2) \cdot P(C_2) + P(W|C_3) P(C_3)$

$= 1 \cdot 1/3 + 0 \cdot 1/3 + 0 \cdot 1/3 = 1/3.$

- Hence, to switch is the best strategy (from ① & ②)