LECTURE: DISCRETE RANDOM VARIABLES

It is not unusual to assign a value to a certain event, like the amount you win when you play the lottery, or the number of heads when you toss 10 coins. This is called a **random variable**

1. Random Variables

Definition:

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

It's generally a quantity we wish to measure.

Definition:

A **discrete random variable** Y is a random variable which can only take on a finite or countable set of distinct values.

Note: For countable set, think for example taking values in $\{0, 1, 2, 3, \cdots\}$

In this section, we will mainly focus on discrete random variables.

Here are some examples of random variables:

(1) The number of people in Providence who prefer chocolate to vanilla ice cream

(2) The number of defective light bulbs out of a shipment of 1000 light bulbs.

(3) The number of times you play a slot machine in Las Vegas until you win.

Example 1:

Let S be the sample space representing the flip of two fair coins.

Let Y = the number of heads flipped.

Then Y is a discrete random variable, since it can only have the values 0, 1, or 2.

	H	T
H	2	1
T	1	0

Notation: Uppercase letters (such as Y) are used to designate random variables whereas lowercase letters (such as y) represent values that a random variable can take.

The expression $(Y = y)$ is shorthand for the set of all points for which the random variable Y takes on the value y.

Example: In the two-coin-toss problem, the possible values of Y are 0, 1, and 2, so we have:

$$(Y = 0) = \{(T, T)\}$$

$$(Y = 1) = \{(H, T), (T, H)\}$$

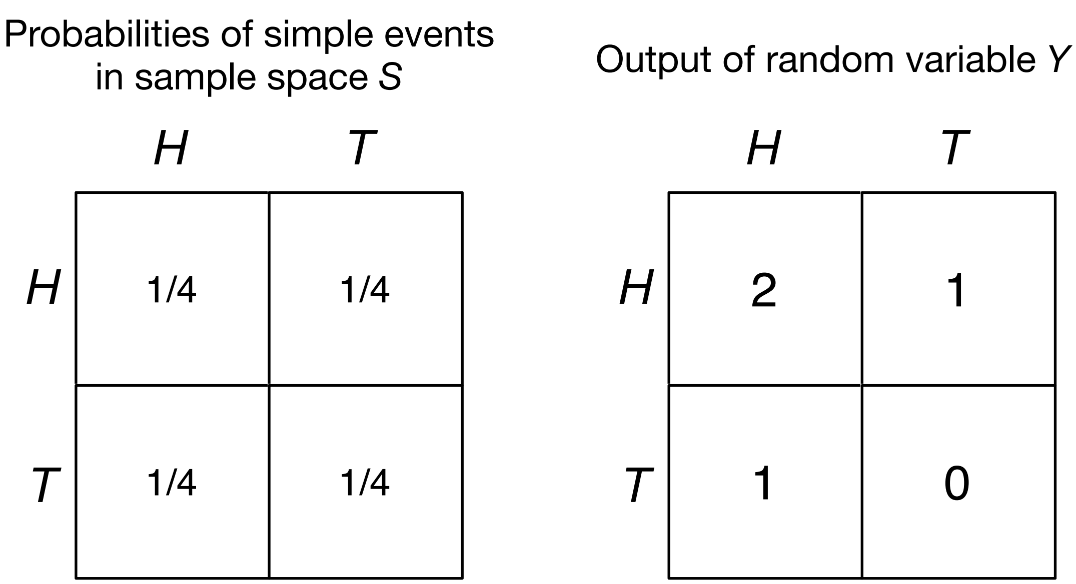
$$(Y = 2) = \{(H, H)\}$$

Upshot: Since $(Y = y)$ is an event in our sample space, we can talk about its probability, i.e. $P(Y = y)$. In fact, the point of random variables is to do just this!

Definition:

$P(Y = y)$ is the probability that Y takes the value y

Example: Back to our two-coin-toss problem, let's look at the probabilities of the random variable Y. Each simple event in our sample space has probability 1/4.



$$P(Y = 0) = 1/4$$
$$P(Y = 1) = 1/4 + 1/4 = 1/2$$
$$P(Y = 2) = 1/4$$

Definition:

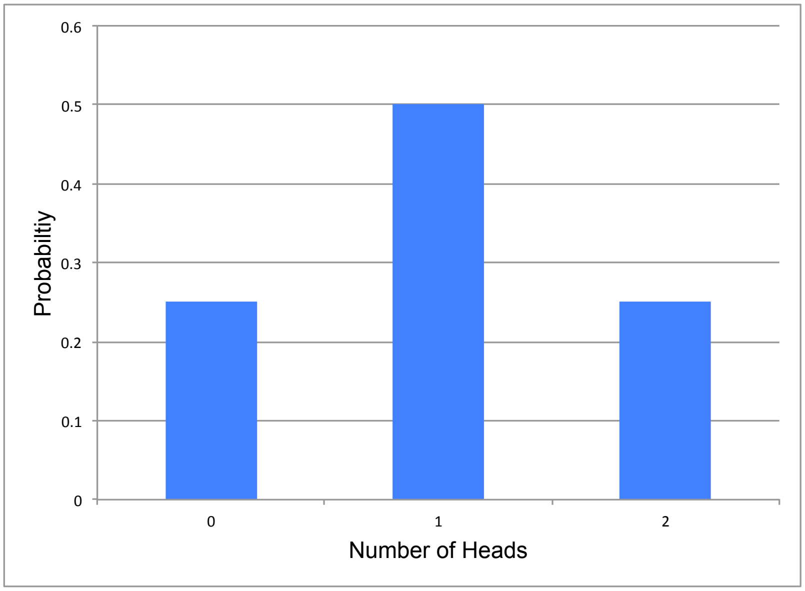
The **probability mass function** (pmf) is $p(y) = P(Y = y)$, that is the probability that Y takes the value y

The pmf can be represented as a function, table, or graph which gives the values $p(y) = P(Y = y)$ for all possible values y which Y can take.

Example: In the two-coin-flip example, we can represent the pmf of Y in a table:

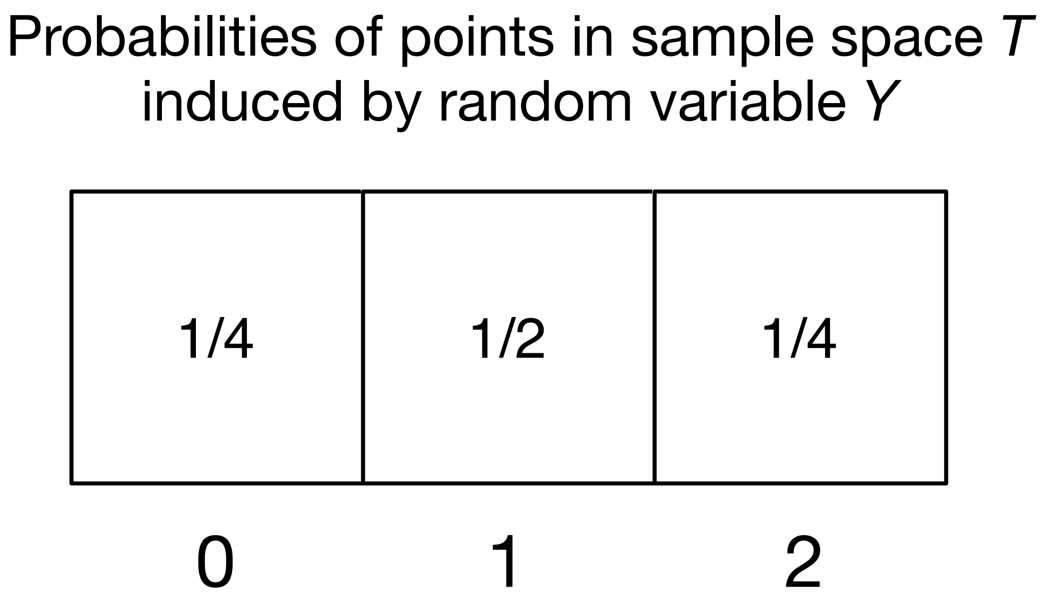
y	$p(y)$
0	1/4
1	1/2
2	1/4

We can also represent the pmf graphically as a histogram



Important Observation: The random variable Y induces/creates a probability distribution on the sample space of outputs $T = \{0, 1, 2\}$.

More precisely, the probabilities of the sample points in T are the probabilities $p(y)$ for $y = 0, 1, 2$. We can illustrate this new sample space in a picture.



Definition:

T is called the **sample space induced by a random variable**

More often or not, we care much more about the sample space T than the original sample space Y. Think for example in gambling, where you care more about the money you'll win/lose than the actual outcome of your game

Since a discrete random variable induces a probability distribution, the following must be true.

Theorem:

For any discrete random variable Y:

$$0 \leq p(y) \leq 1 \text{ for all } y$$
$$\sum_y p(y) = 1$$

Example 2:

Let S be the sample space representing the rolls of two six-sided dice. Consider the following two random variables:

(1) X = the sum of the two dice

(2) Y = the larger of the two die rolls

Two random variables on the sample space of two die rolls

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

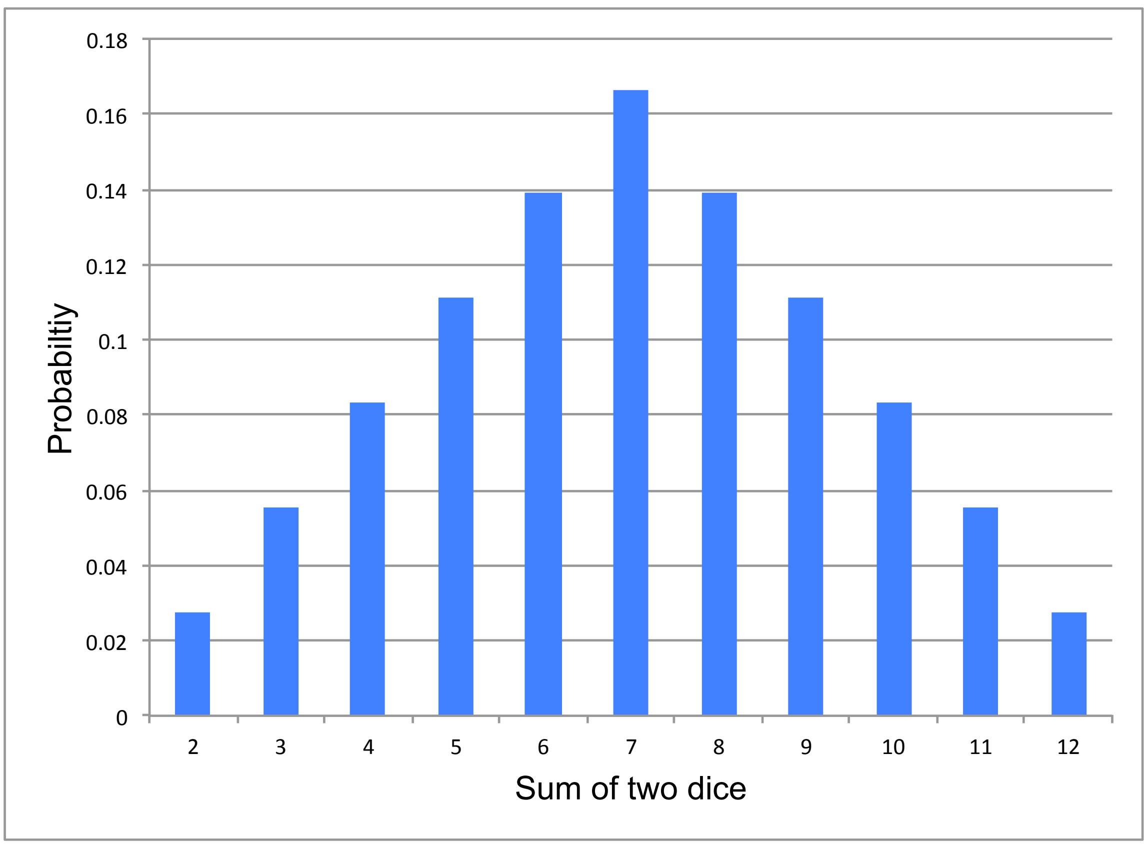
X

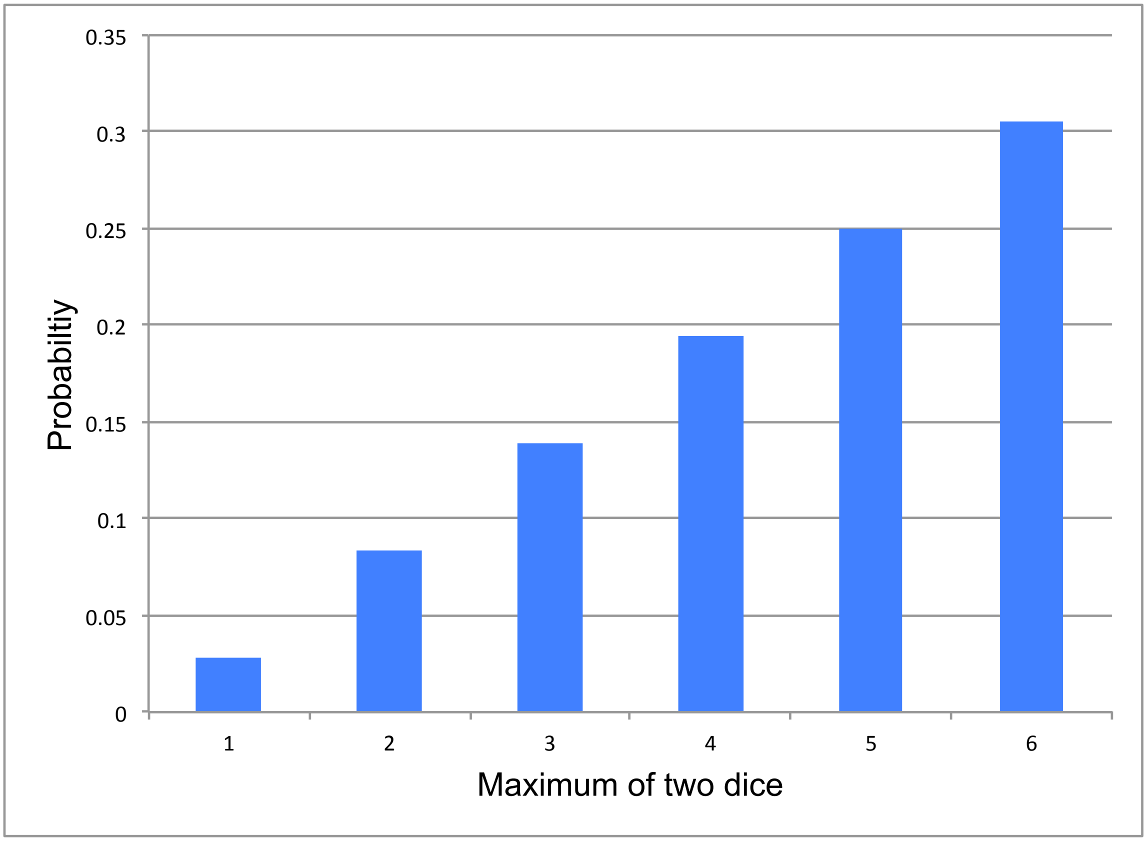
	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	2	3	4	5	6
3	3	3	3	4	5	6
4	4	4	4	4	5	6
5	5	5	5	5	5	6
6	6	6	6	6	6	6

Y

The random variable X induces a probability distribution on the set of integers $\{2, 3, 4, \ldots, 12\}$, and the random variable Y induces a probability distribution on the set of integers $\{1, 2, 3, 4, 5, 6\}$.

Let's look at the pmfs of both random variables using histograms.





Both of these distributions are nonuniform, even though the underlying distribution of the two dice is uniform. .

We can also write the pmfs in table form. For the random variable Y, we have:

y	$p(y)$
1	$1/36$
2	$3/36$
3	$5/36$
4	$7/36$
5	$9/36$
6	$11/36$

The pmf for X can be expressed similarly.

2. Expected Value

Given a discrete random variable, we can definite its mean, or expected value.

Definition:

The **expected value/mean** of a discrete random variable Y is

$$E(Y) = \sum_y yP(Y = y)$$

Where the sum is taken over all possible values y can take.

We can think of the expected value as a weighted average of the values of Y with each possible output y weighted by its probability $p(y)$.

Other Interpretation: Think of a random variable Y as an observation from an experiment. Suppose we perform the experiment n times, and observe n values of Y, which we shall designate $y_1, y_2, \ldots, y_n$. Then for large n,

$$\frac{y_1 + y_2 + \cdots + y_n}{n} \approx E(Y)$$

where the approximation "gets better" as n gets larger, i.e. as we perform more experiments.

The quantity on the left hand side is known as the *empirical mean* or *sample mean* and looks like what we likely learned in high school. The expected value is, in a sense, the limit of the empirical mean as the sample size approaches infinity. We will make this more precise later in the course, but this is a good concept to keep in mind.

Example 3:

Roll a 6-sided die, so $S = \{1, 2, 3, 4, 5, 6\}$

And let X be simply the number that you get, so if you roll a 5, then $X = 5$

Then the expected value of X is:

$$E(X) = \sum_{x=1}^{6} xP(X = x) = \sum_{x=1}^{6} x\left(\frac{1}{6}\right) = \frac{1}{6}\sum_{x=1}^{6} x = \frac{21}{6} = 3.5$$

Where used that $P(X = x) = 1/6$ for all x

Note that the expected value of 3.5 is not a possible value of X, i.e. we cannot roll a 3.5 on a single die. Given our "long term average" interpretation, this is saying that we expect the empirical average to approach 3.5 as the number of rolls increases, not that a 3.5 is the most likely die roll.

Example 4:

Let Y be the random variable above representing the maximum of two dice. What is the expected value of Y?

To find the expected value, we do a weighted average using the probabilities in the table above.

$$\begin{aligned} E(Y) &= \sum_{y=1}^{6} yP(Y = y) \\ &= 1\left(\frac{1}{36}\right) + 2\left(\frac{3}{36}\right) + 3\left(\frac{5}{36}\right) + 4\left(\frac{7}{36}\right) + 5\left(\frac{9}{36}\right) + 6\left(\frac{11}{36}\right) \\ &= \frac{1 + 6 + 15 + 28 + 45 + 66}{36} \\ &= \frac{161}{36} \approx 4.47 \end{aligned}$$

3. Properties of Expectation

Linearity of Expectation:

Let X and Y be two random variables and let a and b be constants. Then

$$E(aX + bY) = aE(X) + bE(Y)$$

Corollary:

If $X_1, X_2, \ldots, X_n$ are random variables, then

$$E(X_1 + X_2 + \cdots + X_n) = E(X_1) + E(X_2) + \cdots + E(X_n)$$

Linearity of expectation is a really nice property since it does not require the random variables to be independent.