Random Variables And Distributions

Consider the following random experiments:

- Tossing a fair coin and observing H/T
- Rolling a die and observing even/odd
- Drawing a marble from a bag with 5 Red and 5 White, and observing the color
- Picking a number at random from the set $\{0, 1\}$

Q: What do these experiments have in common?

A: They are the same, in the sense that they have the same **probability distribution** (for each experiment, we have two possible outcomes, and each of those outcomes occur with probability 1/2).

Random Variables

We would like to capture the essence of random experiments with a mathematical model that we can study.

To do so, we associate numbers to the outcomes. This association is called a **random variable** — the value that the variable takes varies depending on the (random) outcome of the experiment.

We usually denote random variables by capital letters X, Y (or $X_1, X_2, X_3, \ldots, X_n$ for multiple random variables).

Example: For tossing a coin *once*, let X be the number of H appearing on that coin toss. When the coin comes up Heads, $X = 1$, and when the coin comes up Tails, $X = 0$. We have: $P(X = 0) = 0.5$ and $P(X = 1) = 0.5$.

Examples Of Random Variables

Example: For tossing a coin *twice*, let X be the number of H that we get. The following table shows the value of X associated with each of the 4 outcomes:

Outcome	X (# Heads)
HH	2
HT	1
TH	1
TT	0

Compute the probabilities of $X = 0$, $X = 1$, and $X = 2$.

$P(X = 0) = P(TT) = 0.25$

$P(X = 1) = P(HT \text{ or } TH) = 0.50$

$P(X = 2) = 0.25$

Examples Of Random Variables

Example: Roll a single die. Let X be the number that comes up. What are the possible values for X, and what are their corresponding probabilities?

The possible values for X are $\{1, 2, 3, 4, 5, 6\}$. The probabilities are outlined in the table below:

X	P(X)
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	1/6

Note: When observing a random experiment, some type measurement is taken. This measurement can be thought of as the outcome of the random variable.

Examples Of Random Variables

Example: Roll a pair of dice, and let X be the **sum** of the two numbers that come up. What are the possible values for X? Compute the probabilities:

$$P(X = 2) \qquad P(X = 4) \qquad P(X = 7) \qquad P(X = 11)$$

The possible values for X are $\{2, 3, 4, \ldots, 12\}$.

$P(X = 2) = P(1, 1) = 1/36.$

$P(X = 4) = P(1, 3) + P(2, 2) + P(3, 1) = 3/36.$

$P(X = 7) = 6/36.$

$P(X = 11) = 2/36.$

Examples Of Random Variables

Example: A bag of marbles contains 3 Red and 5 White. Select 2 marbles at random, and let X denote the number of Red marbles. What are the possible values of X, and their corresponding probabilities?

Possible values: $\{0, 1, 2\}$. Probabilities (add up to 1):

X	P(X)
0	$\frac{\binom{3}{0}\binom{5}{2}}{\binom{8}{2}} \approx 0.357$
1	$\frac{\binom{3}{1}\binom{5}{1}}{\binom{8}{2}} \approx 0.536$
2	$\frac{\binom{3}{2}\binom{5}{0}}{\binom{8}{2}} \approx 0.107$

Discrete And Continuous Random Variables

Consider the following experiments, and the defined random variables for each:

- Tossing 5 coins. $X = \#$ of heads
- Grading an exam. $X = \#$ of passing scores
- Cars entering highway. $X = \#$ cars in one hour
- Randomly chosen student. $X =$ student's height
- Lifespan of Monarch butterfly
- Fuel economy of a car (mpg)

The first 3 are called **discrete** random variables.

The last 3 are called **continuous** random variables.

Probability Distributions

We will focus on discrete random variables (studying continuous RV's requires calculus!).

We already saw that for discrete random variable X, we can associate to each of its possible values the **probability** with which that value occurs.

This is called the **probability distribution** of the random variable.

Note: A probability distribution must obey the same rules as probabilities we've been studying so far:

- $0 \leq P(X) \leq 1$ for each possible value of X
- the total probability must add up to 1

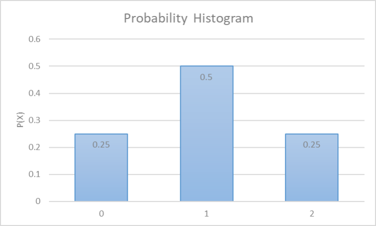
Probability Distributions

Example: Toss 2 coins and let X be the # of Heads. The probability distribution of X can be represented as a piecewise function:

$$P(X = k) = \begin{cases} 0.25 & k = 0 \\ 0.50 & k = 1 \\ 0.25 & k = 2 \end{cases}$$

It is also possible to represent this as a table, or as a **probability histogram**:

X	P(X)
0	0.25
1	0.50
2	0.25



Probability Distributions

Example: Consider a random variable Y with the following distribution:

Y	P(Y)
0	0.40
1	0.30
2	0.15
3	0.10
4	0.03
5	0.02

What is $P(Y \geq 3)$? What about $P(Y < 3)$?

$P(Y \geq 3) = P(Y = 3, 4, 5) = 0.10 + 0.03 + 0.02 = 0.15$

Notice that $Y < 3$ is the complement of $Y \geq 3$, so we can simply do $1 - 0.15 = 0.85$.

Probability Distributions

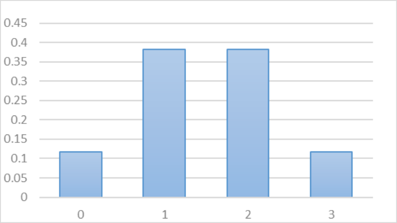
Example: Select 3 cards at random from a standard deck, and let X be the number of black cards. Find the probability distribution of X. First, $X \in \{0, 1, 2, 3\}$.

$$P(X = 0) = \frac{\binom{26}{0}\binom{26}{3}}{\binom{52}{3}} \approx 0.118$$

$$P(X = 1) = \frac{\binom{26}{1}\binom{26}{2}}{\binom{52}{3}} \approx 0.382$$

$$P(X = 2) = \frac{\binom{26}{2}\binom{26}{1}}{\binom{52}{3}} \approx 0.382$$

$$P(X = 3) = \frac{\binom{26}{3}\binom{26}{0}}{\binom{52}{3}} \approx 0.118$$



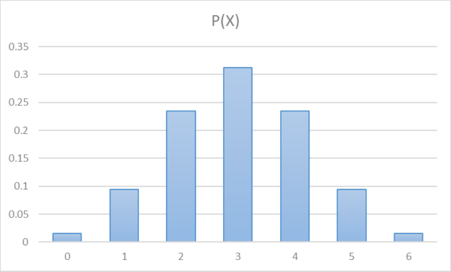
Probability Distributions

Example: Toss a coin 6 times, and let X be the number of Heads you observe. Then the probability that X is equal to k (between 0 and 6) is given by the formula

$$P(X = k) = \frac{\binom{6}{k}}{2^6} \quad \text{for} \quad 0 \leq k \leq 6$$

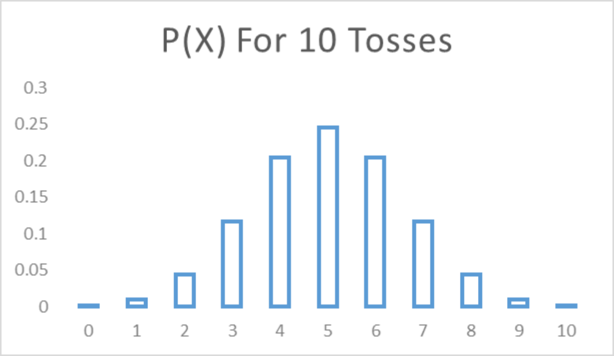
Using this, we can quickly construct the probability table and the probability histogram, so we can visualize the distribution better:

X	P(X)
0	0.016
1	0.094
2	0.234
3	0.312
4	0.234
5	0.094
6	0.016



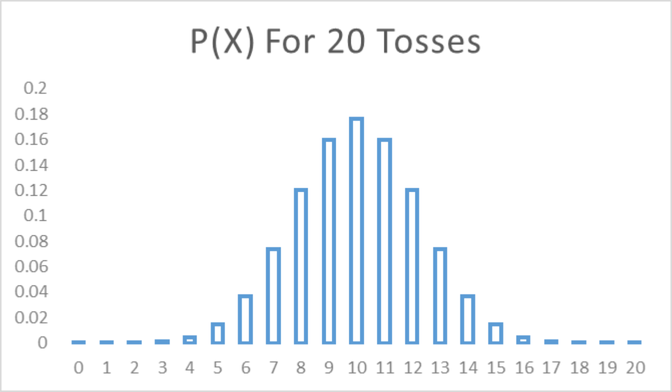
Probability Histograms

Example: The following histogram shows the probability distribution for X = number of Heads on 10 coin tosses.



Probability Histograms

Example: The following histogram shows the probability distribution for X = number of Heads on 20 coin tosses.



Note: As the number of coins increases, the peaks start to resemble a bell curve!

An Infinite Sum

Consider the following **equality**:

$$\boxed{\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \frac{1}{32} + \cdots\cdots = 1}$$

where the sum on the LHS has **infinitely many terms**!

Q: How can this possibly be true?

A: Consider the experiment: toss a coin until you see the first Heads, and let X be the number of tosses required.

The possible values for X are going to be $\{1, 2, 3, 4, 5, \ldots\}$, and the LHS is the sum of the probabilities of these values! Since the total probability must be 1, the equality holds.