Probability

Probability is difficult to formulate in precise mathematical terms (though it is well worth doing so!), but we can make do with some intuitive characterizations.

Probability

Probability is difficult to formulate in precise mathematical terms (though it is well worth doing so!), but we can make do with some intuitive characterizations.

Definition 1. *Let X be the outcome of some event.*

- *The outcome of X is a* random event *if we cannot predict it, but the distribution of such outcomes has a regular pattern over many repetitions.*

Probability

Probability is difficult to formulate in precise mathematical terms (though it is well worth doing so!), but we can make do with some intuitive characterizations.

Definition 1. *Let X be the outcome of some event.*

- *The outcome of X is a* random event *if we cannot predict it, but the distribution of such outcomes has a regular pattern over many repetitions.*

- X *is called a* random variable.

- X *is called a* random variable.
- *The* sample space *is the collection all possible values* X *could take.*

- X *is called a* random variable*.*
- *The* sample space *is the collection all possible values* X *could take.*
- *The* probability *of a certain outcome is the proportion of times it occurs after many repetitions.*

- X *is called a* random variable.
- *The* sample space *is the collection all possible values* X *could take.*
- *The* probability *of a certain outcome is the proportion of times it occurs after many repetitions.*

Example 2. *We flip a coin. It comes up heads with probability .5. Does this mean if we flip the coin an even number of times, we get* $\frac{1}{2}$ *heads?*

- X *is called a* random variable*.*

- *The* sample space *is the collection all possible values X could take.*

- *The* probability *of a certain outcome is the proportion of times it occurs after many repetitions.*

Example 2. *We flip a coin. It comes up heads with probability .5. Does this mean if we flip the coin an even number of times, we get $\frac{1}{2}$ heads? Does this even mean that over N flips, as N gets large the number of heads approaches $N/2$?*

- X *is called a* random variable.
- *The* sample space *is the collection all possible values* X *could take.*
- *The* probability *of a certain outcome is the proportion of times it occurs after many repetitions.*

Example 2. *We flip a coin. It comes up heads with probability .5. Does this mean if we flip the coin an even number of times, we get* $\frac{1}{2}$ *heads? Does this even mean that over* N *flips, as* N *gets large the number of heads approaches* $N/2$*?*

Probability formalities

Probability formalities

While we have yet to completely define probability, we can still list properties which it must obey.

Probability formalities

While we have yet to completely define probability, we can still list properties which it must obey. (These are topics from Chapter 11 which we cover briefly.)

Probability formalities

While we have yet to completely define probability, we can still list properties which it must obey. (These are topics from Chapter 11 which we cover briefly.)

- For any event A, $0 \le P(A) \le 1$. Probability 1 means A is certain to happen, probability 0 means A is certain not to happen.

Probability formalities

While we have yet to completely define probability, we can still list properties which it must obey. (These are topics from Chapter 11 which we cover briefly.)

- For any event A, $0 \leq P(A) \leq 1$. Probability 1 means A is certain to happen, probability 0 means A is certain not to happen.

- If there are D outcomes in the sample space which are a priori equally likely, then the chance of achieving one

of N of these outcomes is $\frac{N}{D}$.

of N of these outcomes is $\frac{N}{D}$.

- $P(A \text{ doesn't happen}) = 1 - P(A)$.

of N of these outcomes is $\frac{N}{D}$.

- $P(A$ doesn't happen$) = 1 - P(A)$.
- If event A and event B have no outcomes in common,

$$P(A \text{ or } B) = P(A) + P(B).$$

of N of these outcomes is $\frac{N}{D}$.

- $P(A$ doesn't happen$) = 1 - P(A)$.
- If event A and event B have no outcomes in common,

$$P(A \text{ or } B) = P(A) + P(B).$$

- If the outcome of event X is unrelated to the outcome of event Y (they are *independent*) then

$$P(X = A \text{ and } Y = B) = P(X = A) \times P(Y = B)$$

Example 3. *You roll two dice; each die has an equal probability* $(1/6)$ *of showing any number out of* $\{1, 2, 3, 4, 5, 6\}$*. What is the probability of getting a 12?*

Example 3. *You roll two dice; each die has an equal probability* $(1/6)$ *of showing any number out of* $\{1, 2, 3, 4, 5, 6\}$*. What is the probability of getting a 12? An 11?*

Example 3. *You roll two dice; each die has an equal probability* $(1/6)$ *of showing any number out of* $\{1,2,3,4,5,6\}$. *What is the probability of getting a 12? An 11? A 7?*

Example 3. *You roll two dice; each die has an equal probability* $(1/6)$ *of showing any number out of* $\{1, 2, 3, 4, 5, 6\}$*. What is the probability of getting a 12? An 11? A 7?*

Example 4. *Toss two coins (a nickel and a penny). If we are told that at least one came up heads, what is the probability of the other coming up heads?*

Example 5. *Monty Hall easy version. There are 3 doors; you don't know what is behind any of them. You are told that there is a car behind one door, and a goat behind the other two. You get whatever is behind the door you choose as a prize. So you pick a door at random. What is your chance of getting a car?*

Example 6. [Monty Hall problem] *Monty Hall wants to confuse you a little bit. You get to pick a door.* Then *(whichever door you pick, whether there is a goat behind it or a car) Monty opens a different door that has a goat behind it. (He can always do this, even if you chose a goat door, because there are 2 goats).*

Now Monty gives you the opportunity to switch doors to the other unopened door. Should you switch?

More flipping coins

More flipping coins

The example of flipping coins is as simple as possible but lets us see many of the features which are prominent in the general case.

Example 7. *Do 4 flips. Let X be the outcome.*

- *What is the sample space?*

More flipping coins

The example of flipping coins is as simple as possible but lets us see many of the features which are prominent in the general case.

Example 7. *Do 4 flips. Let X be the outcome.*

- *What is the sample space?*
- *What is the probability of each outcome? (Key point: each outcome equally likely).*

More flipping coins

The example of flipping coins is as simple as possible but lets us see many of the features which are prominent in the general case.

Example 7. *Do 4 flips. Let X be the outcome.*

- *What is the sample space?*
- *What is the probability of each outcome? (Key point: each outcome equally likely).*

- *Let Y be the random variable giving the* number of heads.
- *What is the sample space?*

- *Let Y be the random variable giving the* number of heads.
- *What is the sample space?*
- *How many outcomes in the original sample space give each Y?*

- *Let Y be the random variable giving the* number of heads.
- *What is the sample space?*
- *How many outcomes in the original sample space give each Y?*
- *What are the probabilities of each Y in this sample space?*

- *Let Y be the random variable giving the* number of heads.
- *What is the sample space?*
- *How many outcomes in the original sample space give each Y?*
- *What are the probabilities of each Y in this sample space?*

Summary of answers to last part above:

Y	0	1	2	3	4
# Outcomes	1	4	6	4	1
Probability	.0625	.25	.375	.25	.0625

Digression: The number of possible ways to get Y heads for N coin flips

Digression: The number of possible ways to get Y heads for N coin flips

We now digress into one of the most beautiful areas of elementary mathematics. Our digression will quickly lead us back to very familiar territory! The number of ways one can get Y heads in N coin flips is well-studied in mathematics, first by Pascal. It is called a *binomial coefficient*, denoted $\binom{N}{Y}$.

Digression: The number of possible ways to get Y heads for N coin flips

We now digress into one of the most beautiful areas of elementary mathematics. Our digression will quickly lead us back to very familiar territory! The number of ways one can get Y heads in N coin flips is well-studied in mathematics, first by Pascal. It is called a *binomial coefficient*, denoted $\binom{N}{Y}$.

Example 8. *Write down the first few binomial coefficients in a triangle (Pascal's triangle).*

Digression: The number of possible ways to get Y heads for N coin flips

We now digress into one of the most beautiful areas of elementary mathematics. Our digression will quickly lead us back to very familiar territory! The number of ways one can get Y heads in N coin flips is well-studied in mathematics, first by Pascal. It is called a *binomial coefficient*, denoted $\binom{N}{Y}$.

Example 8. *Write down the first few binomial coefficients in a triangle (Pascal's triangle). What patterns*

do you notice:

do you notice:

- *What do the rows add up to?*

do you notice:

- *What do the rows add up to?*
- *Are they symmetric?*

do you notice:

- *What do the rows add up to?*
- *Are they symmetric?*
- *How can you find the values in a row from the previous row?*

do you notice:

- *What do the rows add up to?*
- *Are they symmetric?*
- *How can you find the values in a row from the previous row?*
- *What are other immediate patterns?*

do you notice:

- *What do the rows add up to?*
- *Are they symmetric?*
- *How can you find the values in a row from the previous row?*
- *What are other immediate patterns? (Odds and evens)*

do you notice:

- *What do the rows add up to?*
- *Are they symmetric?*
- *How can you find the values in a row from the previous row?*
- *What are other immediate patterns? (Odds and evens)*

Now for the fun part: plot the values of the rows as a histogram, where the blocks each have area $\frac{1}{2^N}$, so that the total area is one.

the total area is one. What do we see? - The Bell Curve is approximating the rows of Pascal's triangle!!! (and thus modeling coin flips!).

the total area is one. What do we see? - The Bell Curve is approximating the rows of Pascal's triangle!!! (and thus modeling coin flips!). Or conversely, the numbers in Pascal's triangle are telling us about the Bell Curve (normal distribution)!

Back to coin flips

Example 9. *Let Y be the number of heads after 10 successive coin flips. The probabilities for Y are:*

Back to coin flips

Example 9. *Let Y be the number of heads after 10 successive coin flips. The probabilities for Y are:*

Y	*0*	*1*	*2*	*3*	*4*	*5*	*6*	*7*
Probability	*.001*	*.010*	*.044*	*.117*	*.205*	*.246*	*.205*	*.11*

Back to coin flips

Example 9. *Let Y be the number of heads after 10 successive coin flips. The probabilities for Y are:*

Y	0	1	2	3	4	5	6	7
Probability	.001	.010	.044	.117	.205	.246	.205	.11

What is the probability of between 3 and 7 heads?

Back to coin flips

Example 9. *Let Y be the number of heads after 10 successive coin flips. The probabilities for Y are:*

Y	0	1	2	3	4	5	6	7
Probability	.001	.010	.044	.117	.205	.246	.205	.11[illegible]

What is the probability of between 3 and 7 heads?

Example 10. *All 80 of us flip coins at once. Let Y be the total number of heads.*

Back to coin flips

Example 9. *Let Y be the number of heads after 10 successive coin flips. The probabilities for Y are:*

Y	0	1	2	3	4	5	6	7
Probability	.001	.010	.044	.117	.205	.246	.205	.11[illegible]

What is the probability of between 3 and 7 heads?

Example 10. *All 80 of us flip coins at once. Let Y be the total number of heads.*

- *We probably won't get exactly 40 heads. The probability of that is .089*

- *We probably won't get exactly 40 heads. The probability of that is .089*
- *But* $P(35 \leq Y \leq 45) = .781$

- *We probably won't get exactly 40 heads. The probability of that is .089*

- *But* $P(35 \leq Y \leq 45) = .781$ *and*

- $P(30 \leq Y \leq 50) = .982.$

- *We probably won't get exactly 40 heads. The probability of that is .089*
- *But* $P(35 \leq Y \leq 45) = .781$ *and*
- $P(30 \leq Y \leq 50) = .982.$

Note that as the number of trials gets very large, the probability of any one outcome gets very small. It starts being less relevant to ask for example what is the probability that you would get exactly 31 heads out of 80.

- *We probably won't get exactly 40 heads. The probability of that is .089*
- *But* $P(35 \leq Y \leq 45) = .781$ *and*
- $P(30 \leq Y \leq 50) = .982.$

Note that as the number of trials gets very large, the probability of any one outcome gets very small. It starts being less relevant to ask for example what is the probability that you would get exactly 31 heads out of 80. The answer is just too small to matter.

- *We probably won't get exactly 40 heads. The probability of that is .089*

- *But* $P(35 \leq Y \leq 45) = .781$ *and*

- $P(30 \leq Y \leq 50) = .982.$

Note that as the number of trials gets very large, the probability of any one outcome gets very small. It starts being less relevant to ask for example what is the probability that you would get exactly 31 heads out of 80. The answer is just too small to matter. What works better is asking the chances that the number of heads will be in some range.

better is asking the chances that the number of heads will be in some range. This change of emphasis (from “individual” measurement to ranges) is important when we look at continuous variables.

better is asking the chances that the number of heads will be in some range. This change of emphasis (from "individual" measurement to ranges) is important when we look at continuous variables.

Continuous (vs. discrete) random variables.

- If a random variable X takes on a finite number of possible values, it is called a *discete* random variable.

Continuous (vs. discrete) random variables.

- If a random variable X takes on a finite number of possible values, it is called a *discete* random variable. All of our examples so far have been discrete.

Continuous (vs. discrete) random variables.

- If a random variable X takes on a finite number of possible values, it is called a *discete* random variable. All of our examples so far have been discrete.

- If a random variable X takes on a *range* of possible values, it is called a *continuous* random variable.

Continuous (vs. discrete) random variables.

- If a random variable X takes on a finite number of possible values, it is called a *discete* random variable. All of our examples so far have been discrete.

- If a random variable X takes on a *range* of possible values, it is called a *continuous* random variable.

- The sample space is still the collection of possible values for X.

Continuous (vs. discrete) random variables.

- If a random variable X takes on a finite number of possible values, it is called a *discete* random variable. All of our examples so far have been discrete.
- If a random variable X takes on a *range* of possible values, it is called a *continuous* random variable.
- The sample space is still the collection of possible values for X.

- It no longer makes sense to ask what $P(X = a)$ is, but rather what is $P(a \leq X \leq b)$.

- It no longer makes sense to ask what $P(X = a)$ is, but rather what is $P(a \leq X \leq b)$.

Example 11. *Let* X *be a number randomly chosen between* 0 *and* 1. *(Note that it is hard to randomly choose numbers in this fashion.)*

- It no longer makes sense to ask what $P(X = a)$ is, but rather what is $P(a \leq X \leq b)$.

Example 11. *Let X be a number randomly chosen between 0 and 1. (Note that it is hard to randomly choose numbers in this fashion.)*

- *What is $P(X = .3)$*

- *If we assume all choices between 0 and 1 are equally likely, what is $P(X \geq .5)$.*

- It no longer makes sense to ask what $P(X = a)$ is, but rather what is $P(a \leq X \leq b)$.

Example 11. *Let X be a number randomly chosen between 0 and 1. (Note that it is hard to randomly choose numbers in this fashion.)*

- *What is $P(X = .3)$*
- *If we assume all choices between 0 and 1 are equally likely, what is $P(X \geq .5)$.*
- *What is $P(X \leq .3$ or $X \geq .6)$?*

- It no longer makes sense to ask what $P(X = a)$ is, but rather what is $P(a \leq X \leq b)$.

Example 11. *Let X be a number randomly chosen between 0 and 1. (Note that it is hard to randomly choose numbers in this fashion.)*

- *What is $P(X = .3)$*
- *If we assume all choices between 0 and 1 are equally likely, what is $P(X \geq .5)$.*
- *What is $P(X \leq .3$ or $X \geq .6)$?*

The probabilities end up corresponding to (ratios of) lengths within these state spaces. Probabilities can also correspond to (ratios of) areas.

The probabilities end up corresponding to (ratios of) lengths within these state spaces. Probabilities can also correspond to (ratios of) areas.

Example 12. *If a dartboard is 8 inches in radius and the center circle is $\frac{1}{2}$ inch in radius, what is the probability that a dart thrown at random will hit the center square?*

The probabilities end up corresponding to (ratios of) lengths within these state spaces. Probabilities can also correspond to (ratios of) areas.

Example 12. *If a dartboard is 8 inches in radius and the center circle is* $\frac{1}{2}$ *inch in radius, what is the probability that a dart thrown at random will hit the center square? What is the probability that a dart thrown at random will hit in the 20-point wedge?*

The probabilities end up corresponding to (ratios of) lengths within these state spaces. Probabilities can also correspond to (ratios of) areas.

Example 12. *If a dartboard is 8 inches in radius and the center circle is* $\frac{1}{2}$ *inch in radius, what is the probability that a dart thrown at random will hit the center square? What is the probability that a dart thrown at random will hit in the 20-point wedge?*

Example 13. [Fun example] *You and your housemates have a two-bathroom house, but only one hot shower* *can be taken at a time. You have one housemate whose first class is at the same time as yours, and unfortunately you both like to start your day in the same half-hour window.*

can be taken at a time. You have one housemate whose first class is at the same time as yours, and unfortunately you both like to start your day in the same half-hour window. What are the chances that you get to have an uninterrupted hot shower?

can be taken at a time. You have one housemate whose first class is at the same time as yours, and unfortunately you both like to start your day in the same half-hour window. What are the chances that you get to have an uninterrupted hot shower?

The most common way in which areas represent probabilities is through areas under *density curves*. Here the random variable is listed on the x-axis, and $P(a < X < b)$ will end up being the area under the density curve between a and b.

can be taken at a time. You have one housemate whose first class is at the same time as yours, and unfortunately you both like to start your day in the same half-hour window. What are the chances that you get to have an uninterrupted hot shower?

The most common way in which areas represent probabilities is through areas under *density curves*. Here the random variable is listed on the x-axis, and $P(a < X < b)$ will end up being the area under the density curve between a and b. (For those who have had calculus, this should ring a bell (pardon the pun)).

can be taken at a time. You have one housemate whose first class is at the same time as yours, and unfortunately you both like to start your day in the same half-hour window. What are the chances that you get to have an uninterrupted hot shower?

The most common way in which areas represent probabilities is through areas under *density curves*. Here the random variable is listed on the x-axis, and $P(a < X < b)$ will end up being the area under the density curve between a and b. (For those who have had calculus, this should ring a bell (pardon the pun)).

- Example: uniform.

- Example: uniform.
- Example: sum of two numbers chosen randomly.

- Example: uniform.
- Example: sum of two numbers chosen randomly.
- Example: normal.

- Example: uniform.
- Example: sum of two numbers chosen randomly.
- Example: normal.

We will expand on and use this idea more down the road.