Week Four: Continuous Distributions

•••

CS 217

Discrete Distributions

- Last week we discussed discrete distributions, which shows the probability of outcomes with **finite values**.
- If I flip a coin, there are only two possible outcomes in the sample space.
  - *What are they?*

Discrete Distributions

- Last week we discussed **discrete distributions**, which shows the probability of outcomes with **finite values**.
- If I flip a coin, there are only two possible outcomes in the sample space.
  - Either heads or tails.
- For a discrete distribution, we obtain our results by **counting**.
- If I flip ten coins, I will be counting the number of times that heads comes up.
- If Aaron Judge gets up to bat ten times, I will be counting the number of hits he has.

Discrete Distributions

- What is the **probability mass function** for a discrete distribution?
- What is the **cumulative distribution function** for a discrete distribution?

Discrete Distributions

- The **probability mass function** gives us the probability that a discrete distribution is **equal** to a given outcome.
- The **cumulative distribution function** gives us the probability that a discrete distribution is **less than or equal** to a given outcome.

Continuous Distributions

- A **continuous distribution** is when our outcome is **continuous** with an infinite sample space rather than **discrete** with a finite sample space.
- Examples of continuous variables are distance, age, and temperature.
- We obtain our results by **measuring** rather than **counting**.
- Say I wanted to measure how long it took me to run a mile.
- What possible outcomes are in the sample space?

Continuous Distributions

- A **continuous distribution** is when our outcome is **continuous** with an infinite sample space rather than **discrete** with a finite sample space.
- Examples of continuous variables are distance, age, and temperature.
- We obtain our results by **measuring** rather than **counting**.
- Say I wanted to measure how long it took me to run a mile.
- What possible outcomes are in the sample space?
  - Technically, there are an *infinite* number of possibilities in the sample space.

Continuous Distributions

- Say it takes me 8 minutes and 14 seconds to run a mile. How do I record this?
- It could be:
  - 8 minutes 14 seconds
  - 8 minutes 14.2 seconds
  - 8 minutes 14.23 seconds
  - 8 minutes 14.229 seconds
  - 8 minutes 14.2294 seconds
  - 8 minutes 14.22936 seconds
  - Etc...
- The sample space is infinite because there is no limit to how precise our measurement can be.
- The measurement of a continuous variable is restricted by the methods used, or by the accuracy of the measuring instruments.

Continuous Distributions

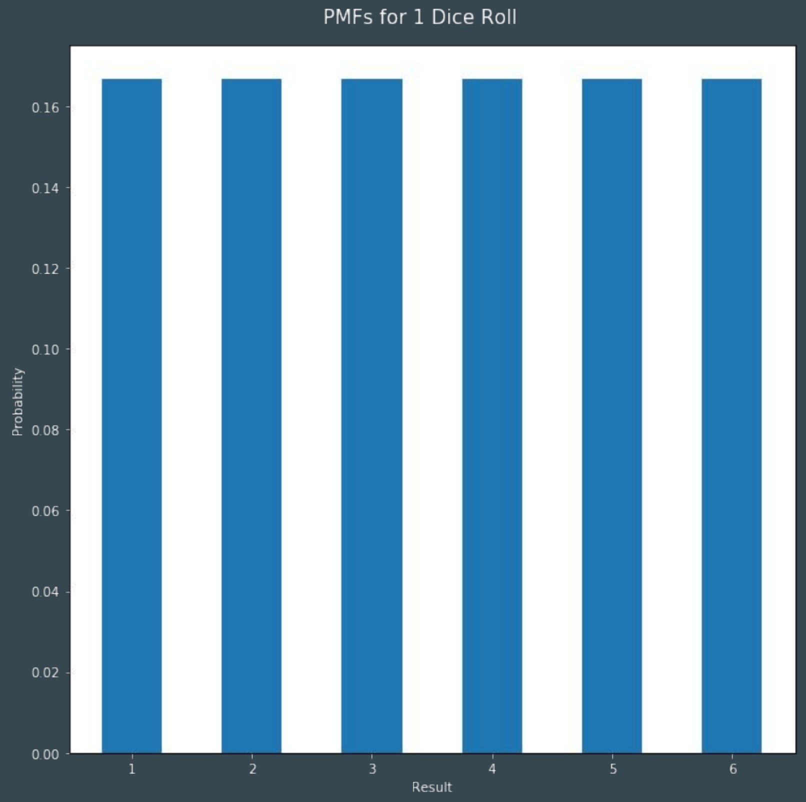
- The **probability density function** is the continuous analogue of the probability mass function
- It is technically the relative likelihood of a given value in a distribution.
- Because the number of possibilities are infinite, the possibility of getting any given value is zero.
- Thus we want to measure the probability of our outcome being *in between* two values.
- We can do this by taking the **integral** of the **probability density function**.

Continuous Distributions

- The **probability density function** is the continuous analogue of the probability mass function
- It is technically the relative likelihood of a given value in a distribution.
- Because the number of possibilities are infinite, the possibility of getting any given value is zero.
- Thus we want to measure the probability of our outcome being *in between* two values.
- We can do this by taking the **integral** of the **probability density function**.
- Continuous distributions, like discrete distributions, have **cumulative distribution functions** that represent the probability of something being **less than or equal to** a given value

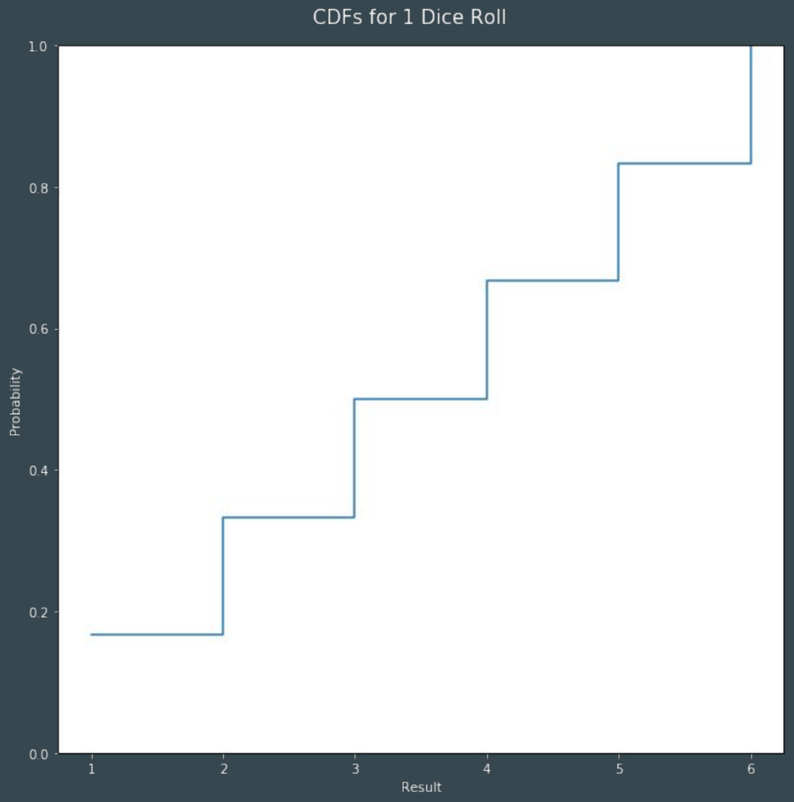
The Uniform Distribution

- Let's take a **uniform** distribution as an example of the difference between discrete and continuous distributions
- Let's first use the example of a single die as a **discrete uniform distribution** between 1 and 6, since those are the only possible outcomes
- The sample space is 1, 2, 3, 4, 5, and 6, all which have an equal probability of occuring.



The Uniform Distribution

- Let's take a **uniform** distribution as an example of the difference between discrete and continuous distributions
- Let's first use the example of a single die as a **discrete uniform distribution** between 1 and 6, since those are the only possible outcomes
- The sample space is 1, 2, 3, 4, 5, and 6, all which have an equal probability of occuring.

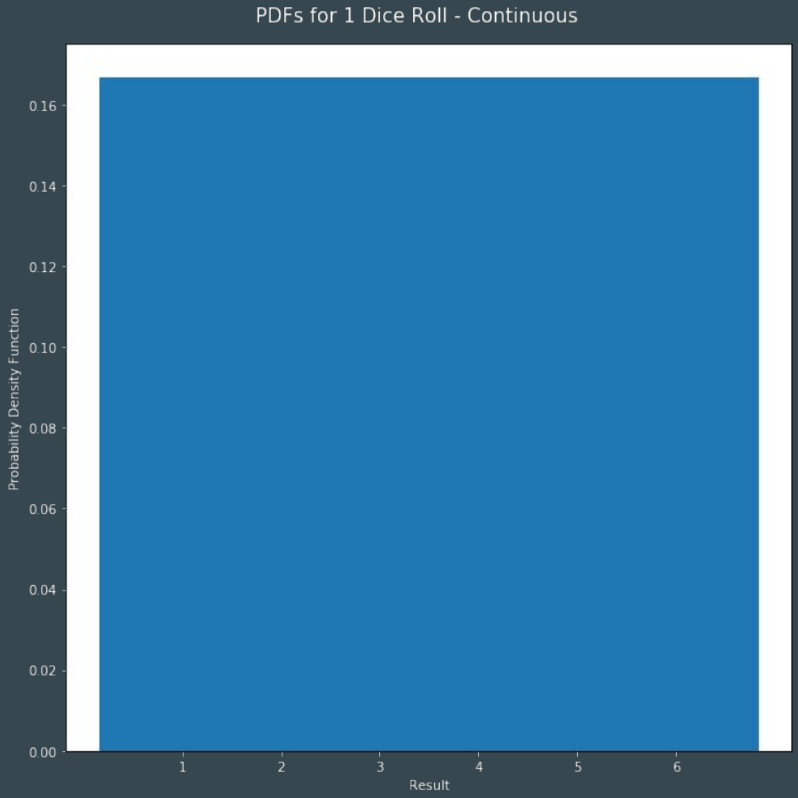


The Uniform Distribution

- Now let's look at a continuous uniform distribution.
- The continuous uniform distribution has **two inputs**: the minimum value and maximum value.
- A value has an equal probability of falling anywhere between these two values.

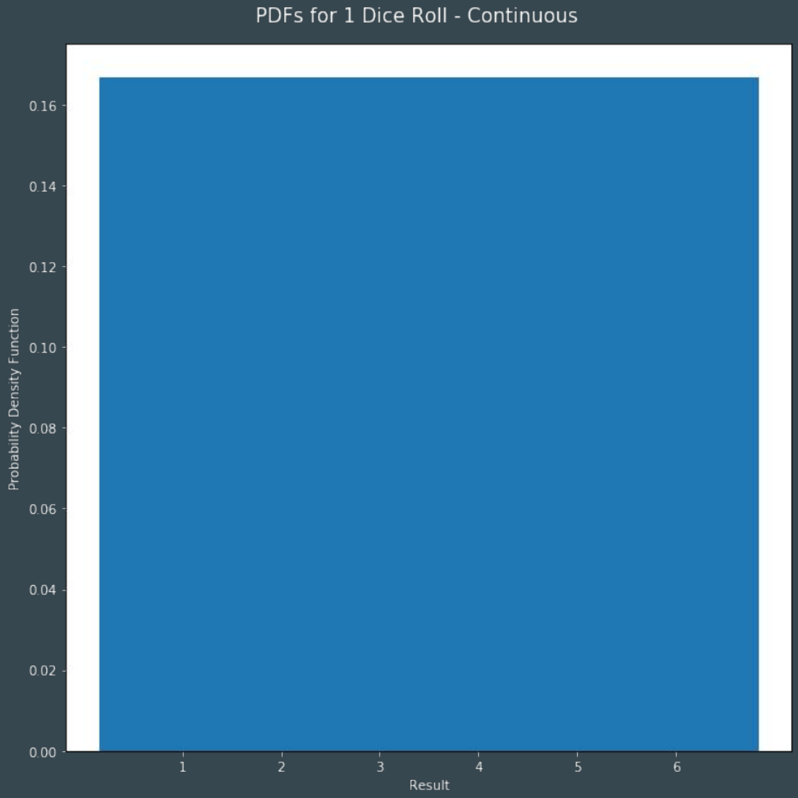
The Uniform Distribution

- Let's now look at a **continuous uniform distribution**, where a random variable has an equal chance of being any value from 0.5 to 6.5
- The sample space is the range from 0.5 to 6.5
- The **probability density function** for this distribution is constant at ⅙.



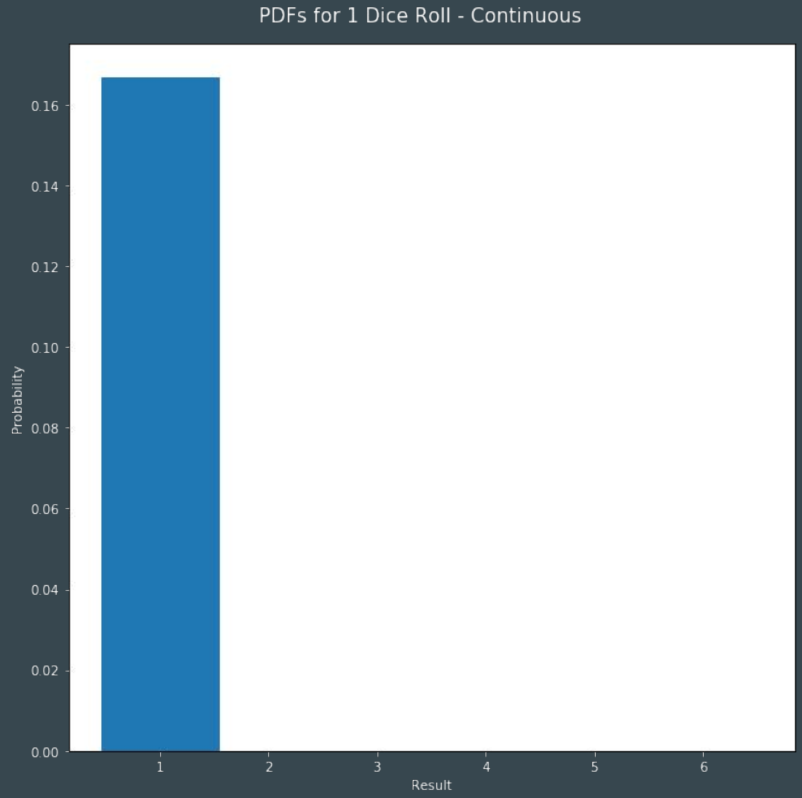
The Uniform Distribution

- Say that we have a computer pick a number between 0.5 and 6.5, and the integer that the result rounds to is the equivalent of landing on the rounded value.
- So if it picks somewhere between 0.5 and 1.5, it is the equivalent of getting a rolling a 1 with our die.



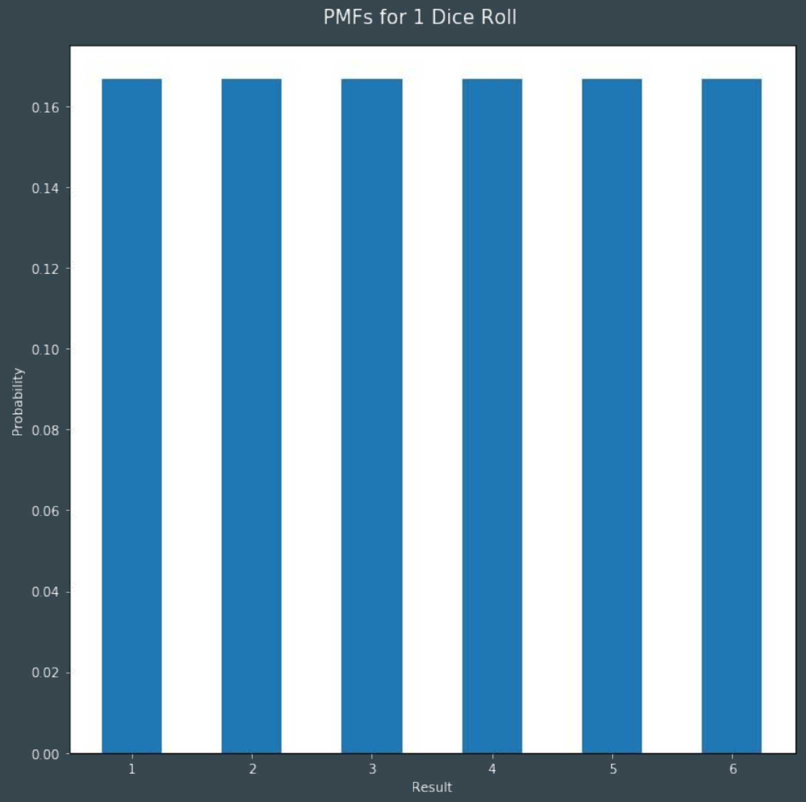
The Uniform Distribution

- Say that we have a computer pick a number between 0.5 and 6.5, and the integer that the result rounds to is the equivalent of landing on the rounded value.
- So if it picks somewhere between 0.5 and 1.5, it is the equivalent of getting a rolling a 1 with our die.
- The integral of the **probability density function** here is (1.5 - 0.5) * ⅙, so the **probability** of getting a 1 is 6.



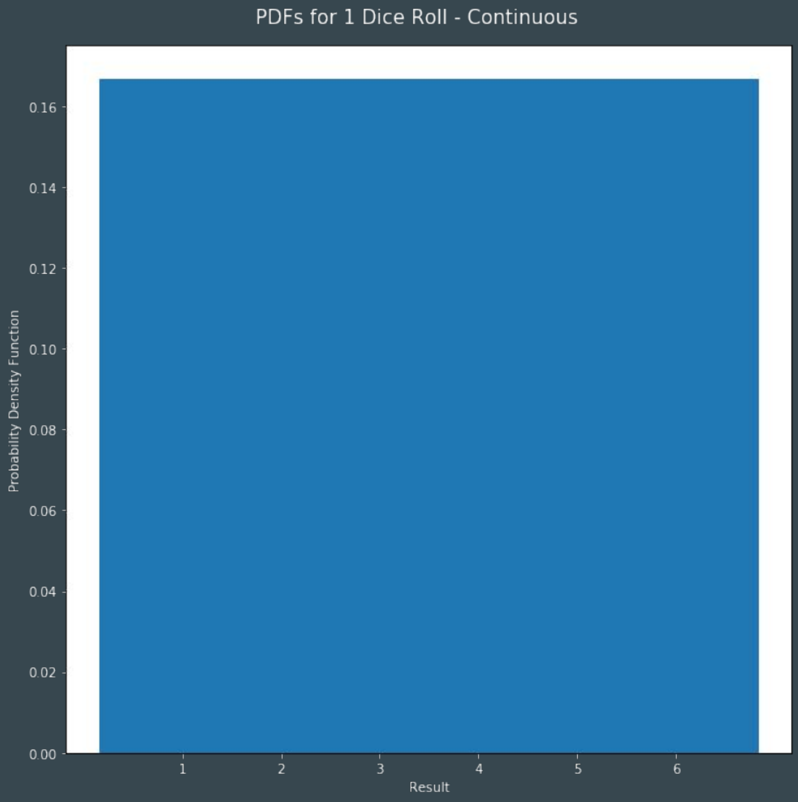
The Uniform Distribution

- In a **discrete distribution**, you can find the total probability by adding up all of the distinct PMFs
- Here, we have 6 outcomes in our sample space, each with a PMF of ⅙.
- These add up to a total probability of 1.



The Uniform Distribution

- In a **discrete distribution**, you can find the total probability by adding up all of the distinct PMFs
- Here, we have 6 outcomes in our sample space, each with a PMF of ⅙.
- These add up to a total probability of 1.
- Here, we find the integral value in this graph, so 6 * ⅙ , for a total probability of 1.

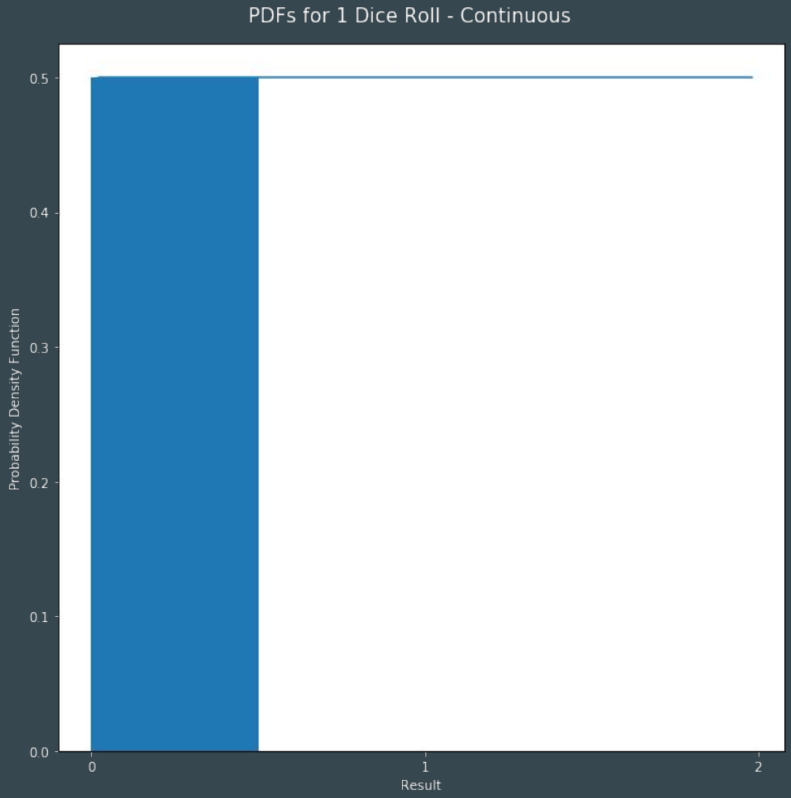


The Uniform Distribution

- Now say that we had a continuous uniform distribution between 0 and 2.
- The PDF of this distribution would be a constant of 0.5
- What is the probability of getting between 0 and 0.5?

The Uniform Distribution

- Now say that we had a continuous uniform distribution between 0 and 2.
- The PDF of this distribution would be a constant of 0.5
- What is the probability of getting between 0 and 0.5?
- The area to the right is 0.5 * (0.5 - 0), or 0.25.



Probability Density Function

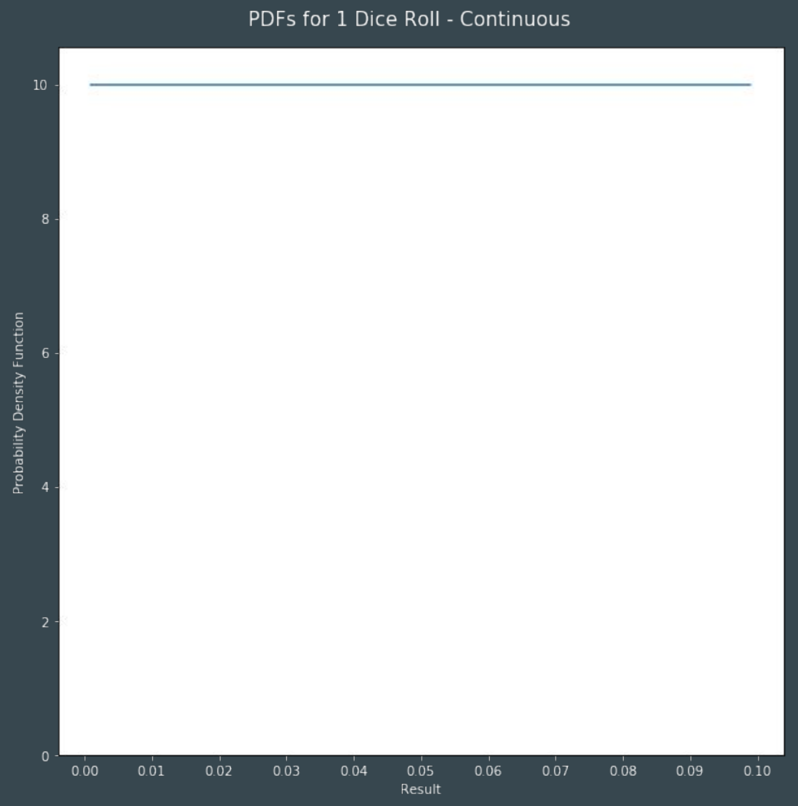
- The probability density function is a constant for the uniform distribution, but is much more complicated for other distributions.
- Mathematically, you have to take the integral of this function between two points in order to find the probability of a distribution being in between two values.
- We won't be doing integrals in this class, but conceptually it is essential to be aware that the probability density function is **not equal** to the probability for a given value.
- In fact, the probability density function can even be above 1 sometimes (whereas the total probability of a discrete or continuous distribution or a PMF of a continuous distribution will never be greater than 1).

Probability Density Function

- Visually, the probability density function is most useful for knowing the most common values, or modes, of a distribution.

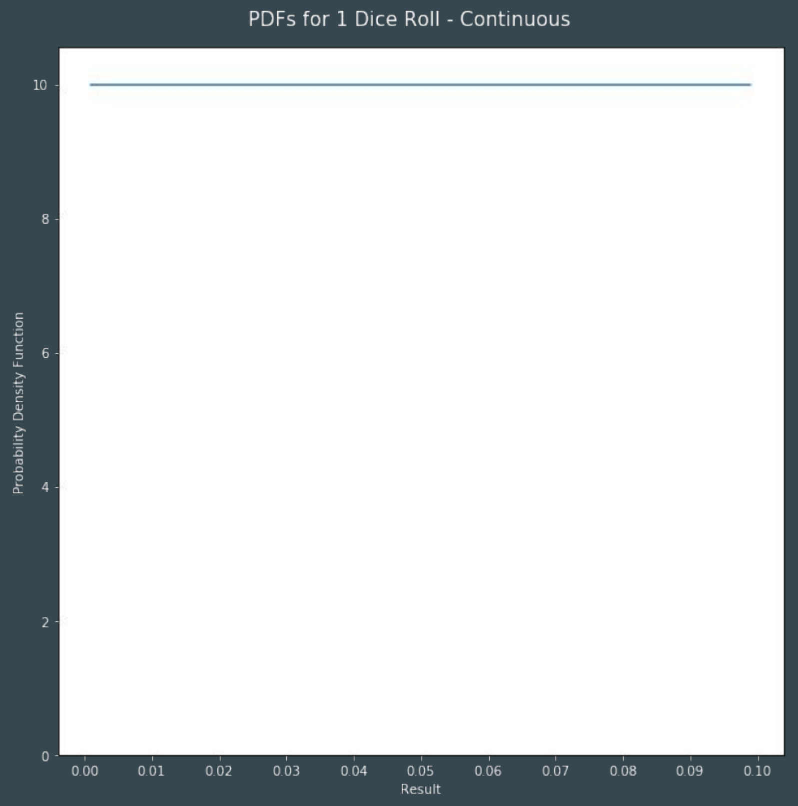
The Uniform Distribution

- Here is a uniform probability distribution between 0 and 0.1, which has a probability density function of 10.
- What is the probability of getting a value between 0 and 0.05 in this distribution?



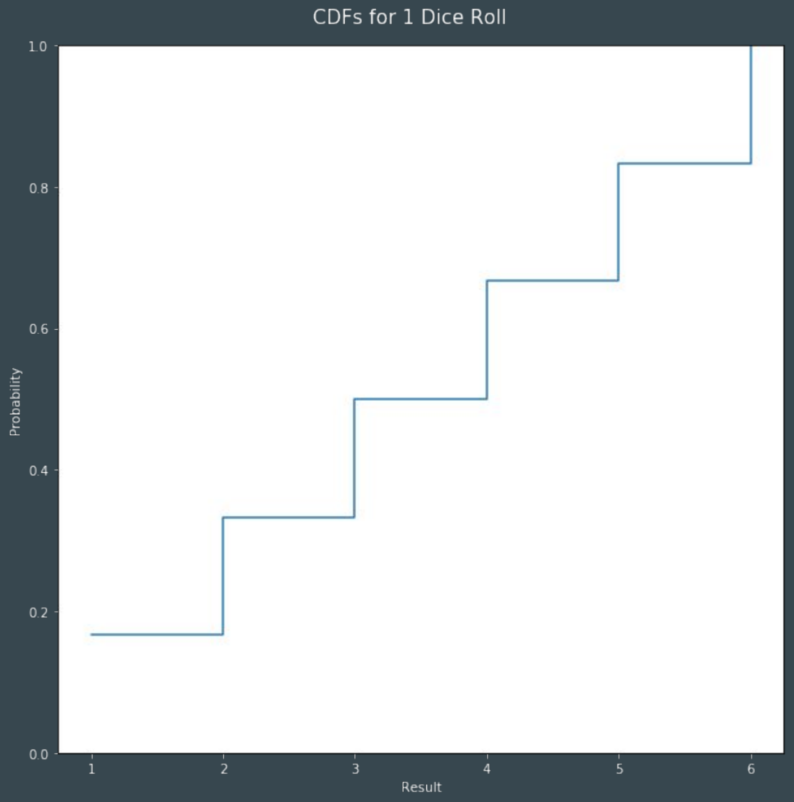
The Uniform Distribution

- Here is a uniform probability distribution between 0 and 0.1, which has a probability density function of 10.
- What is the probability of getting a value between 0 and 0.05 in this distribution?
- 10 * (0.05 - 0), or 0.5 (50%).



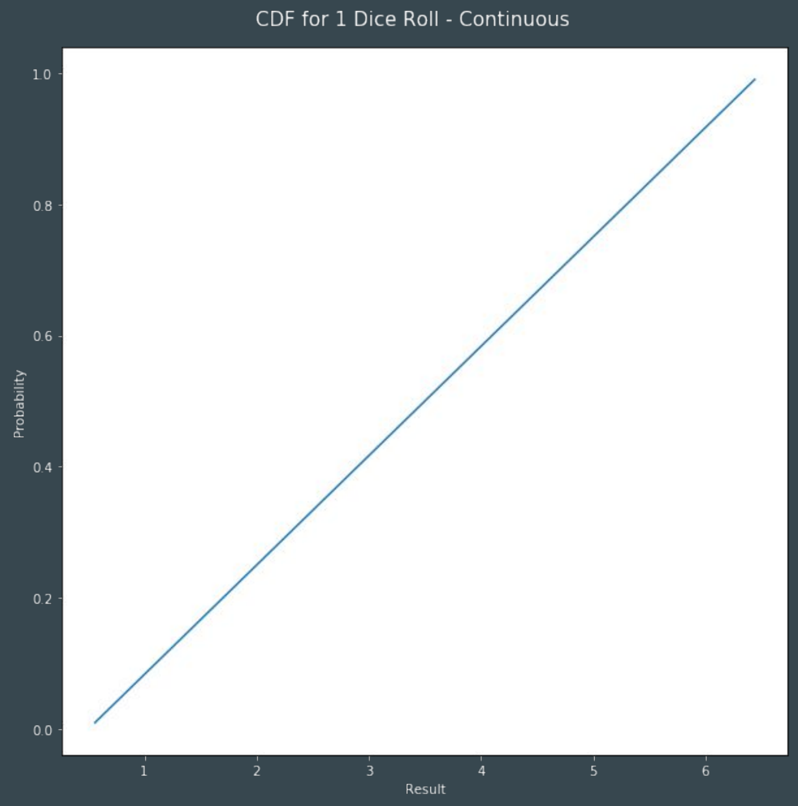
The Uniform Distribution

- In a **discrete uniform distribution**, the CDF looks like a step ladder, as the probability increases at every value in the sample space.



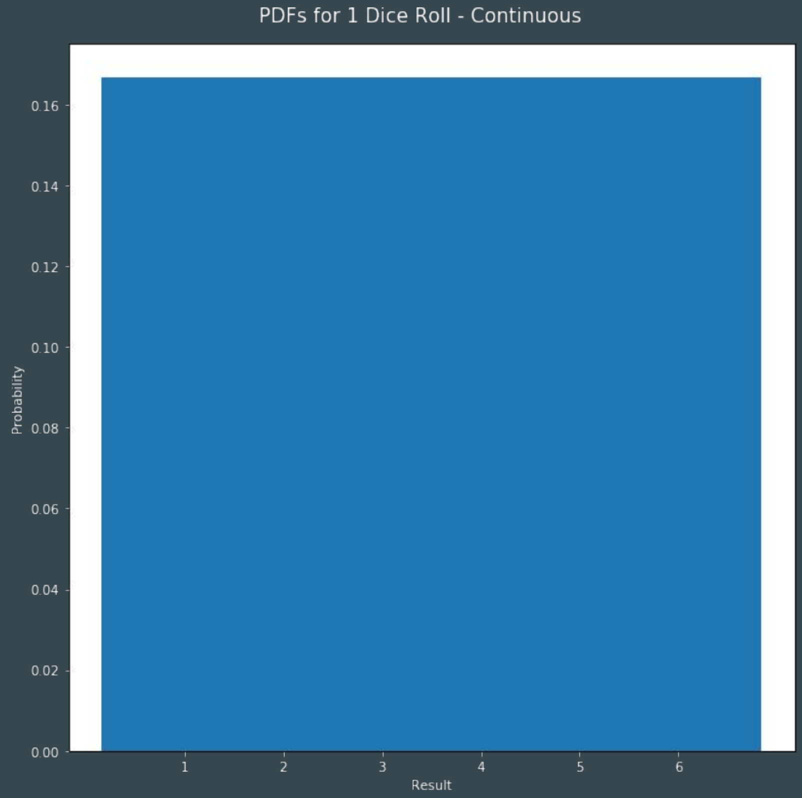
The Uniform Distribution

- In a **continuous uniform distribution**, the CDF is a straight line, as the probability increases linearly between 0.5 and 6.5.



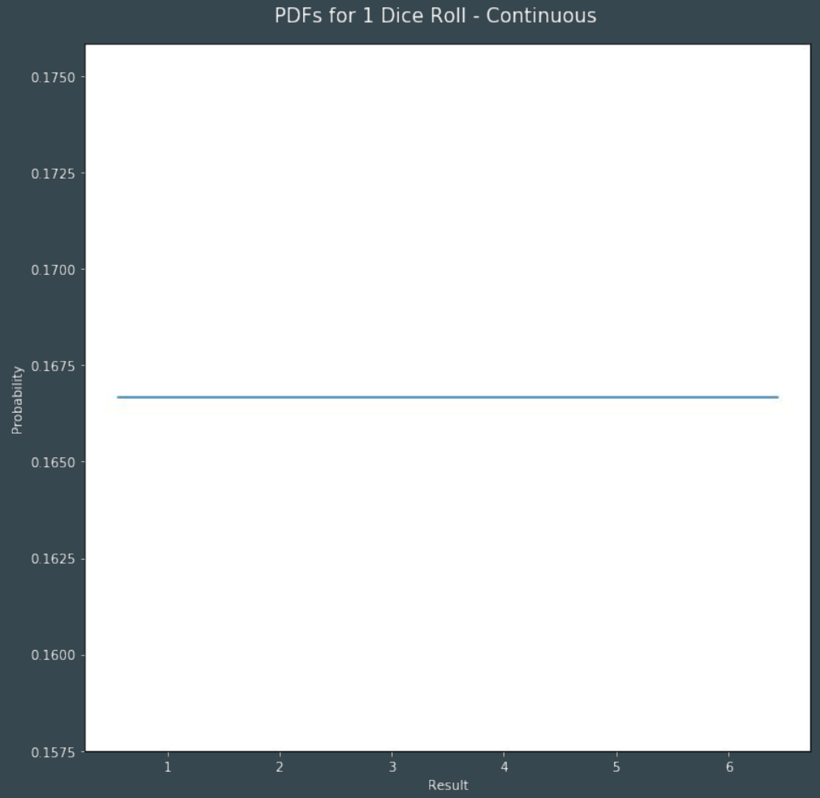
The Uniform Distribution

- I showed the PDF as a bar graph to compare it to the PMF graph, but it's typically viewed as a line graph.



The Uniform Distribution

- I showed the PDF as a bar graph to compare it to the PMF graph, but it’s typically viewed as a line graph.



Poisson Distribution

- A **poisson distribution** measures the probability of a given number of events happening in a fixed interval of time (as opposed to the **binomial distribution** which measures the probability of a given number of events happening in a fixed **number of trials**)
- With the poisson distribution, there is the assumption that the occurence of each event is independent from each other.
- This means that the occurrence of the next event happening is **not dependent** on the previous event occurring.
- This is called being **memoryless.**

Poisson Distribution

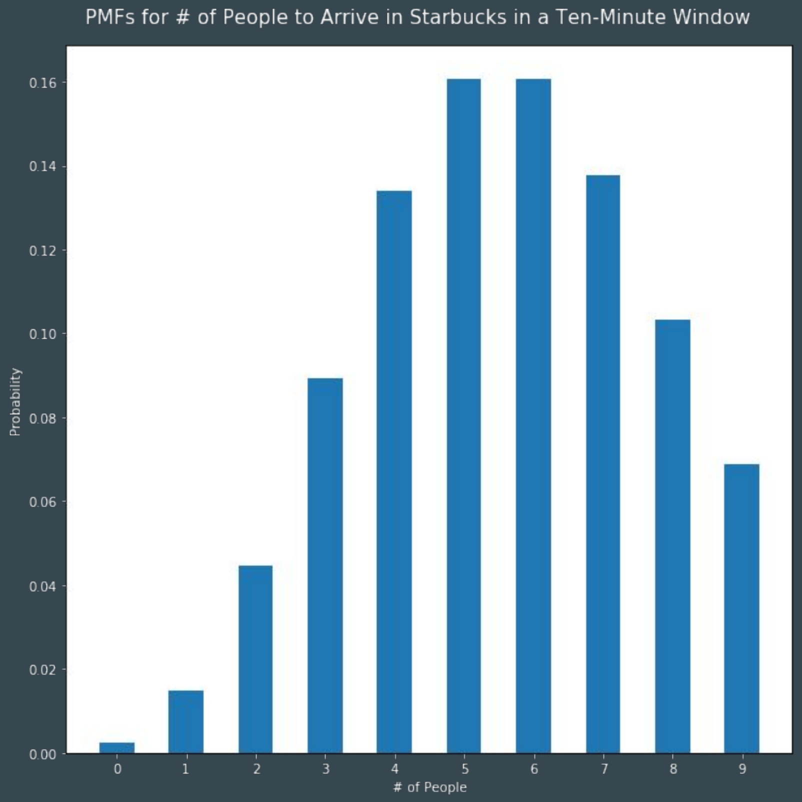
- An example of the Poisson distribution is the number of babies born in a hospital per hour, since the time one baby is born has nothing to do with when another baby is born
- Babies can be born at any point of the day, on any day of the week, so this rate will be constant 24/7.
- Another example is the number of customers who walk into a Starbucks every ten minutes. What assumptions would have to hold for this to be a **Poisson process**?

Poisson Distribution

- An example of the Poisson distribution is the number of babies born in a hospital per hour, since the time one baby is born has nothing to do with when another baby is born
- Babies can be born at any point of the day, on any day of the week, so this rate will be constant 24/7.
- Another example is the number of customers who walk into a Starbucks every ten minutes. What assumptions would have to hold for this to be a **Poisson process**?
    - The rate of customers entering the Starbucks is constant throughout all hours of the day.
    - One customer entering the store has no bearing on another customer entering the store.

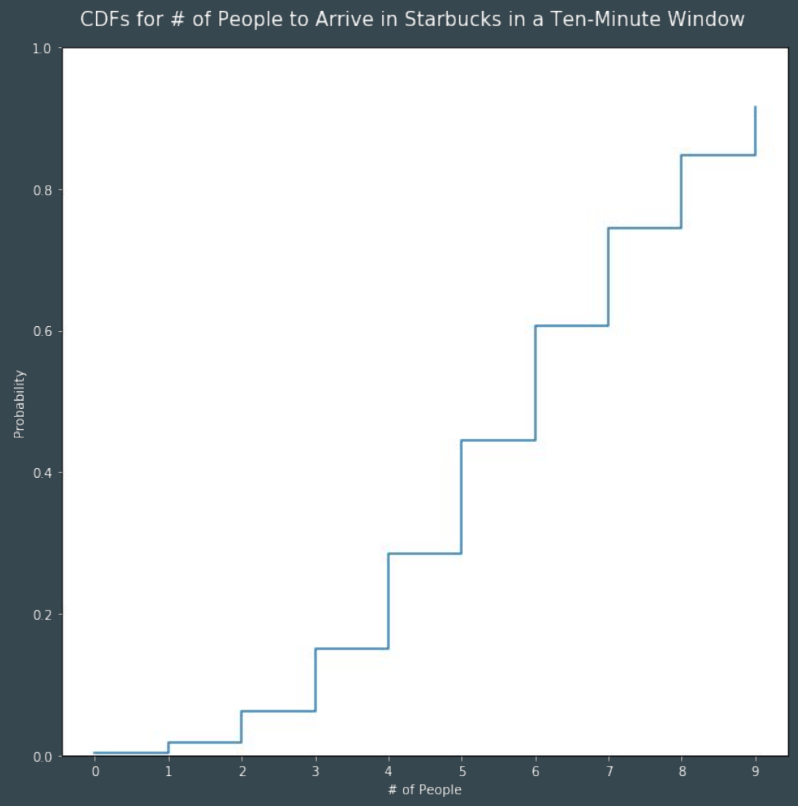
Poisson Distribution

- Let's say, on average, six people walk into Starbucks every ten minutes.
- To the right are the PMF values for the number of people who will walk into Starbucks in a given ten-minute window.
- The Poisson distribution technically has **one input**, lambda, which is also the *expected value* of the distribution.
- But the lambda value implies a time range, which here is ten minutes.
- Here the *lambda value* is 6, or six people in a ten-minute window.



Poisson Distribution

- Let’s say, on average, six people walk into Starbucks every ten minutes.
- To the right are the CDF values for the number of people who will walk into Starbucks in a given ten-minute window.
- Here the *lambda value* is 6, or six people in a ten-minute window
- Here is the corresponding CDF for the Poisson distribution.



Exponential Distribution

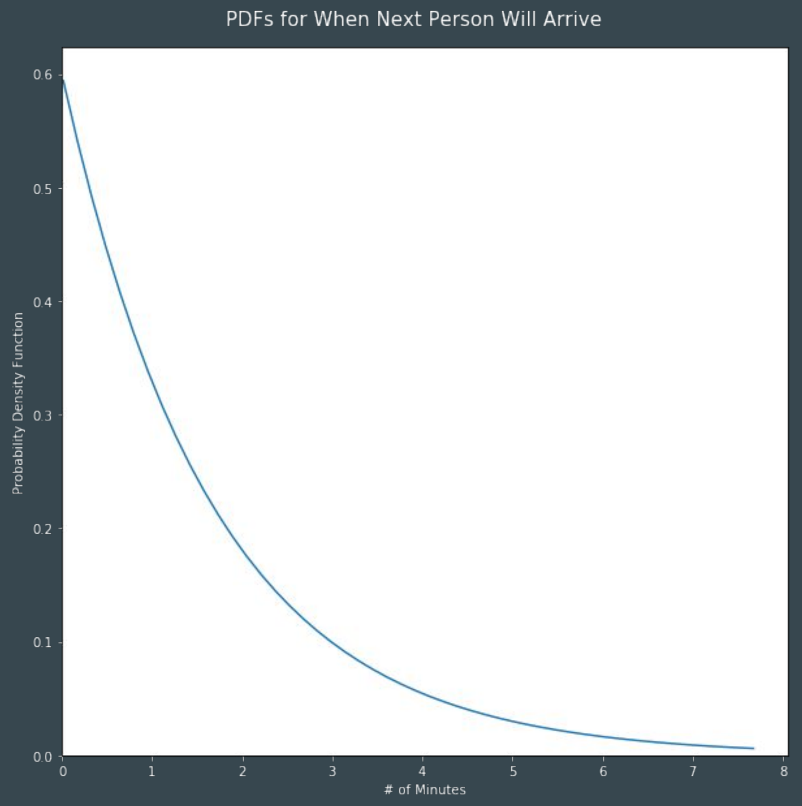
- The exponential distribution models the inverse of this. How long will it take for the next person to arrive?
- It uses **theta** as an input, which is the inverse of the lambda value from the exponential function. Here that value is ⅙, meaning that the expected value for how long it takes for the next person to come is ⅙ of a ten-minute window.
- It is a **continuous distribution**.

Exponential Distribution

- Both the Poisson and exponential distribution are **memoryless** in that what has happened in the past does not affect what happens in the future.
- If a person hasn't arrived in the past fifteen minutes, the probability that a person arrives in the next ten minutes will be unaffected.
- If ten people have arrived in the past fifteen minutes, the expected number of people that will arrive in the next ten minutes will remain the same.

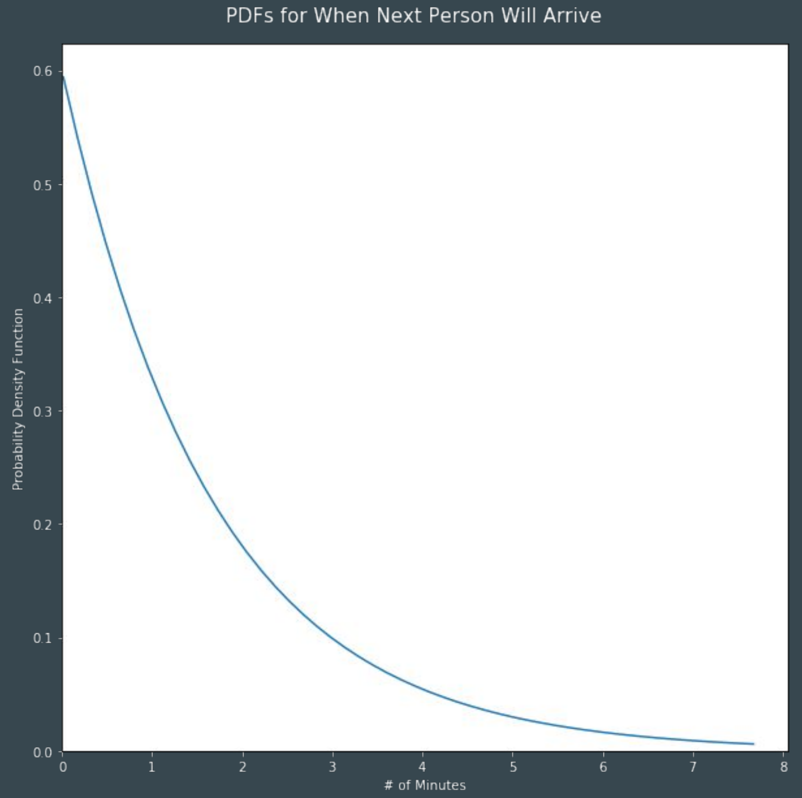
Exponential Distribution

- Let's say, on average, six people arrive in a Starbucks every ten minutes in a Poisson process.
- Here the theta value is 10/6, which is the inverse of 6 people every ten minutes.
- We expect the next person to come in 10/6 of a minute, or 1.667 minutes.



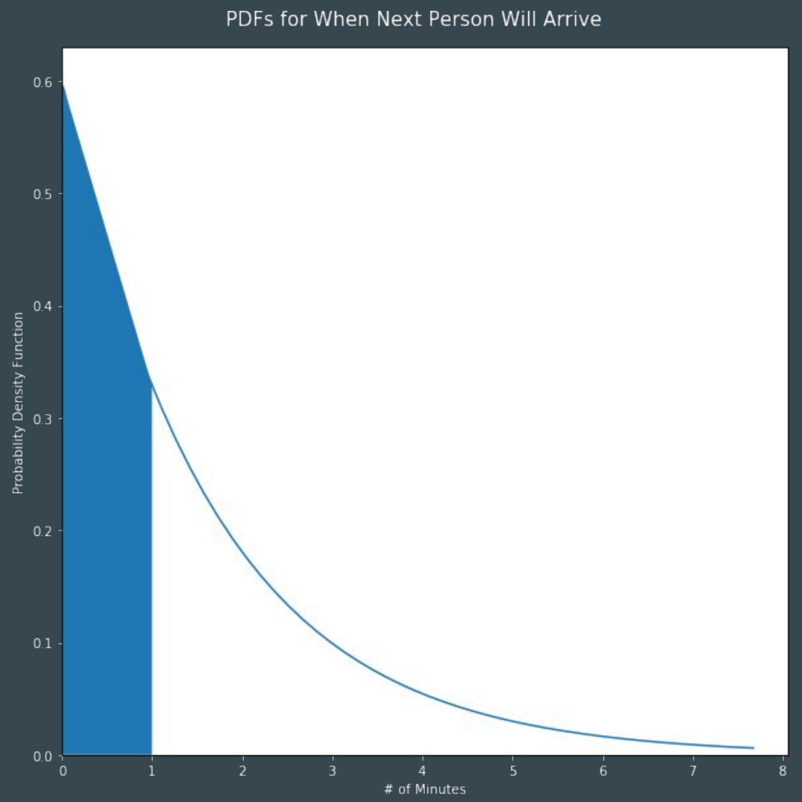
Exponential Distribution

- Note here that the PDF is no longer constant.
- The probability that the next person will arrive in one minute is the area under the curve between 0 and 1.



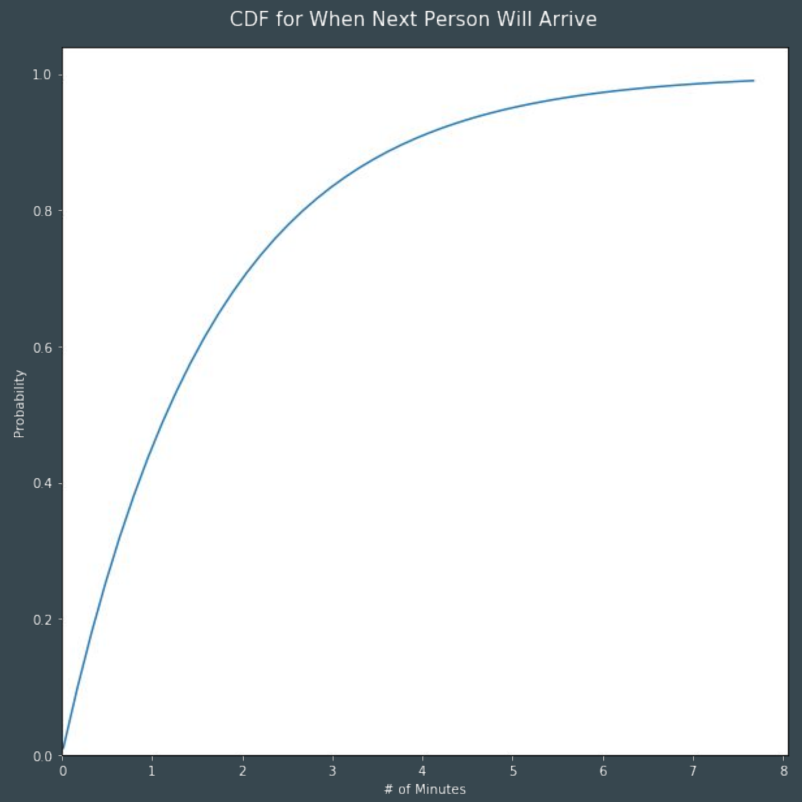
Exponential Distribution

- Note here that the PDF is no longer constant.
- The probability that the next person will arrive in one minute is the area under the curve between 0 and 1.
- This value is approximately **45%** (which we will confirm later).



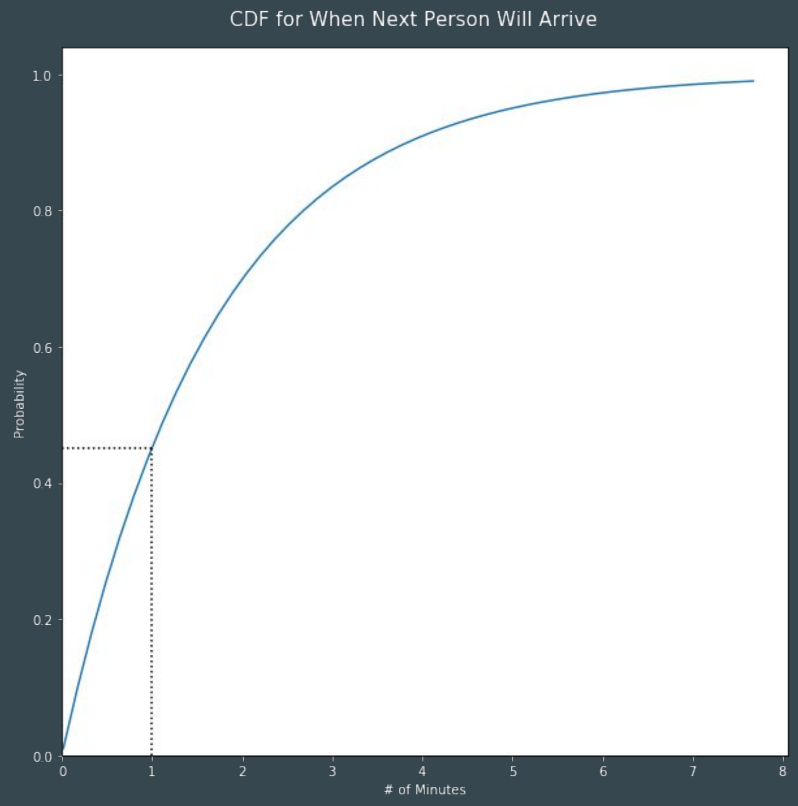
Exponential Distribution

- To the right is the CDF of the exponential distribution, which is essentially a flipped version of the PDF graph.
- Again like with a discrete distribution, we can use this to calculate the probability of someone arriving **at or before** a given time period.



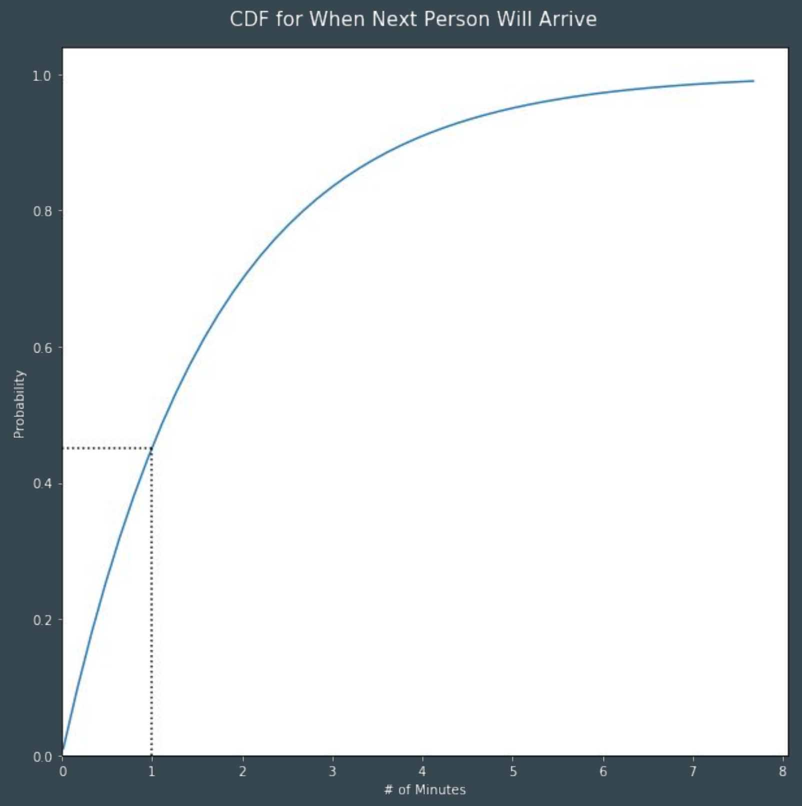
Exponential Distribution

- For example, we can visually confirm that there is around a 45% chance that someone will arrive in the next minute.



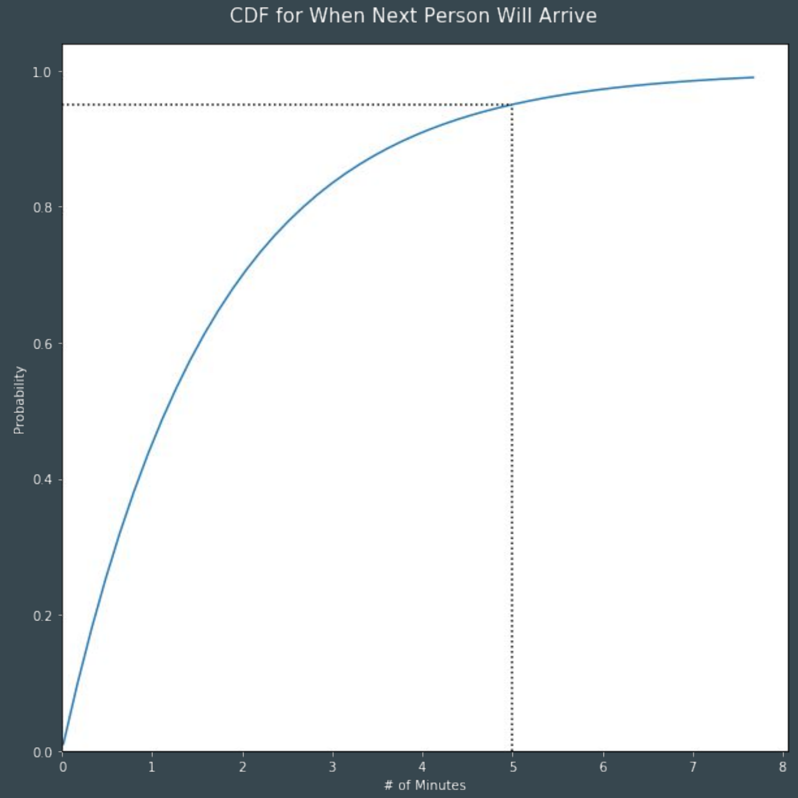
Exponential Distribution

- We can also use the CDF to calculate the **percentile rank** of a distribution.
- In this example, someone arriving in exactly a minute would be in the **45th percentile** of the distribution
- 45% of the time someone arrives sooner, and 55% of the time, someone arrives later.



Exponential Distribution

- Someone arriving in exactly five minutes would be in approximately the **95th percentile** of the distribution.
- 95% of the time somebody would arrive sooner, and 5% of the time somebody would arrive later.

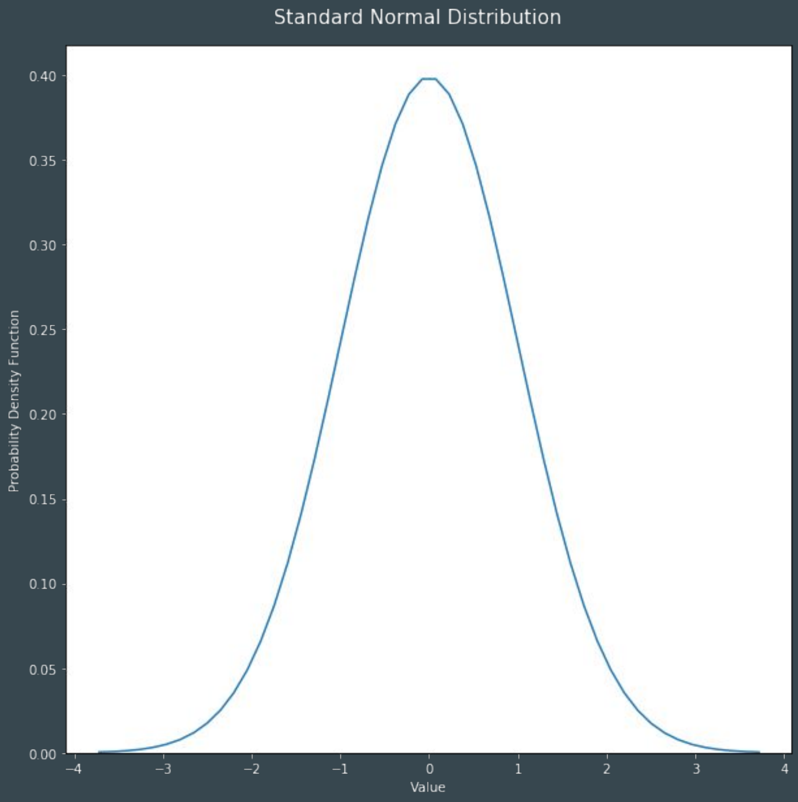


The Normal Distribution

- The **normal distribution** is by far the most important continuous distribution, for reasons we will explore throughout the course
- A normal distribution is a continuous distribution where the the data tends to cluster around a central value with no skew or bias
- Common examples of the **normal distribution** include height among males (or females), and SAT scores.
- It has two parameters, the mean and standard deviation.

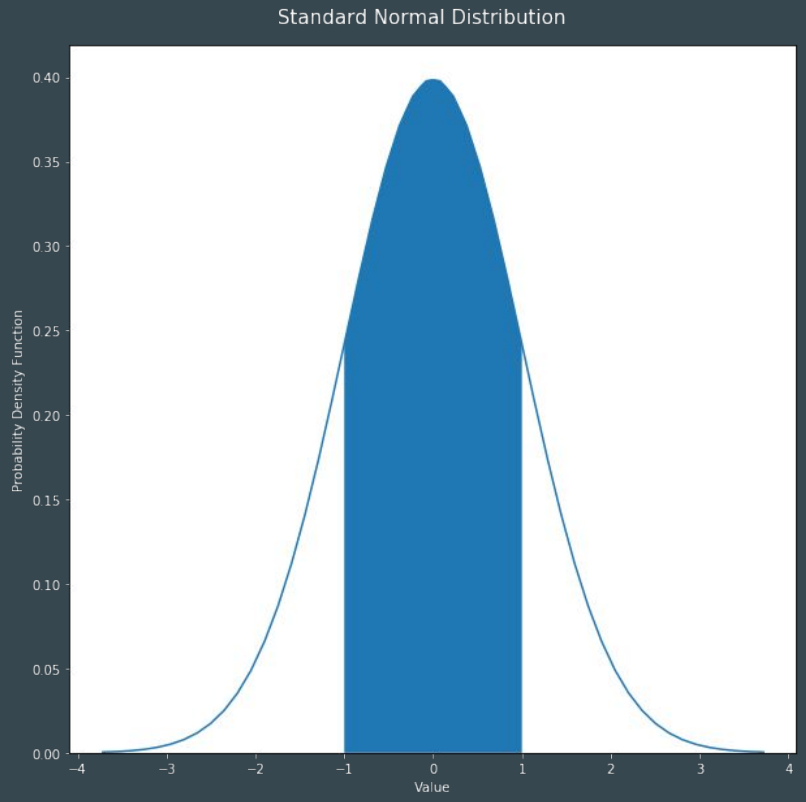
The Normal Distribution

- The **standard normal distribution** has a mean of 0 and a standard deviation of 1



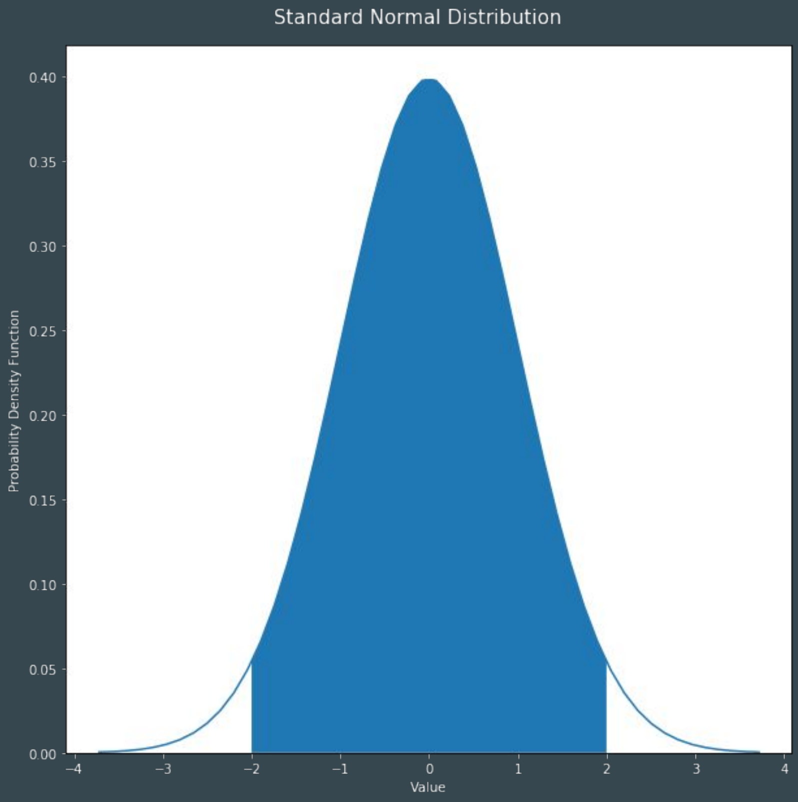
The Normal Distribution

- 68% of the data in a normal distribution will fall within **one standard deviation** of the mean
- With a mean of 0 and a standard deviation of 1, this means that the data has a 68% chance of falling between **-1** (0 -1) and **1** (0 + 1)
- Here this means that **68% of the area under the bell curve** is filled out.



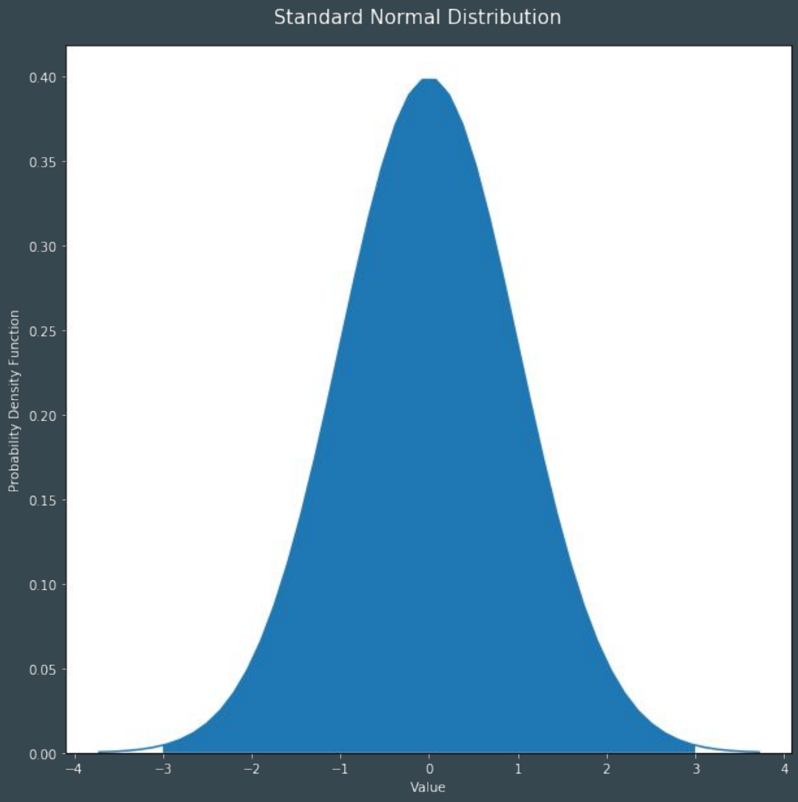
The Normal Distribution

- 95% of the data in a normal distribution will fall within **two standard deviations** of the mean
- With a mean of 0 and a standard deviation of 1, this means that the data has a 95% chance of falling between **-2** (0 -1) and **2** (0 + 1)
- Here this means that **95% of the area under the bell curve** is filled out.

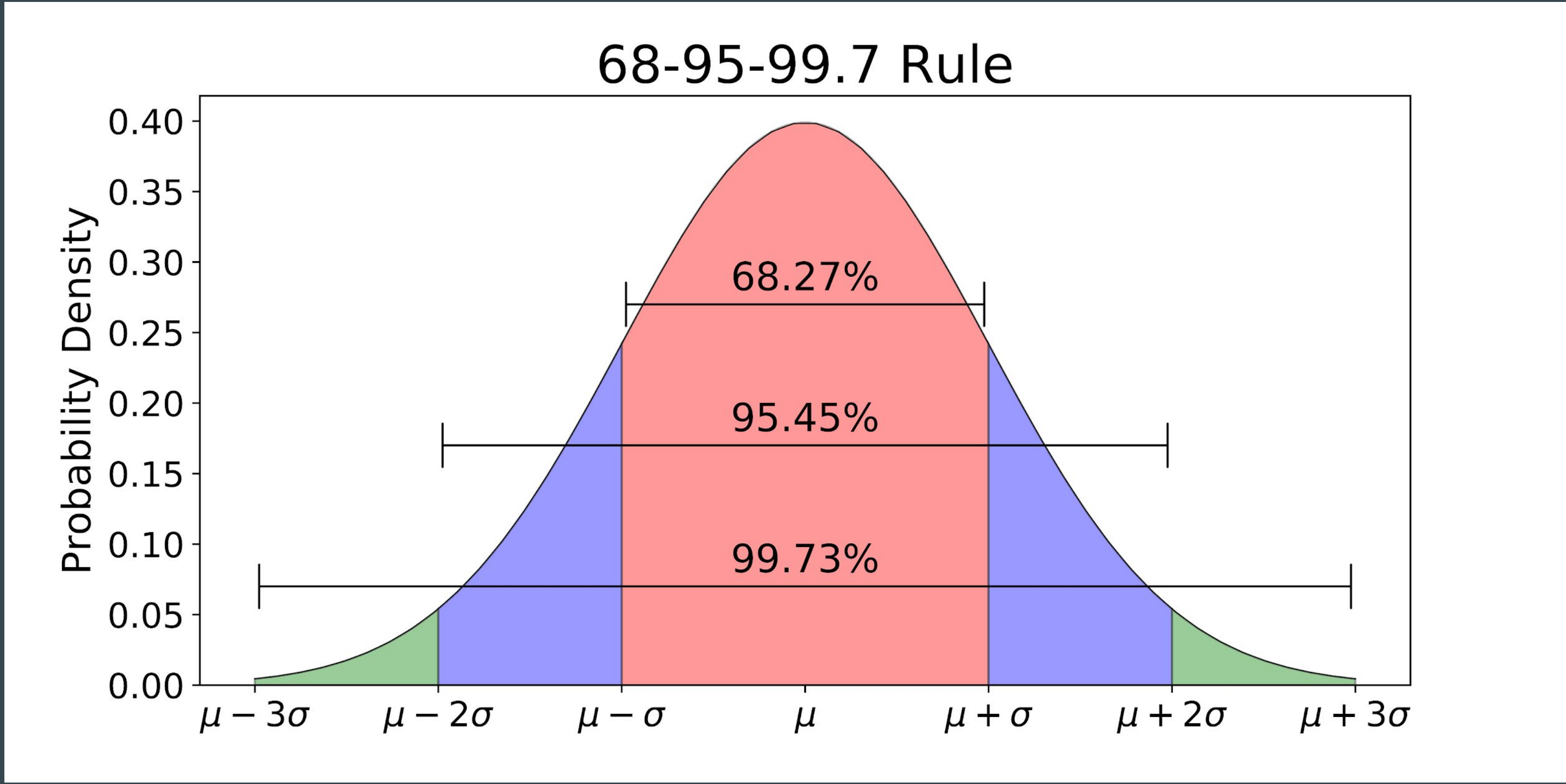


The Normal Distribution

- 99% of the data in a normal distribution will fall within **three standard deviations** of the mean
- With a mean of 0 and a standard deviation of 1, this means that 99% of the data will fall between ***-3*** (0 -3) and **3** (0 + 3)



Normal Distribution



Normal Distribution

- We can also confirm this by looking at each value's percentiles.
- What is the difference in the percentile value between -1 and 1?
- What is the difference in the percentile value between -2 and 2?
- What is the difference in the percentile value between -3 and 3?

Value	Percentile
-3	0.001
-2	0.022
-1	0.158
0	0.5
1	0.841
2	0.977
3	0.998

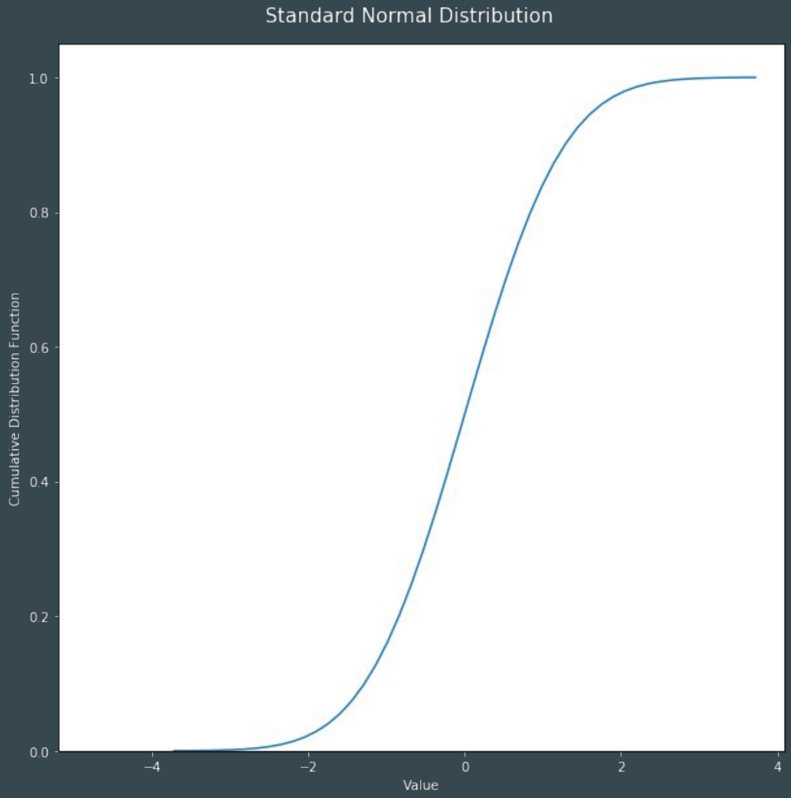
Normal Distribution

- We can also confirm this by looking at each value's percentiles.
- What is the difference in the percentile value between -1 and 1?
  - ~68.2%
- What is the difference in the percentile value between -2 and 2?
  - ~95.5%
- What is the difference in the percentile value between -3 and 3?
  - ~99.8%

Value	Percentile
-3	0.001
-2	0.022
-1	0.158
0	0.5
1	0.841
2	0.977
3	0.998

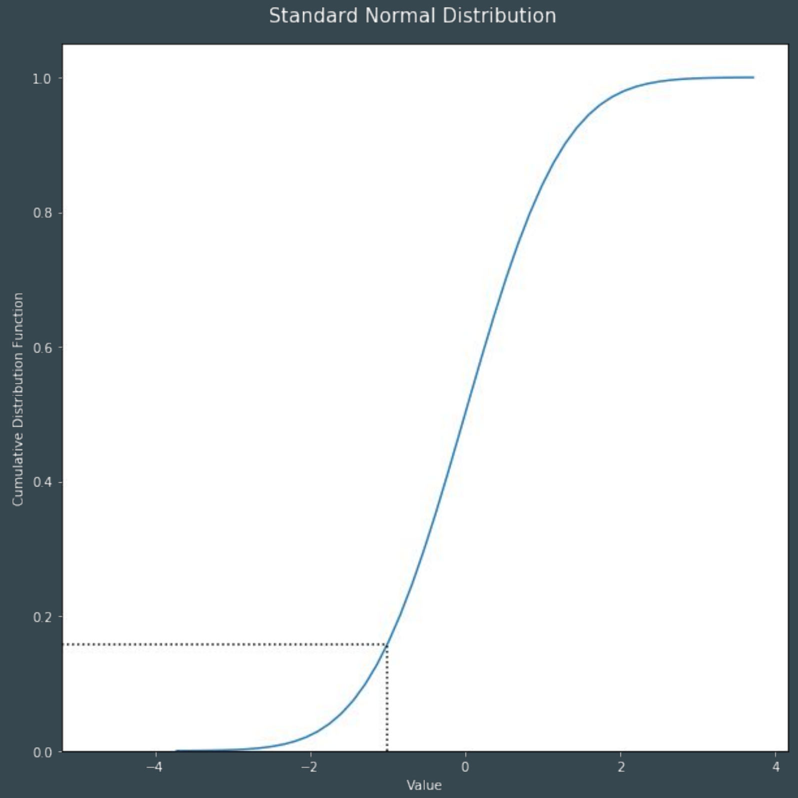
Normal Distribution

- We can also confirm these values via the **cumulative distribution function**.



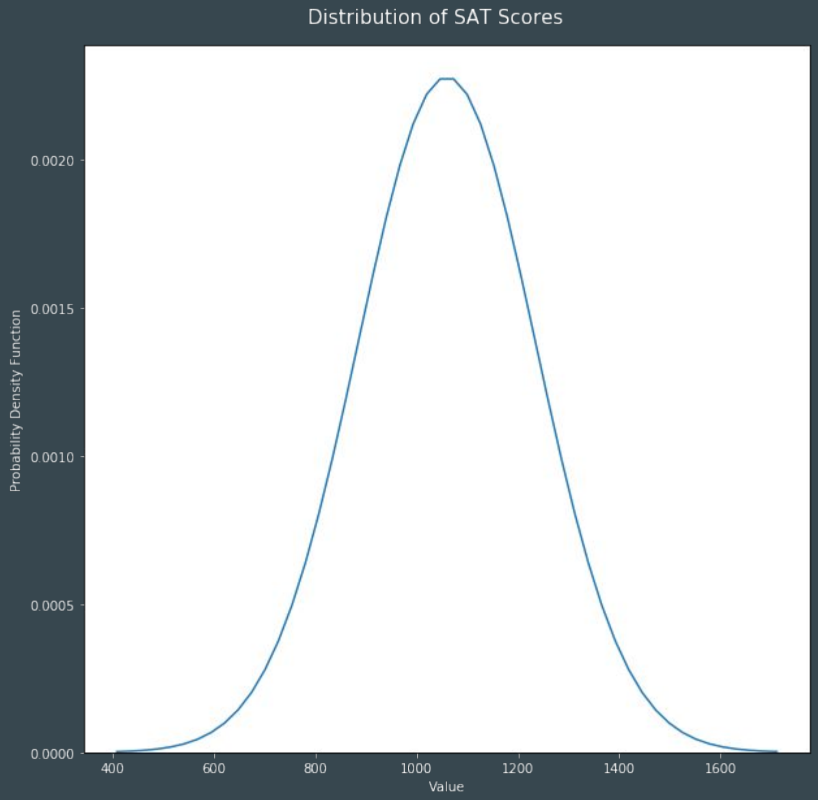
Normal Distribution

- We can also confirm these values via the **cumulative distribution function**.
- For example, -1 is in approximately the 15th percentile.



The Normal Distribution

- Other normal distributions behave the same way as the **standard normal distribution**, except we have different values for the mean and standard deviation.
- For example, we can say that SAT scores are normally distributed with a mean of 1060 and a a standard deviation of 175.



Normal Distribution

- The percentile functions stay the same, but instead of a raw number score, they are the **number of standard deviations away from the mean**.

Value	Percentile
-3	0.001
-2	0.022
-1	0.158
0	0.5
1	0.841
2	0.977
3	0.998

Normal Distribution

- The percentile functions stay the same, but instead of a raw number score, they are the **number of standard deviations away from the mean**.

Value	Score	Percentile
-3	535	0.001
-2	710	0.022
-1	885	0.158
0	1060	0.5
1	1235	0.841
2	1410	0.977
3	1585	0.998

The Z-Score

- Given these principles, we can find the CDF of any given point on a normal distribution if we have the **mean** and **standard deviation** of that normal distribution
- We can do this by seeing how many standard deviations away from the mean a given point is.
- This value is called the **Z-Score**.

$$Z = \frac{X - \mu}{\sigma}$$

The Z-Score

- A Z-score of less than 0 means that that value is less than the mean of the distribution, and a Z-score of more than 1 means that the value is more than the mean of the distribution
- The value of the Z-score is the **number of standard deviations away** a value is from the mean of a distribution.

$$Z = \frac{X - \mu}{\sigma}$$

The Z-Score

- For example, say that SAT scores are distributed with a mean of 1060 and a a standard deviation of 175
- If I get a 1120, what is my Z score?

$$Z = \frac{X - \mu}{\sigma}$$

The Z-Score

- For example, say that SAT scores are distributed with a mean of 1060 and a a standard deviation of 175
- If I get a 1120, what is my Z score?
- My Z-score is (1120-1060)/175, or about 0.34
- This means that I scored about 0.34 standard deviations above the mean score

The Z-Score

- For example, say that SAT scores are distributed with a mean of 1060 and a a standard deviation of 175
- If I get an 800, what is my Z score?

$$Z = \frac{X - \mu}{\sigma}$$

The Z-Score

- For example, say that SAT scores are distributed with a mean of 1060 and a a standard deviation of 175
- If I get an 800, my Z score is (800-1060)/195, or about -1.45
- This means that I scored about 1.45 standard deviations below the mean score.