Probability and Set Theory

Probability and Set Theory
Presented by David Baker
Dec 9th, 2019

// FLATIRON SCHOOL

Outline

What is Probability?

Numbers and Beliefs
Probability and Terms

Set Theory

Jupyter Notebook

Probability

Jupyter Notebook

Permutations and Combinations

Jupyter Notebook

Learning Goals

- ***Define sets, subsets, union, intersection as they are related to set theory***
- ***Explain how sample space and event space form the basis of the study of probability***
- ***Explain the three axioms of probability***
- ***Define the difference between permutations and combinations***
- ***Be able to give examples of independent and dependant events***

What are we even doing?!

One of the most important jobs of a data scientist (statistician) is to attempt to quantify uncertainty.

In order to give data scientists a common language to talk about such a task we employ mathematical language to formalize our assumptions and ensure we know what each other is talking about when trying to solve problems concerning quantifying uncertainty.

What does it mean to say that there is a 45% chance of it raining later today?

- What does the word "chance" mean in this context?
- Where does the number 45 come from here? What does it reflect?
- What is the 45% a reflection of? The day? Feb 3rds? The weather? Our experience of the weather?
- How does this assumption about there being a 45% chance of rain "today" compare and contrast to when we say that a six sided die has a ~50% chance of landing on an even number?
- How do discussions about a week of rain compare to a few rolls of a die?
- Take 10 minutes to discuss these questions in a group of four!

What does it mean to say that there is a 45% chance of it raining later today?

- Frequentism
- Bayes
- Likelihood
- Discrete vs Continuous
- Independent vs Dependant Events

Set Theory

Set Theory

- Branch of math dealing with collections of objects

- Define a set as S
- If an element x belongs to S we write $x \in S$.
- If not… $x \notin S$

Subsets

- Set *T* is a subset of *S* if every element of set *T* is also in *S*.
    - $T \subset S$

- S = { Python, R, SQL, Julia, C++ }
- T = {Python, SQL}

$T \subset S$

Set Operations

Union → Elements that are in EITHER S, T, or BOTH

Intersection → ALL of S that is in T

Set Operations

Dave = { R, Python, SQL, bash, LaTeX }

Wachira = { Python, C++, SQL, noSQL}

UNION ??? INTERSECTION? Jot down your answer...

Union → Elements that are in EITHER S, T, or BOTH

Intersection → ALL of S that is in T

Set Operations

Dave = { R, Python, SQL, bash, LaTeX }

Wachira = { Python, C++, SQL, noSQL}

UNION → {R, Python, SQL, bash, LaTeX, C++, noSQL}

Union → Elements that are in EITHER S, T, or BOTH

Intersection → ALL of S that is in T

Set Operations

Dave = { R, Python, SQL, bash, LaTeX }

Wachira = { Python, C++, SQL, noSQL}

Intersection → {Python, SQL}

Union → Elements that are in EITHER S, T, or BOTH

Intersection → ALL of S that is in T

Probability

Sample Space: All possible outcomes of an event*!

S = < 1, 2, 3, 4, 5, 6 > = Ω = Omega

*** event is any declarative (true or false) statement**

Probability

Sample Space: All possible outcomes of an event*!

S = < 1, 2, 3, 4, 5, 6 > = Ω = Omega

Event Space: a SUBSET of sample space

Event of Even Number? Rolling one? Rolling one or six?

* event is any declarative (true or false) statement

Probability

Sample Space: All possible outcomes of an event*!

S = < 1, 2, 3, 4, 5, 6 > = Ω = Omega

Event Space: a SUBSET of sample space

E = < 2, 4, 6> = Event Space of Even Number

* event is any declarative (true or false) statement

Axioms of Probability

1. Probability must always be positive
2. All probabilities sum to 1
3. Independent events in sample space are additive

1. Probability must always be positive

2. All probabilities sum to 1

3. Independent events in sample space are additive

Take five minutes and discuss how these axioms relate to our dice example

Conditional Probability

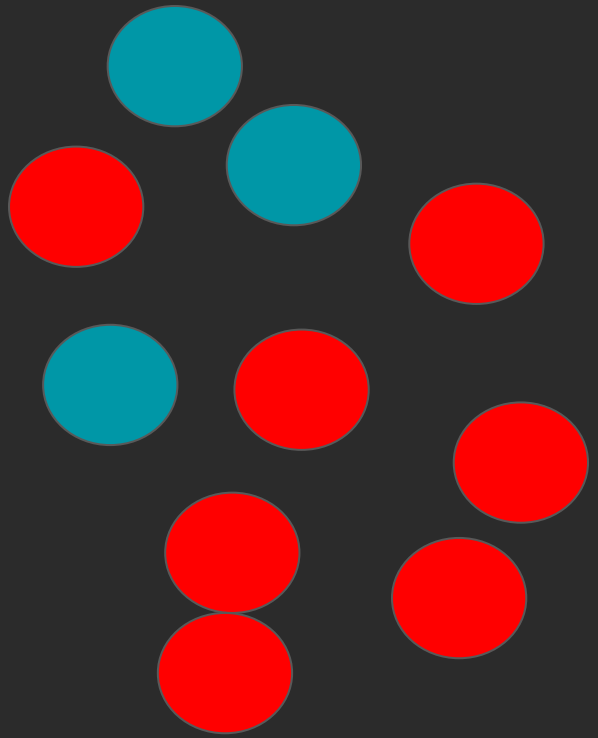
Rolling a die is an INDEPENDENT event

→ One roll of a die does not affect the next roll

Events are DEPENDENT if the outcome of event A has an effect on B.

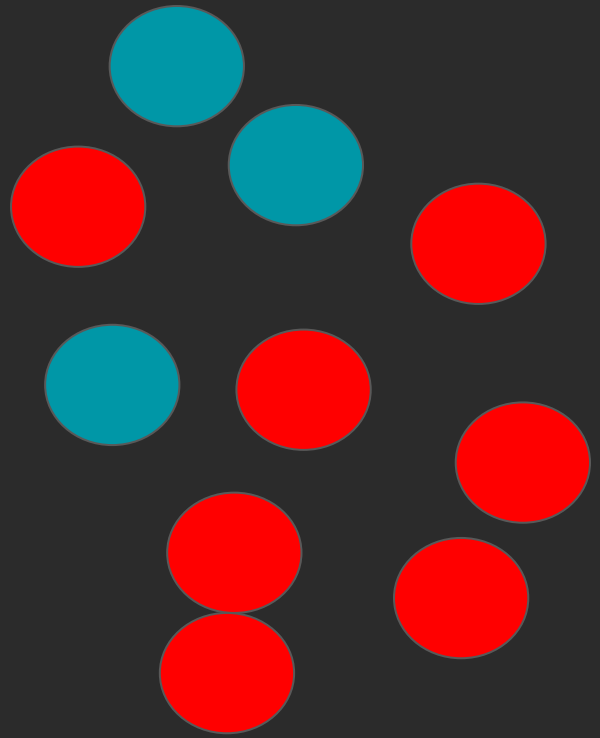
Balls in an Urn

- Ten balls in an urn/jar
- What is probability of pulling a red ball?
- What is probability of teal (green/blue) ball?
- What is probability of pulling a red, keeping it out, then getting a teal (green/blue) ball?



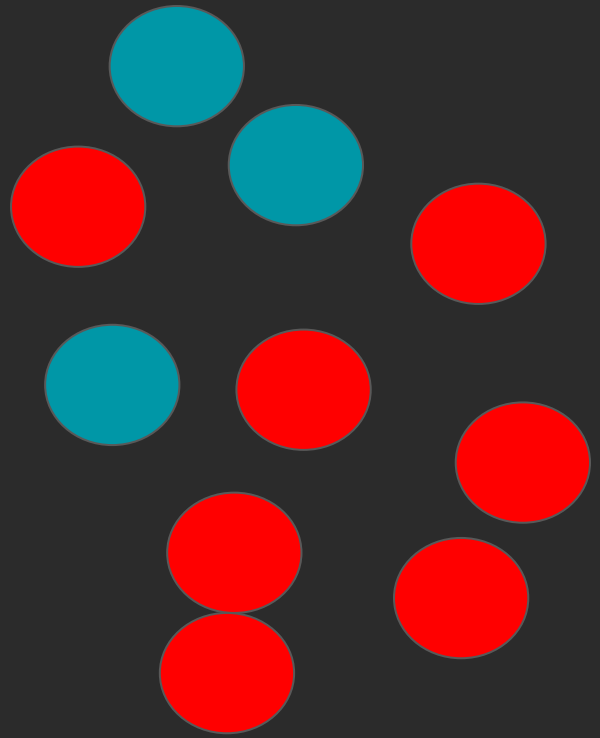
Balls in an Urn

- Ten balls in an urn/jar
- What is probability of pulling a red ball?
- What is probability of teal (green/blue) ball?
- What is probability of pulling a red, keeping it out, then getting a teal (green/blue) ball?



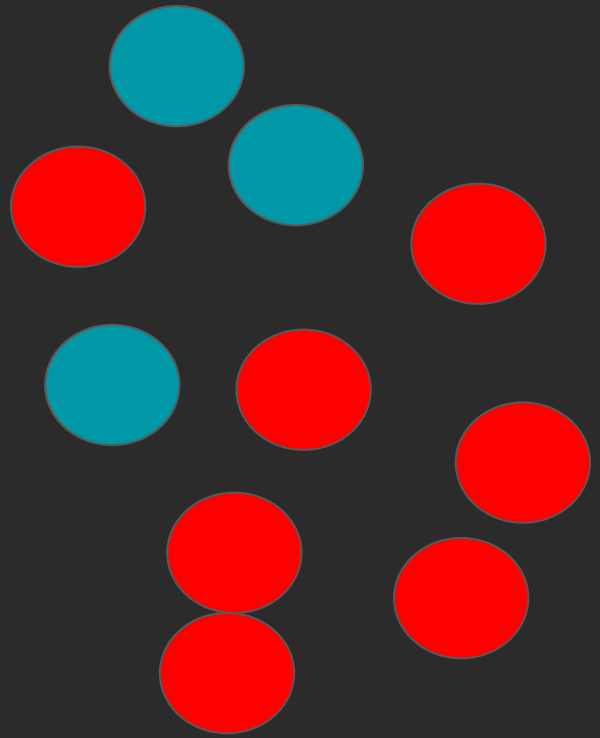
Balls in an Urn

- Ten balls in an urn/jar
- What is probability of pulling a red ball?
- What is probability of teal (green/blue) ball?
- What is probability of pulling a red, keeping it out, then getting a teal (green/blue) ball?

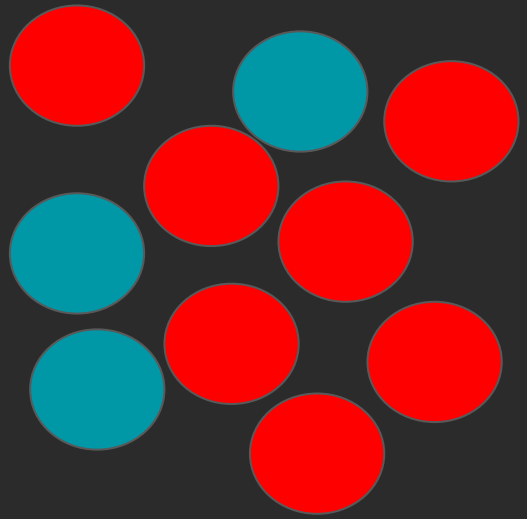


Balls in an Urn

- Ten balls in an urn/jar
- What is probability of pulling a red ball?
- What is probability of teal (green/blue) ball?
- What is probability of pulling a red, keeping it out, then getting a teal (green/blue) ball?



What is probability of pulling a red, keeping it out, then getting a teal (green/blue) ball?



7 / 10

3 / 9

These events are DEPENDENT since first affects second

Combinations and Permutations

When looking at full SET need to know if the order matters or not.

Permutation

→ Order Matters

Combination

→ Order DOES NOT matter

Combinations and Permutations Exercises

See Notebook!