$P(\textbf{PAL})$ Workshop

05/12/2022 17:00 Future Technologies Lab

Contents

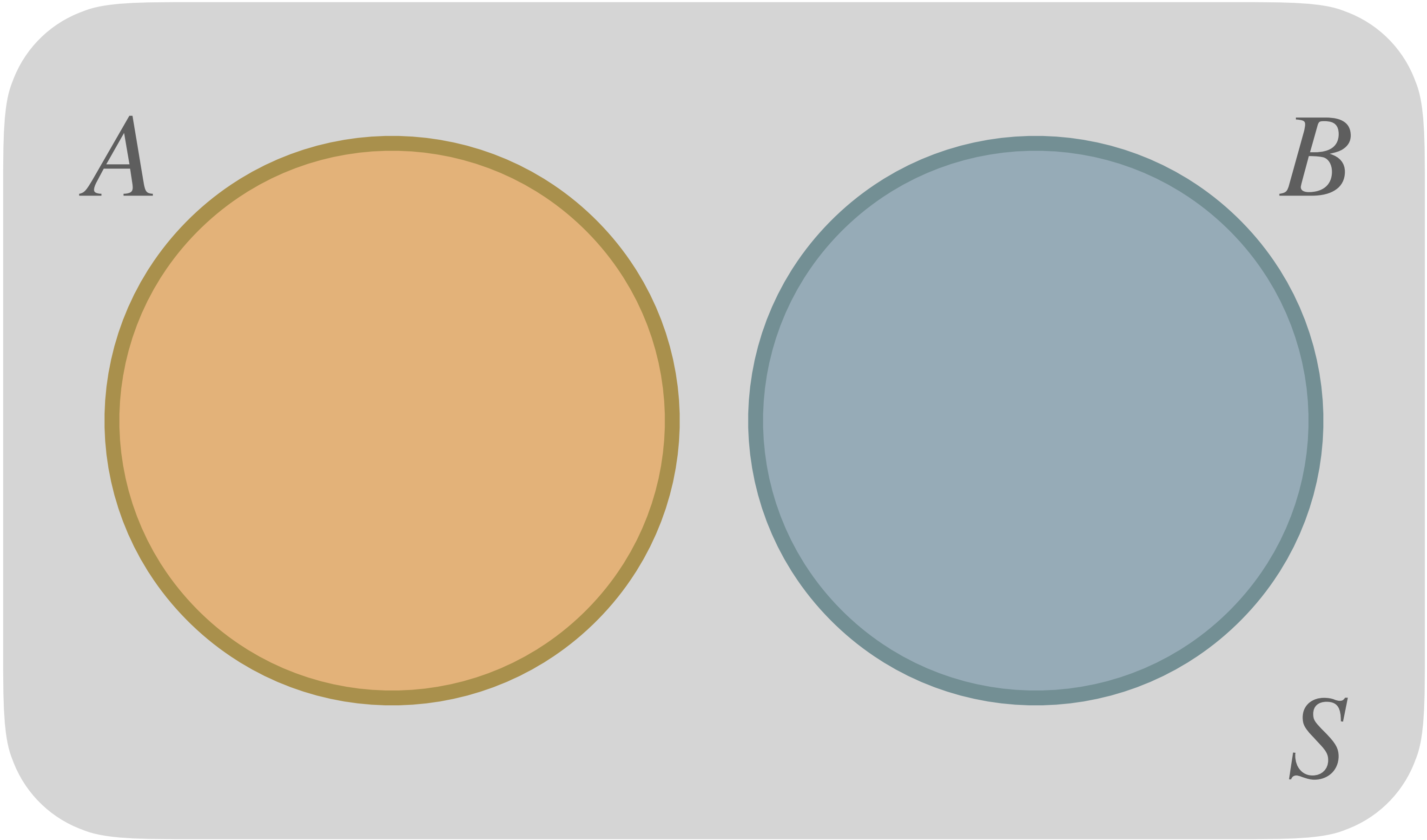
- Laws of probabilities
- Random Variables
- Discrete and Continuous spaces
- The Law of Total Probability
- Measures of Central Tendency and Spread
- Distributions

What is Probability?

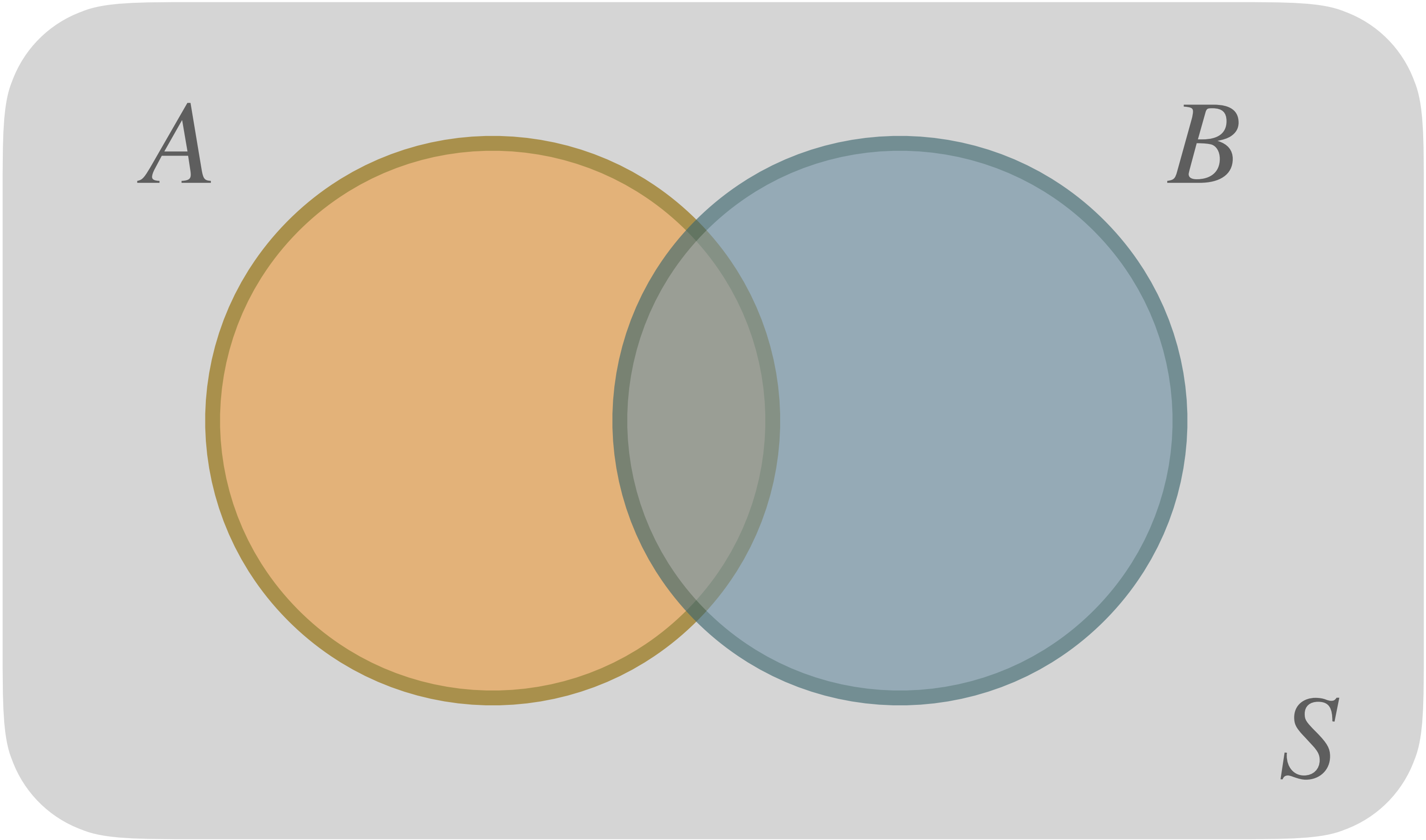
- Concerns the study on uncertainty (loosely speaking).
- For certain types of events, we cannot predict the outcome with certainty in advance, *e.g. tossing a coin or tossing a die.*
- However we know the set of all possible outcomes for these events.
- We would like to use probability to measure the chance of something occurring in an experiment.

Probabilities as Sets

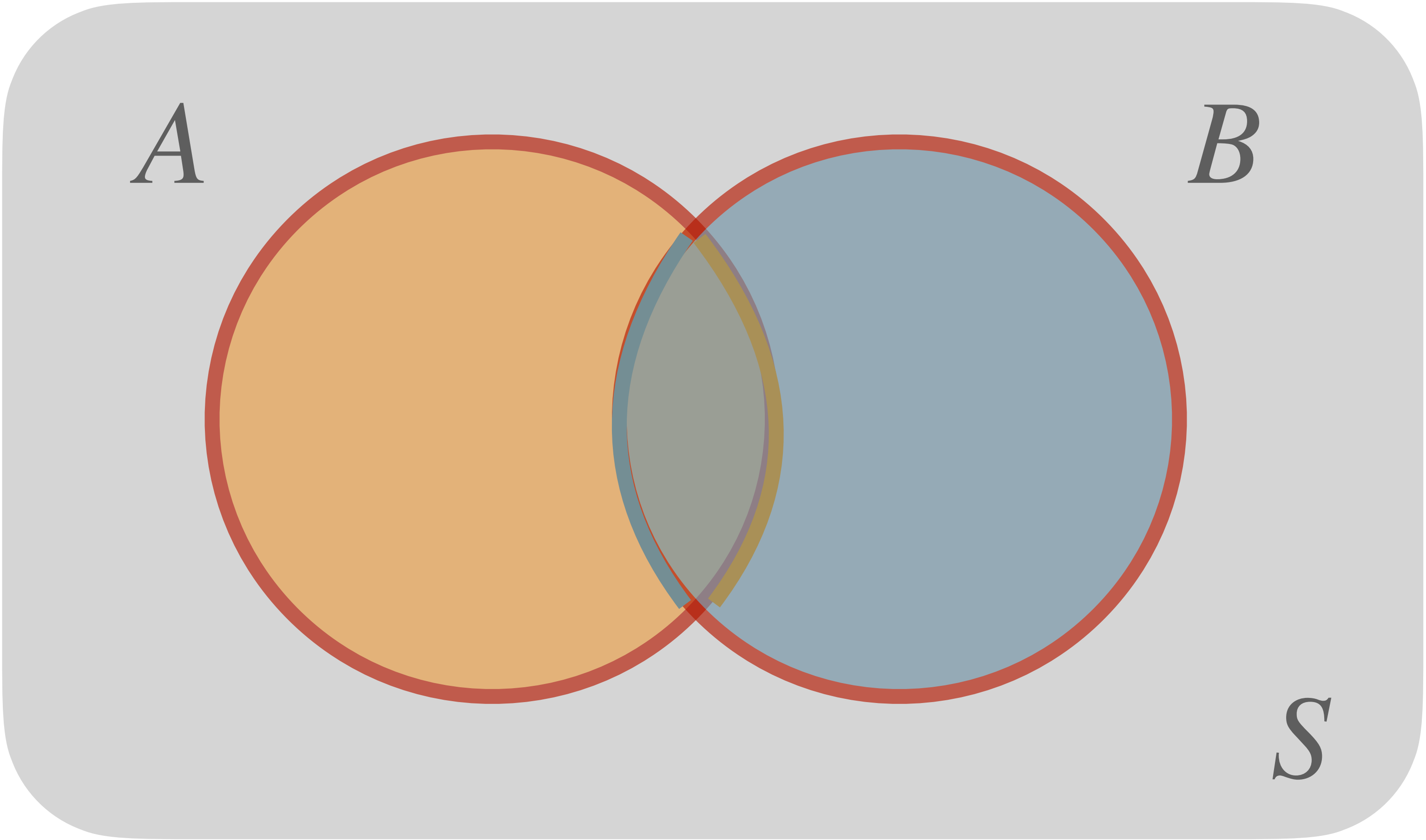
Mutually Exclusive events



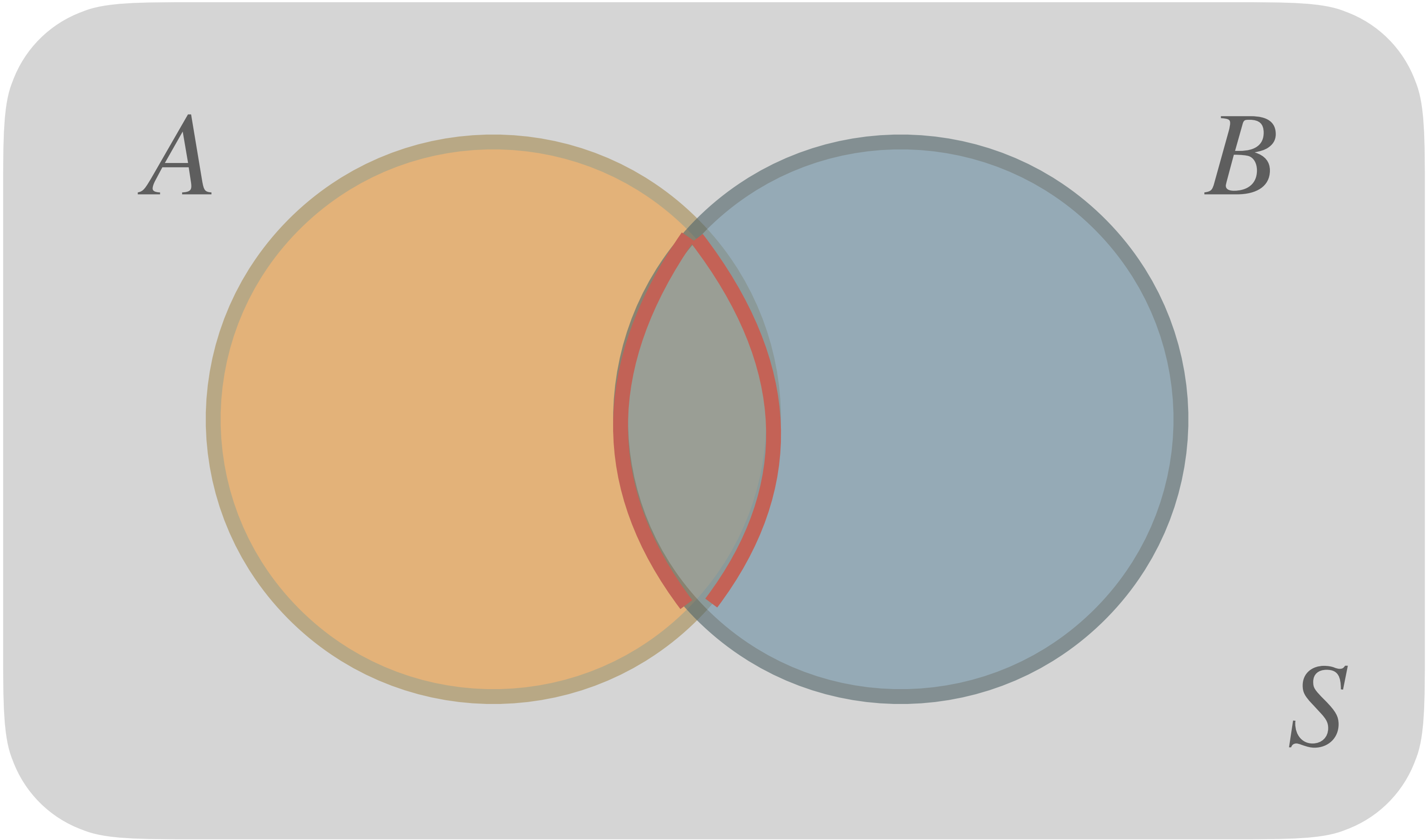
Non-mutually Exclusive events



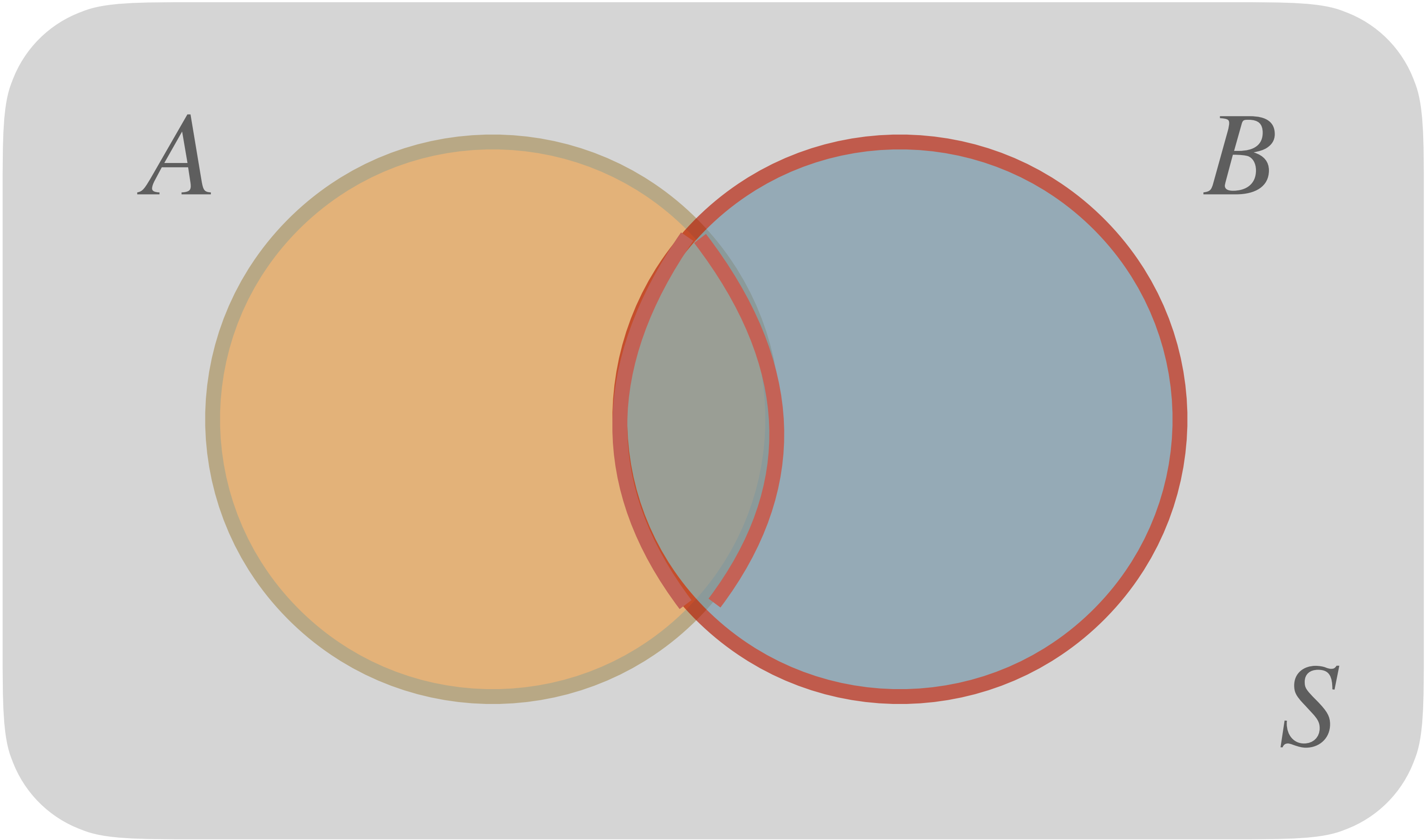
$P(A \cup B) = P(A) + P(B) - P(A \cap B)$



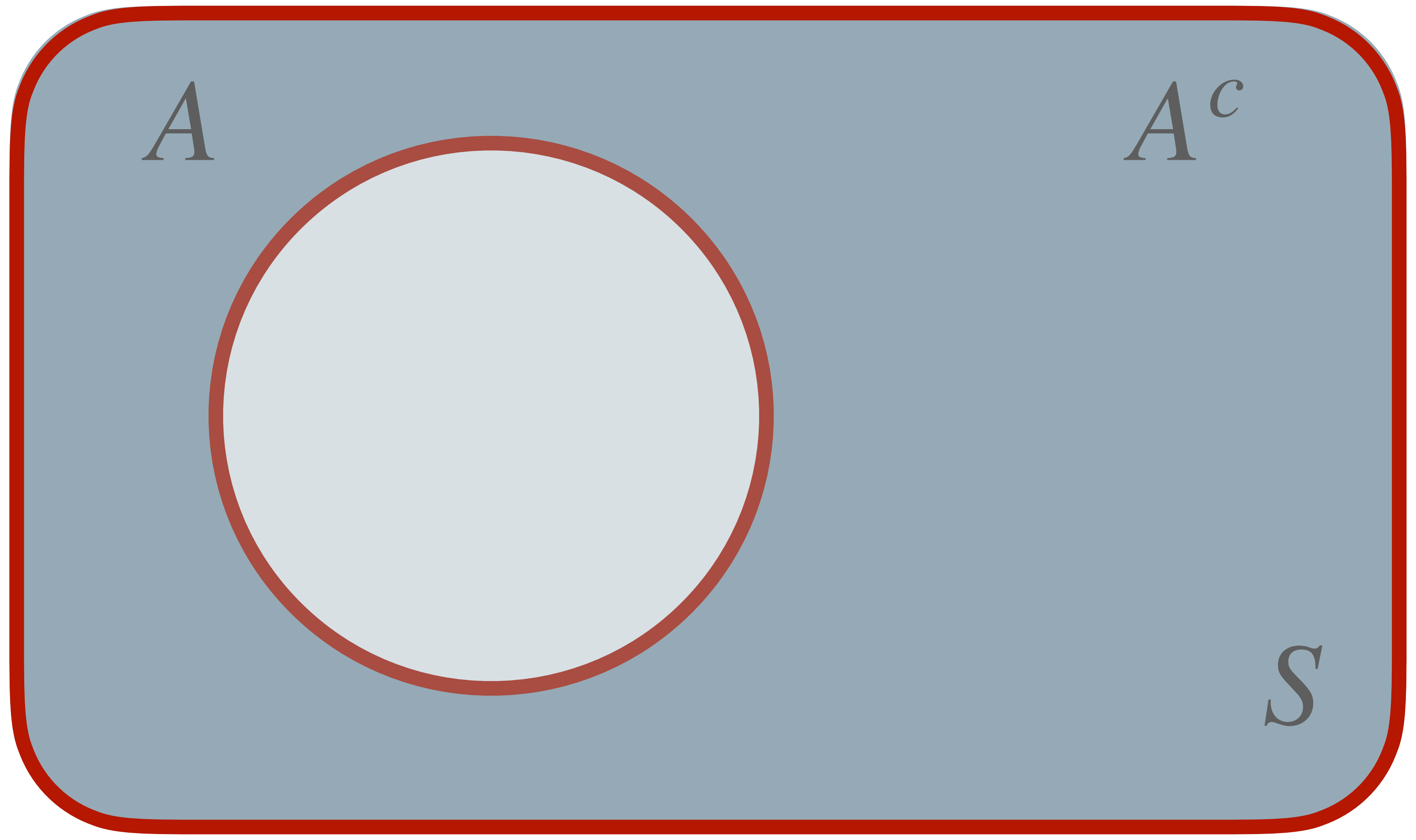
$P(A \cap B) = P(A) \times P(B \mid A)$



$P(A \mid B) = P(A \cap B)/P(B)$



Complement $P(A^c) = P(S) - P(A)$



Random Variables

The Sample Space Ω

- The set of all possible outcomes of the experiment.
- Let's consider the sample space of two successive coin tosses:

  $\Omega = \{hh, tt, ht, th\}$

- A subset of the sample space Ω is called an event E.
- Examples:

  1. Tossing a die. $\Omega = \{1,2,3,4,5,6\}$, $E = \{4\}$ describes *tossing a four.*
  2. Tossing a die. $\Omega = \{1,2,3,4,5,6\}$, $E = \{2,4,6\}$ describes *tossing an even number.*
  3. Tossing a coin twice. $\Omega = \{hh, tt, ht, th\}$, $E = \{tt, th\}$ describes *getting tails in the first toss.*

Random Variables

Power Sets

- Power set - a set of all possible subsets of set A.
- Denoted as $\mathscr{P}(A)$. *Do not confuse with probability!*
- Contains both $\varnothing$ and A.
- A size of a power set is 2^n, where n is the cardinality of set $|\mathscr{P}(A)|$.

Random Variables

Power Sets

- Example:
  - Set $A = \{1,2,3\}$
  - Subsets of set A = $\{\}, \{1\}, \{2\}, \{3\}, \{1,2\}, \{2,3\}, \{1,3\}, \{1,2,3\}$
  - $\mathscr{P}(A) = \{\{\}, \{1\}, \{2\}, \{3\}, \{1,2\}, \{2,3\}, \{1,3\}, \{1,2,3\}\}$

Random Variables

The Event Space $\mathscr{A}$

- The space of all potential *results* of the experiment.
- Describes *all* possible sets of events that can happen.
- For discrete probability distributions, $\mathscr{A}$ is often the power set of Ω.

Random Variables

The Probability Measure P

- A mapping from each event to the degree of belief that the event will occur.
- The probability of a single event lies in the interval $[0,1]$.
- $P(\Omega) = 1$

Random Variables

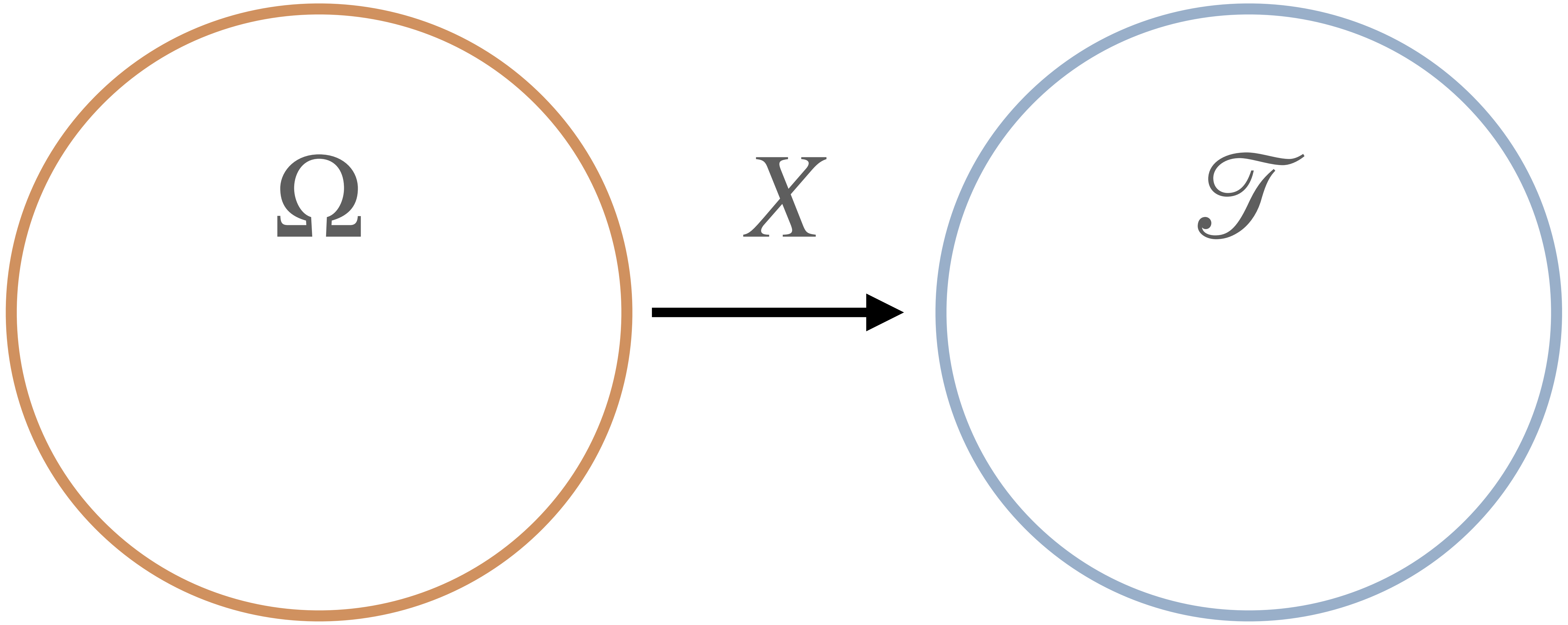
The Target Space $\mathcal{T}$

- A mapping from the sample space to the particular quantities of interest.
- Sounds a bit complicated but in fact this is a very simple idea!

Random Variables
Definition

- The term itself is misleading as it is neither random nor is it a variable.
- In fact it as a function!

$$X:\ \Omega \rightarrow \mathcal{T}$$



Random Variables

Example

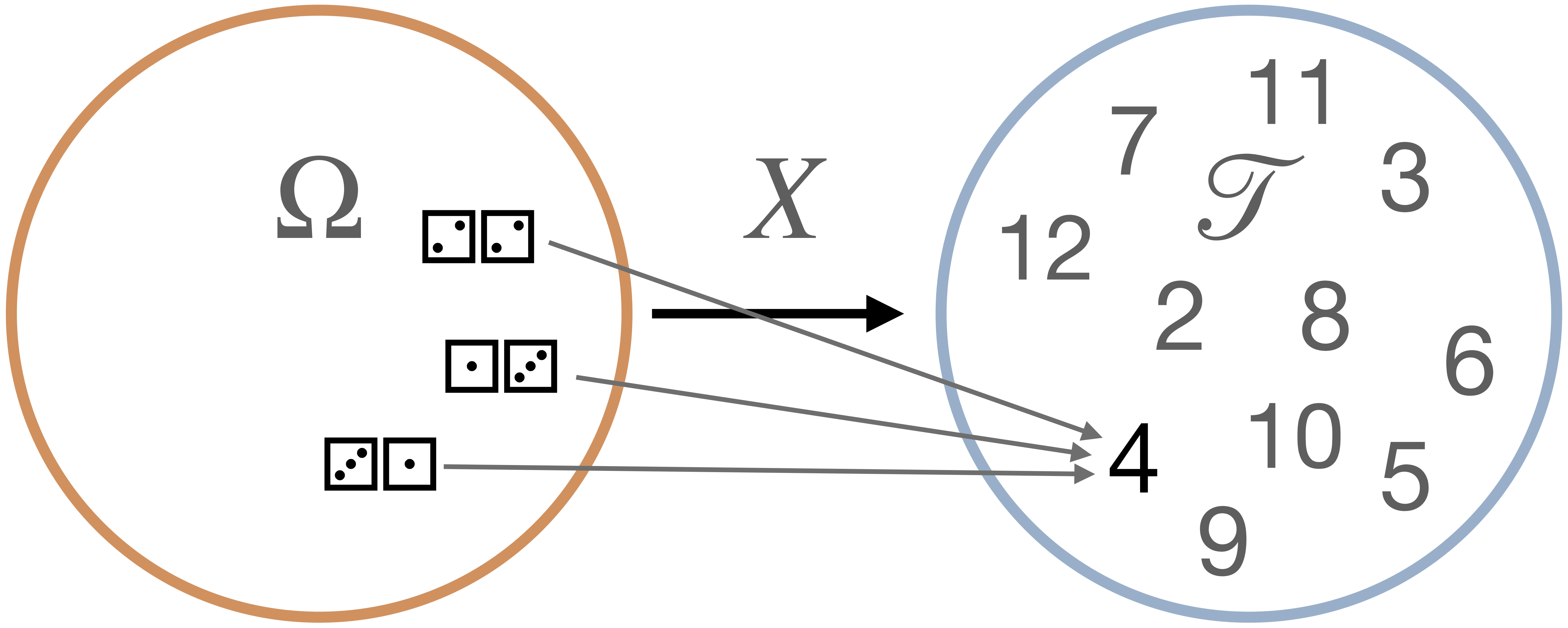
- Consider the set of all possible outcomes of throwing two six-sided dice.
- Then Ω can be represented as:

{(⚀, ⚀); (⚀, ⚁); (⚀, ⚂); (⚀, ⚃); (⚀, ⚄); (⚀, ⚅); (⚁, ⚀); (⚁, ⚁); (⚁, ⚂); (⚁, ⚃); (⚁, ⚄); (⚁, ⚅);
(⚂, ⚀); (⚂, ⚁); (⚂, ⚂); (⚂, ⚃); (⚂, ⚄); (⚂, ⚅); (⚃, ⚀); (⚃, ⚁); (⚃, ⚂); (⚃, ⚃); (⚃, ⚄); (⚃, ⚅),
(⚄, ⚀); (⚄, ⚁); (⚄, ⚂); (⚄, ⚃); (⚄, ⚄); (⚄, ⚅); (⚅, ⚀); (⚅, ⚁); (⚅, ⚂); (⚅, ⚃); (⚅, ⚄); (⚅, ⚅)}

- One possible random variable we can define is the *sum of the tosses.*
- In that case $\mathscr{T}$ will be a set of integers from 2 to 12.

Random Variables

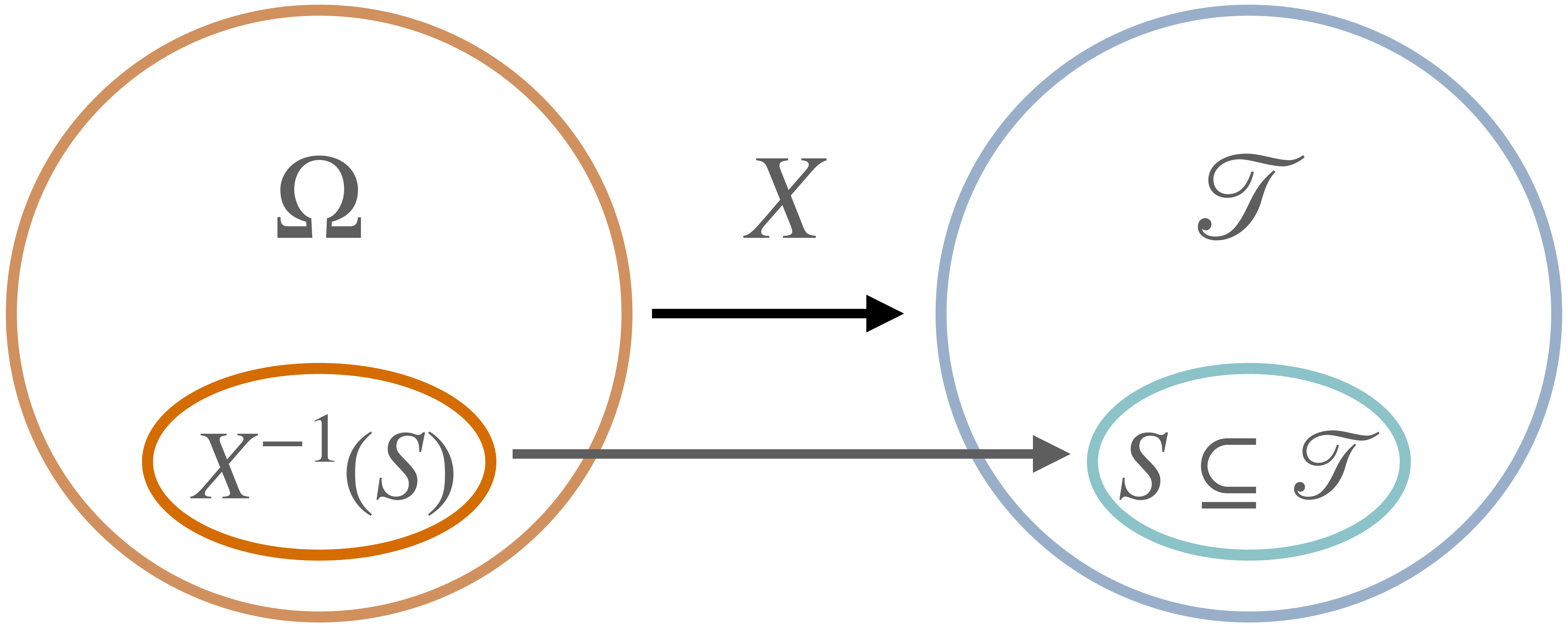
Example



Random Variables

Example

- In general:



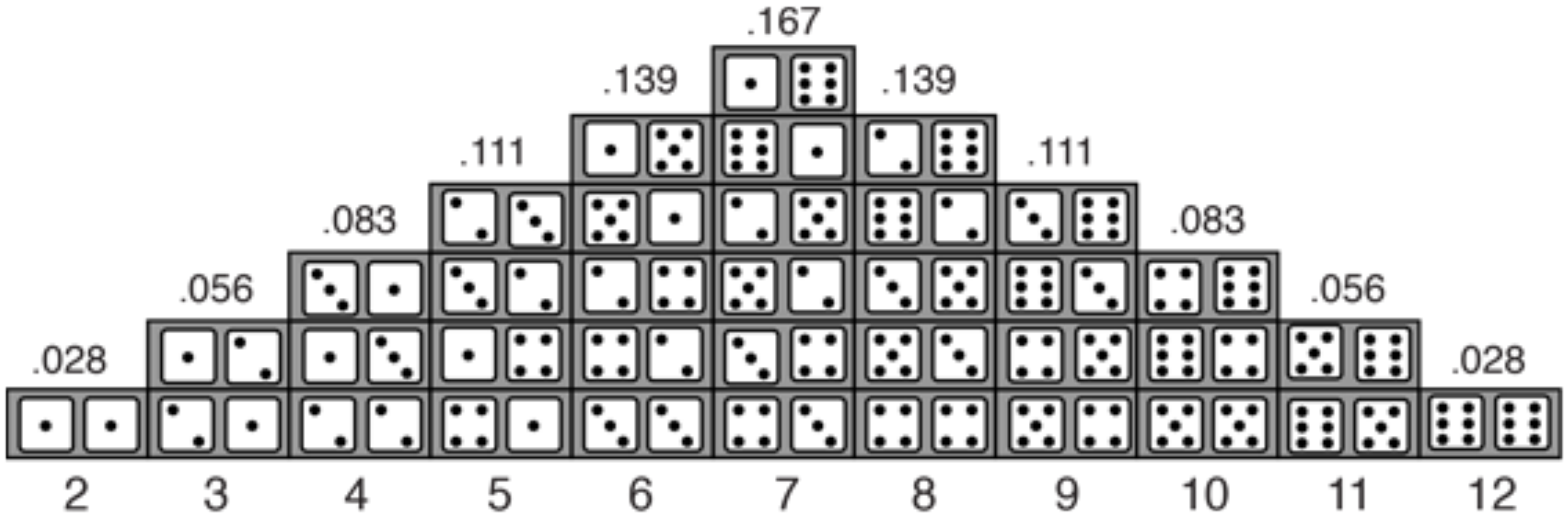
Where X^{-1} is the subset of Ω that if we apply X would lead to S.

The probability of subset S is given by: $P(S) = P(\omega \in \Omega : X(\omega) \in S)$.

Random Variables

Example

- If we assume that the game of dice is fair, i.e., all outcomes are equally likely, then the probability distribution may look like this:



Total number of microstates: 36

Random Variables

One more example

- Experiment: Drawing two coins from the bag (with replacement). The bag contains coins from USA($) and UK(£).
- Four possible outcomes: $\Omega = \{(\$, \$)(£, £), (\$, £), (£, \$)\}$.
- Assume that there are more $ coins in the bag so the probability of drawing a £ coin is 0.3.
- Event of interest: *number of times we draw a £.*

Random Variables

One more example

- Lets define a random variable X that maps Ω to $\mathscr{T}$, which denotes the number of times we draw £ out of the bag.
- From our sample space we can see that we can get 0£, 1£ and 2£s, therefore $\mathscr{T} = \{0,1,2\}$.
- Our random variable X can then be represented like the following:

$$X((£, £)) = 2$$

$$X((£, \$)) = 1$$

$$X((\$, £)) = 1$$

$$X((\$, \$)) = 0$$

Random Variables
One more example

- If we assume that draws are independent of each other, the probability can be calculated as follows:

$$P(X = 2) = P((£, £)) = P(£) \times P(£) = 0.3 \times 0.3 = 0.09$$

$$P(X = 1) = P((£, \$)) + P((\$, £)) = P(£) \times P(\$) + P(\$) \times P(£) = 0.3 \times (1 - 0.3) + (1 - 0.3) \times 0.3 = 0.42$$

$$P(X = 0) = ?$$

Random Variables

One more example

- If we assume that draws are independent of each other, the probability can be calculated as follows:

$$P(X = 2) = P((£, £)) = P(£) \times P(£) = 0.3 \times 0.3 = 0.09$$

$$P(X = 1) = P((£, \$)) + P((\$, £)) = P(£) \times P(\$) + P(\$) \times P(£) = 0.3 \times (1 - 0.3) + (1 - 0.3) \times 0.3 = 0.42$$

$$P(X = 0) = P((\$, \$)) = P(\$) \times P(\$) = (1 - 0.3) \times (1 - 0.3) = 0.49$$

Discrete & Continuous Probabilities

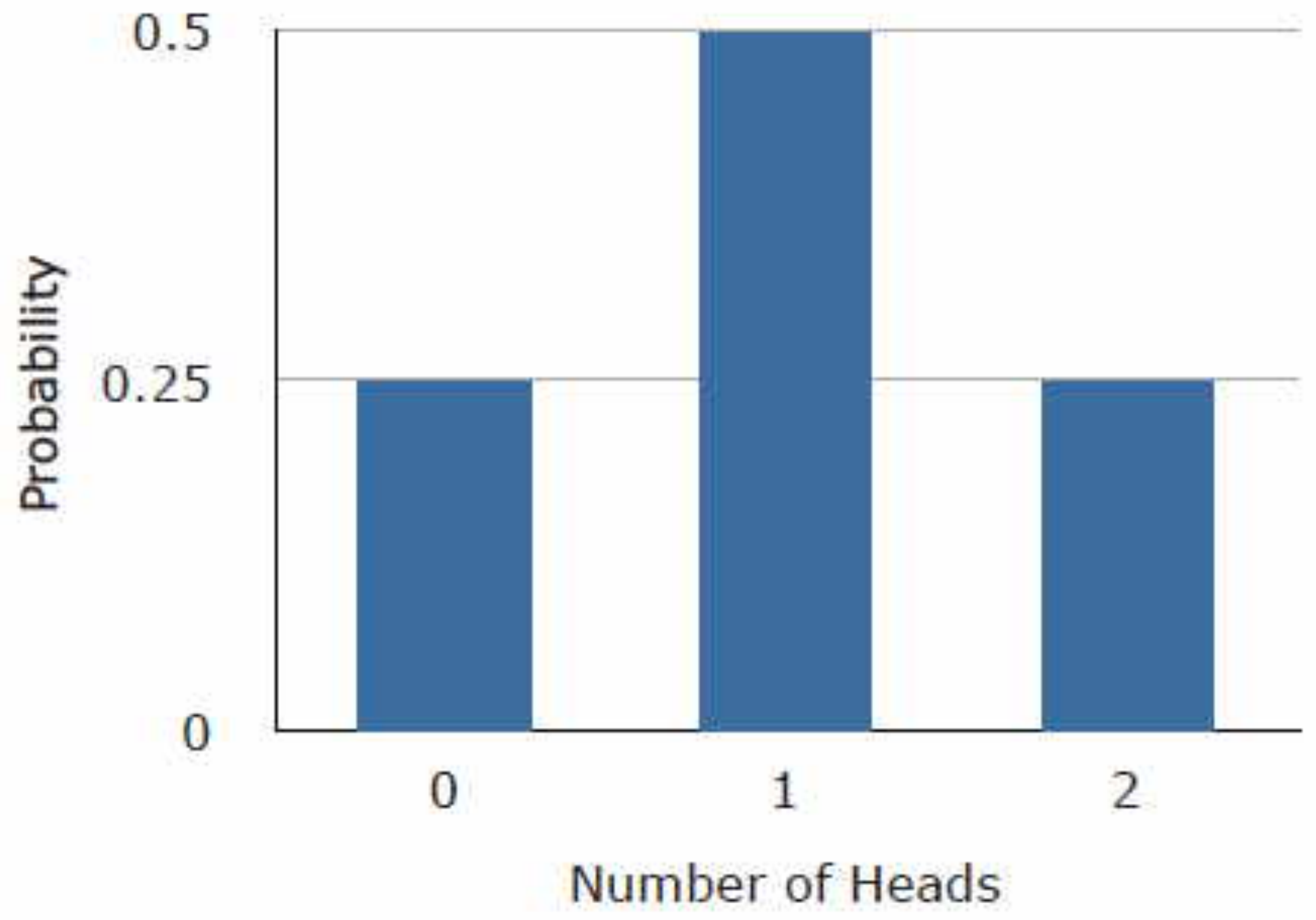
Introduction

- It is important to understand the difference between target space types.
- Discrete random variables:
  - Variable can take on a *discrete* set of values.
  - Value can be obtained by counting.
- Continuous random variables:
  - Variable can take on a *continuous* set of values.
  - Value can be obtained by measuring.

Discrete & Continuous Probabilities

Discrete

- The probability that a random variable X takes a particular value $x \in \mathcal{T}$ is denoted as $P(X = x)$.
- This expression is also called ***probability mass function.***



$$P(X = 0) = 0.25$$
$$P(X = 1) = 0.5$$
$$P(X = 2) = 0.25$$

Discrete & Continuous Probabilities

Discrete

- When the target space is discrete we can imagine the probability distribution of multiple random variables as a multidimensional array of numbers

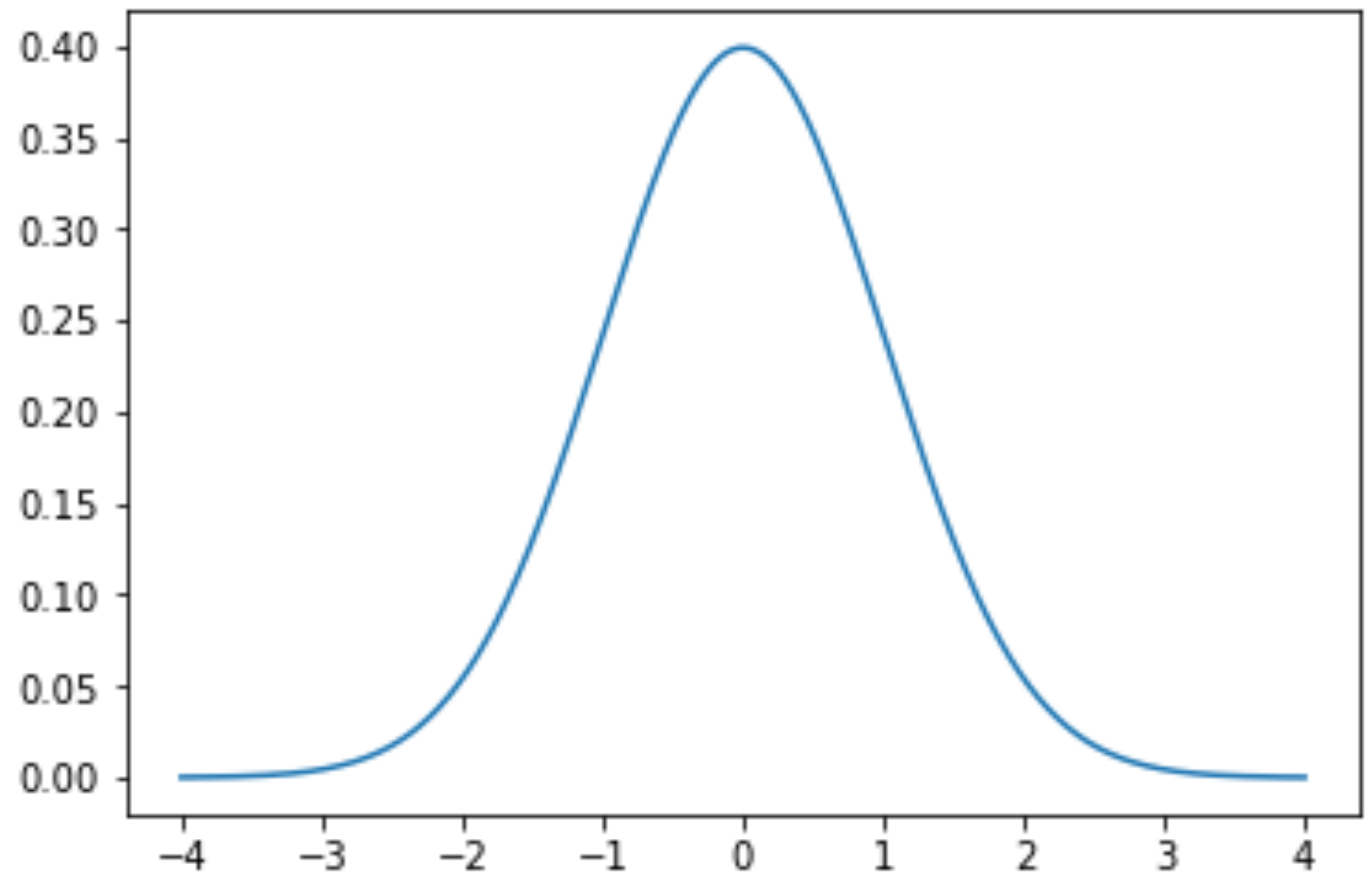
	x_1	x_2	x_3	x_4	x_5
y_3	0.1	0.07	0.06	0.03	0.1
y_2	0.12	0.09	0.02	0.01	0.05
y_1	0.18	0.01	0.11	0.02	0.03

- *Joint probability* is defined as $p(x, y) = P(X = x_i, Y = y_j)$
- *Marginal probability* $p(x)$ represents the probability that X takes the value x_i irrespective to the value of Y.
- *Conditional probability* $p(y \mid x)$ will only consider the value of Y for a particular value of X.

Discrete & Continuous Probabilities

Continuous

- Target spaces are intervals of the real line $\mathbb{R}$.
- A ***probability density function*** is a function whose value at any given point in the sample space can be interpreted as providing a *relative likelihood* that the value of the random variable would be close to that sample.



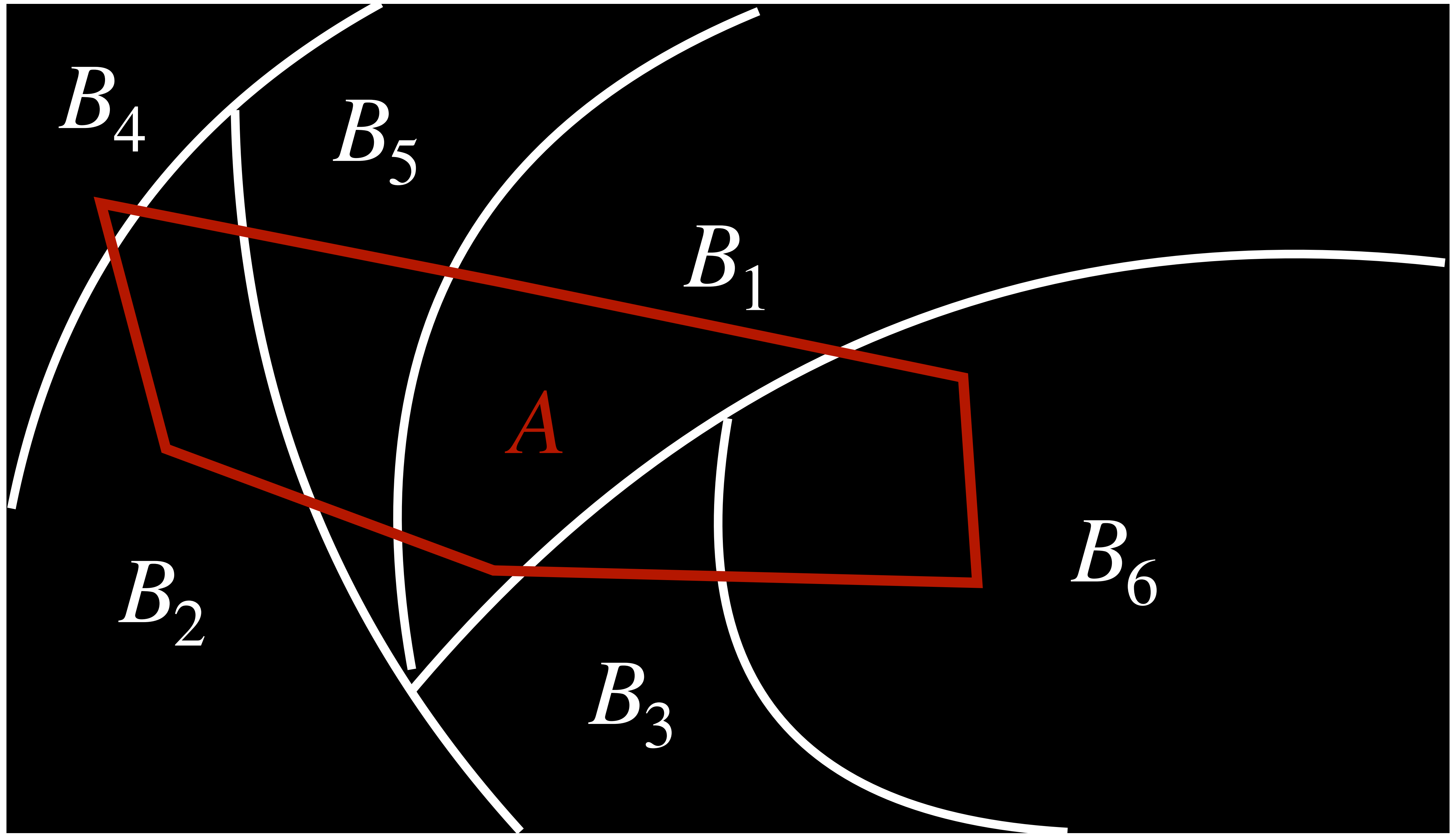
The Law of Total Probability

- Suppose $B_1, ..., B_n$ are mutually exclusive and collectively exhaustive events in a sample space. We can then sum/integrate over the set of states of variable B to get a *marginal* distribution of variable A.

$$P(A) = \sum_i^n P(A \mid B_i)\, P(B_i) = \sum_i^n P(A \cap B_i)$$

The Law of Total Probability

- Mutually exclusive - no overlap.
- Collectively exhaustive - cover the whole space.



The Law of Total Probability

Example

- Three robots are making parts at a factory. We know that:
  - Robot 1 makes 60% of the parts.
  - Robot 2 makes 30% of the parts.
  - Robot 2 makes 10% of the parts.
- Some parts that are produced are defective:
  - Of the parts Robot 1 makes, 7% are defective.
  - Of the parts Robot 2 makes, 15% are defective.
  - Of the parts Robot 3 makes, 30% are defective.
- What is the probability of a randomly selected part being defective?

The Law of Total Probability
Example

- Three robots are making parts at a factory. We know that:
  - Robot 1 makes 60% of the parts. $P(R_1) = 0.6$
  - Robot 2 makes 30% of the parts. $P(R_2) = 0.3$
  - Robot 2 makes 10% of the parts. $P(R_3) = 0.1$
- Some parts that are produced are defective:
  - Of the parts Robot 1 makes, 7% are defective. $P(D \mid R_1) = 0.07$
  - Of the parts Robot 2 makes, 15% are defective. $P(D \mid R_2) = 0.15$
  - Of the parts Robot 3 makes, 30% are defective. $P(D \mid R_3) = 0.3$
- What is the probability of a randomly selected part being defective? $P(D) = ?$

The Law of Total Probability

Example

$$
\begin{aligned}
P(D) &= P(D \mid R_1)P(R_1) + P(D \mid R_2)P(R_2) + P(D \mid R_3)P(R_3) \\
&= P(D \cap R_1) + P(D \cap R_2) + P(D \cap R_3) \\
&= 0.042 + 0.045 + 0.03 \\
&= 0.117
\end{aligned}
$$

Expectation and Variance

Definition

- *Expected Value/Mean* gives the weighted average of all possible outcomes of the random variable. *Is not an expected outcome, but a theoretical mean!*

$$\mathbb{E}(X) = \sum x\, p(x)$$

- *Variance* represents the dispersion, i.e. how far a set of numbers is spread from the mean.

$$Var(X) = \mathbb{E}[(X-\mu)^2] = \sum_x (x-\mu)^2\, p(x)$$

- *Standard Deviation* is simply a square root of the variance

$$\sigma_X = \sqrt{Var(X)}$$

Expected Value

Exercise

- A computer randomly chooses 4 numbers in range $[0,10]$.
- We play a game and try to guess all 4 numbers.
- We pay 3$ for each game.
- If we win we get 10.000$.
- What is out expected profit in a long run?

Expected Value

Exercise

- Let X be a random variable representing profit on each play.
- $X \in \{-3, 9997\}$.
- The probability of making a correct guess…

Expected Value
Exercise

- Let X be a random variable representing profit on each play.
- $X \in \{-3, 9997\}$.
- The probability of making a correct guess is $0.1^4 = 0.0001$.

X	$P(X)$	$\mathbb{E}(X)$
-3	0.9999	?
9997	0.0001	

Expected Value

Exercise

- Let X be a random variable representing profit on each play.
- $X \in \{-3, 9997\}$.
- The probability of making a correct guess is $0.1^4 = 0.0001$.

X	$P(X)$	$\mathbb{E}(X)$
-3	0.9999	-2
9997	0.0001	

Expected Value

One more exercise

- Calculate the expectation for the given random random variable:

x	0	1	2
$p(x)$	0.16	0.48	0.36

$$\mathbb{E}(X) = \sum_x x\,p(x) = 0 \times 0.16 + 1 \times 0.48 + 2 \times 0.36 = 1.2$$

Expected Value

One more exercise

- Calculate the mean for the given random random variable:

x	0	1	2
$p(x)$	0.16	0.48	0.36

$$\mathbb{E}(X) = \sum_x x\, p(x) = 0 \times 0.16 + 1 \times 0.48 + 2 \times 0.36 = 1.2$$

- Now try to calculate the Variance:

$$Var(X) = \sum_x (x - \mu)^2 p(x) = ?$$

Expected Value

One more exercise

- Calculate the mean for the given random random variable:

x	0	1	2
$p(x)$	0.16	0.48	0.36

$$\mathbb{E}(X) = \sum_x x\,p(x) = 0 \times 0.16 + 1 \times 0.48 + 2 \times 0.36 = 1.2$$

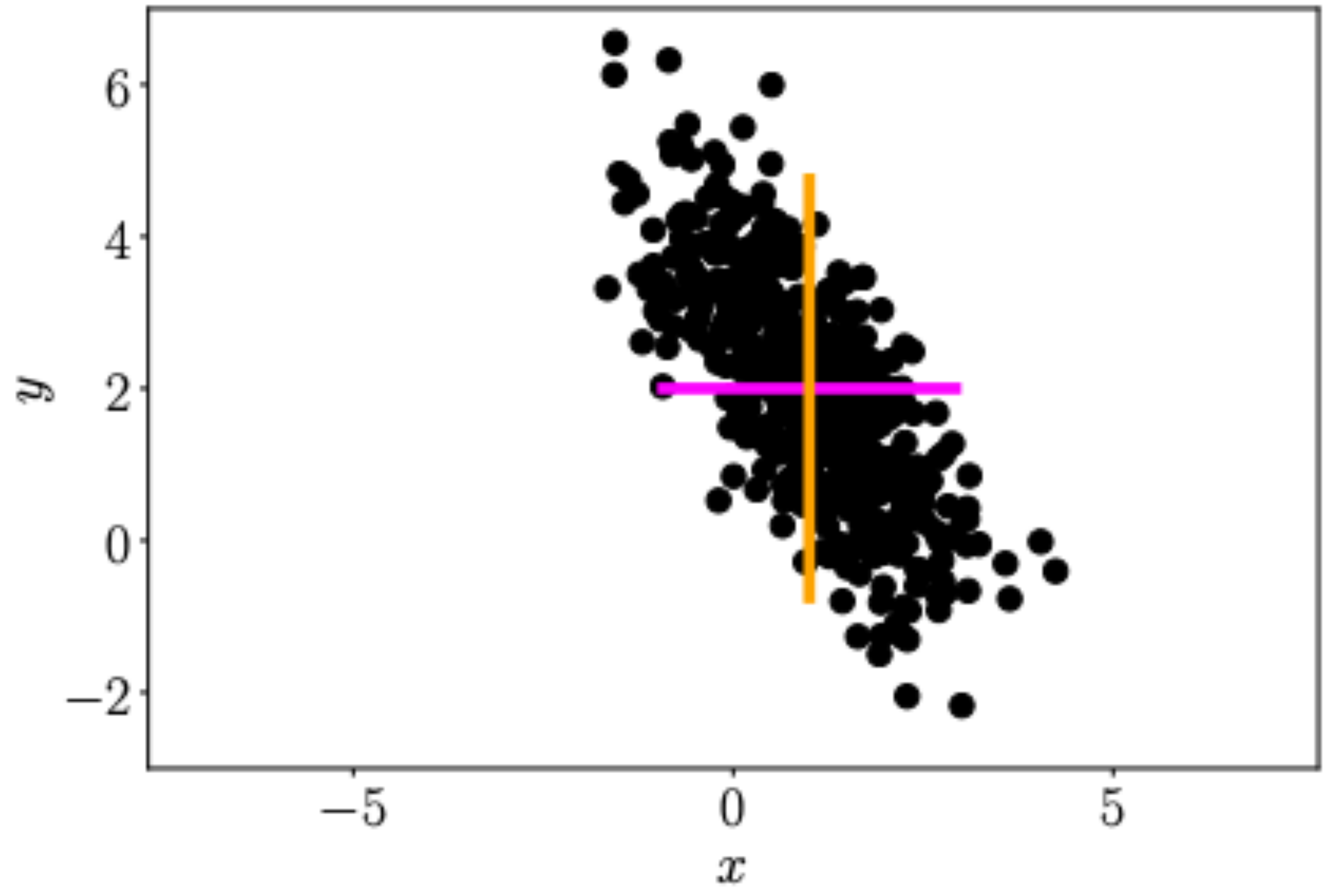
- Now try to calculate the Variance:

$$Var(X) = \sum_x (x - \mu)^2 p(x) = 0.48 \qquad \sigma_X = \sqrt{Var(X)} = \sqrt{48}$$

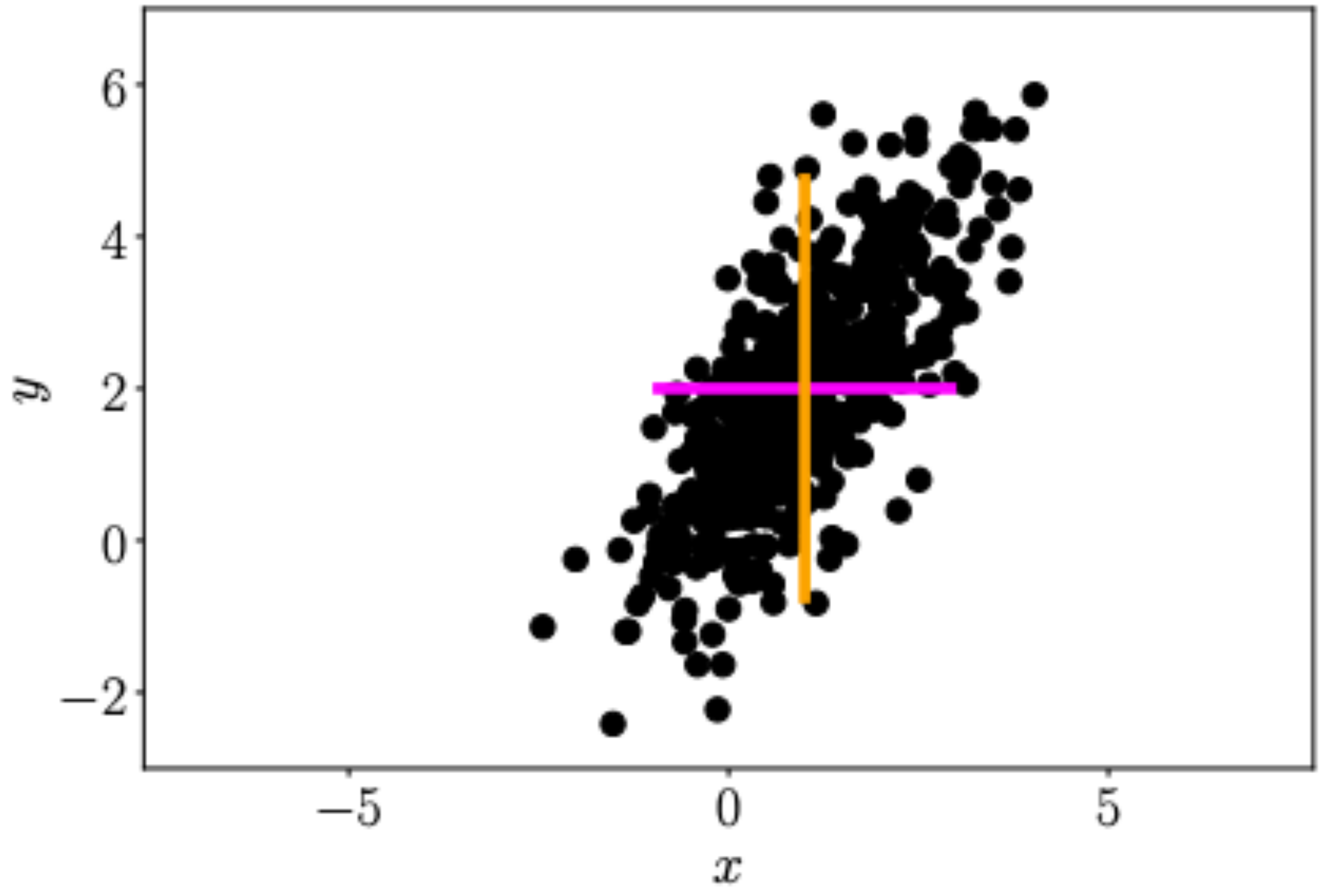
Covariance

- *Covariance* of two univariate random variables $X, Y \in \mathbb{R}$ is given by the expected product of their deviations from their respected means.

$$Cov(X, Y) = \mathbb{E}_{X,Y}[(x - \mathbb{E}_X[x])][(y - \mathbb{E}_Y[y])]$$



(a) x and y are negatively correlated.



(b) x and y are positively correlated.

Correlation

- *Correlation* is the normalized form of Covariance.

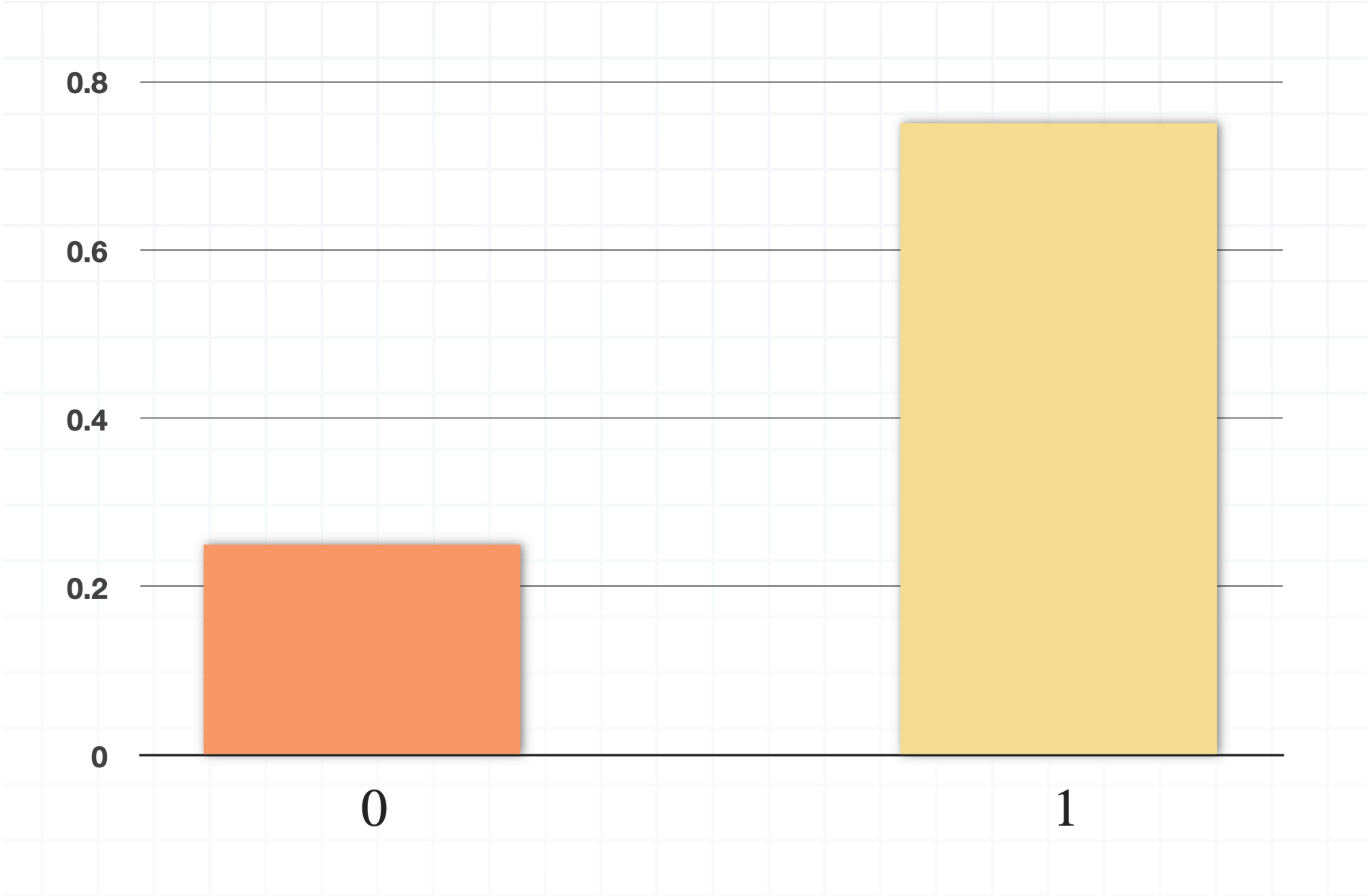
$$Corr[X, Y] = \frac{Cov(X, Y)}{\sqrt{Var(X)Var(Y)}} \in [-1, 1]$$

- Is useful when we want to compare the covariances between different pairs of random variables.

Distributions

Bernoulli distribution

- Models the set of possible outcomes for a *single* experiment.
- $X \in \{0,1\}$
- Parameter $\rho \in [0,1]$ reflects the probability of getting a 1.
- PMF: $f(x; p) = \rho^x(1-\rho)^{1-x}$
- $\mathbb{E}[X] = p$
- Example: tossing a biased coin.



Distributions

Binomial distribution

- A generalization of Bernoulli for $\mathbb{N}$ random variables, i.e. $X \in \mathbb{N}$.
- Parameters $\rho \in [0,1], n \in \mathbb{N} = 0,1,2,3,...$
- PMF: $f(x; p, n) = \binom{n}{k} \rho^x (1 - \rho)^{n-x}$
- $\mathbb{E}[X] = n\rho$

Distributions

Binomial distribution

- Let's have a closer look at the PMF:

$$f(x; p, n) = \binom{n}{k} \rho^x (1 - \rho)^{n-x}$$

- Combination

$$\binom{n}{k} = \frac{n!}{n!(n-k)!}$$

Where n is the total number of possible outcomes,

k is number of items you want to rearrange.

Useful Links

1. Intelligent Systems Lab YouTube channel
2. jbstatistics
3. 3Blue1Brown Probability of Probabilities
4. https://mml-book.github.io/
5. StatQuest!!!