ME 233 Advanced Control II

Lecture 2
Introduction to Probability Theory

(ME233 Class Notes pp. PR1-PR3)

Outline

- Sample Space and Events
- Probability function
- Discrete Random Variables
- Probability mass function, expectation and variance

Sample Space and Events

Assume:

- We do an experiment many times.
  - Each time we do an experiment we call that a ***trial***
- The outcome of the experiment may be different at each trial.

ω_i **:** The i^{th} possible outcome of the experiment

Sample Space and Events

Sample Space Ω :

The space which contains all possible outcomes of an experiment.

$$\Omega = \{\omega_1, \omega_2, \cdots, w_n\}$$

ω_i **:** The i[th] possible outcome of the experiment

Each outcome is an element of Ω

Example: Dice

***Experiment*:**

A situation whose ***outcome*** depends on chance

- throwing a die once

Sample Space Ω

The set of <u>all possible</u> ***outcomes*** of an experiment

$$\Omega = \{ ⚀, ⚁, ⚂, ⚃, ⚄, ⚅ \}.$$

Events

Event S_j :

Is a subset of the union of the sample space Ω and the empty set ϕ

If a sample space has n outcomes:

$$\Omega = \{\omega_1, \omega_2, \cdots, w_n\}$$

There are 2^n events:

$$\mathcal{S} = \{S_1, \cdots, S_{2^n}\}$$

Probability - events

Experiment: throwing a die once

$$\Omega = \{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\}.$$

Outcomes*: elements of the sample space* S

Events*: Are <u>subsets</u> of the sample space* S

An event occurs if any of the outcomes in that event occurs.

Empty subsets are ***null*** *or* ***impossible events***

Probability - events

Experiment: throwing a die once

$$\Omega = \{\text{⚀}, \text{⚁}, \text{⚂}, \text{⚃}, \text{⚄}, \text{⚅}\}.$$

<u>Some events:</u>

- The event E of observing an even number of dots:

$$E = \{\text{⚁}, \text{⚃}, \text{⚅}\}$$

- The event O of observing an odd number of dots:

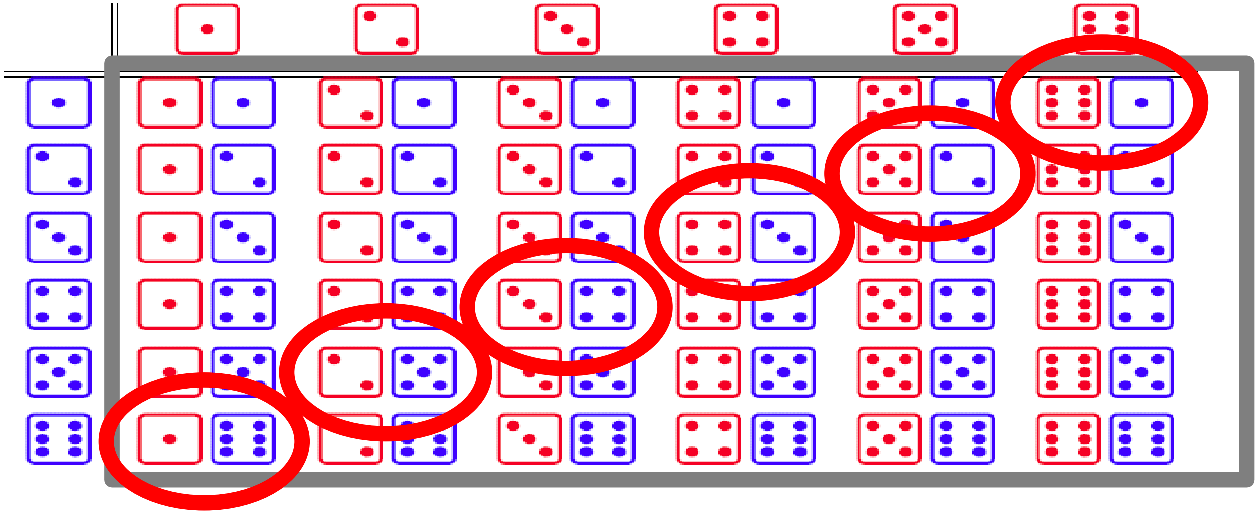
$$O = \{\text{⚀}, \text{⚂}, \text{⚄}\}.$$

Example: throwing a pair of dice

(one red and one blue)

– the sample space has ***36*** outcomes:

$\Omega =$



• *The event* L *of obtaining the number* ***7*** *is*

$$L = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$$

L occurs if any of the outcomes in L occurs.

Union, Complement and Intersection

For a sample space $\Omega = \{\omega_1, \omega_2, \cdots, w_n\}$

And the set of all events $\mathcal{S} = \{S_1, \cdots, S_{2^n}\}$

- Union of two events (or):

$$S_i \cup S_j = \{\omega_m \,|\, \omega_m \in S_i \text{ or } \omega_m \in S_j\}$$

- Intersection of two events (and):

$$S_i \cap S_j = \{\omega_m \,|\, \omega_m \in S_i \text{ and } \omega_m \in S_j\}$$

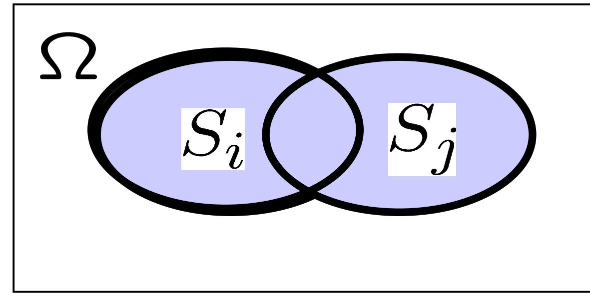
- Complement of an event (not):

$$\backslash S_i = \{\omega_m \,|\, \omega_m \in \Omega \cup \phi \text{ and } \omega_m \notin S_i\}$$

Union, Complement and Intersection

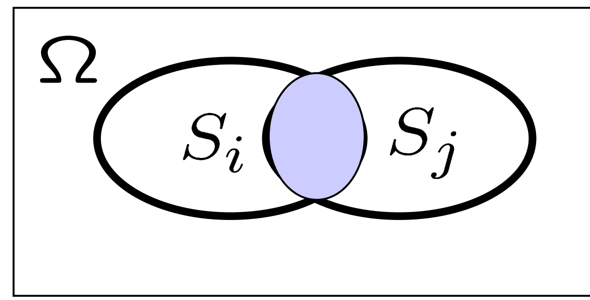
- Union of two events:

$$S_i \cup S_j$$



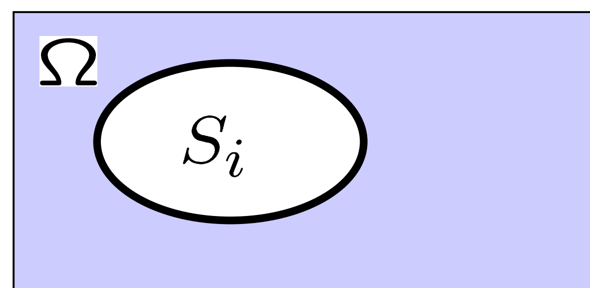
- Intersection of two events:

$$S_i \cap S_j$$



- Complement of an event:

$$\backslash S_i = S_i^c$$

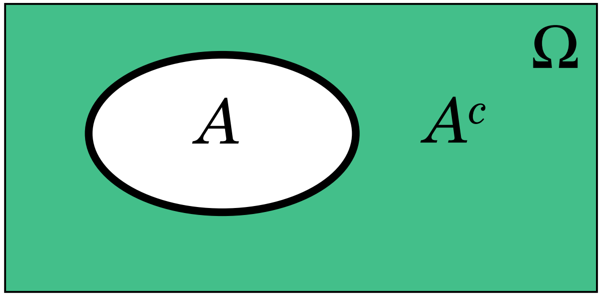


Complement

- The **complement** of an event $\boldsymbol{A}$, denoted by $\boldsymbol{A^c}$, is the set of outcomes that are not in $\boldsymbol{A}$

- $\boldsymbol{A^c}$ occurring means that $\boldsymbol{A}$ *does not occur*

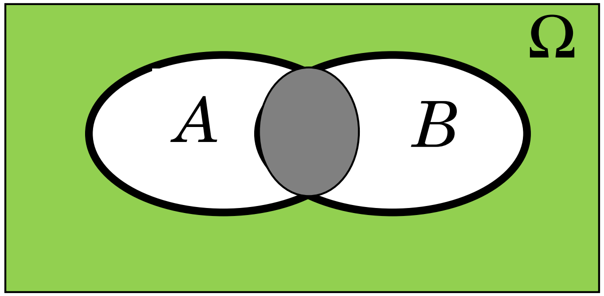
$$A^c = \{\omega \mid \omega \in \Omega \text{ and } \omega \notin A\}$$



Intersection of two events

- The **<u>intersection</u>** of two events $\boldsymbol{A}$ and $\boldsymbol{B}$, denoted by $\boldsymbol{A} \cap \boldsymbol{B}$, is the set of outcomes that are in $\boldsymbol{A}$, **<u>and</u>** $\boldsymbol{B}$.

- *If* the event $\boldsymbol{A} \cap \boldsymbol{B}$ *occurs, then* ***<u>both</u>*** $\boldsymbol{A}$ *and* $\boldsymbol{B}$ *occur*



- *E*vents $\boldsymbol{A}$ and $\boldsymbol{B}$ are <u>mutually exclusive</u> if they cannot both occur at the same time, i.e. if

$$A \cap B = \emptyset$$

Example of Intersection of two events

Experiment: throwing of a dice once

$$\Omega = \{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\}.$$

- Events E and O are mutually exclusive

$$E = \{⚁, ⚃, ⚅\} \qquad O = \{⚀, ⚂, ⚄\}.$$

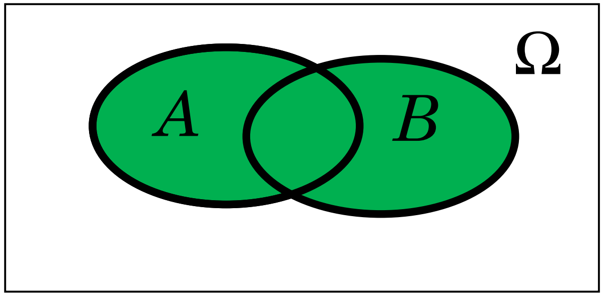
$$E \bigcap O = \varnothing$$

Union of two events

- The **union** of two events $\boldsymbol{A}$ and $\boldsymbol{B}$, denoted by

$\boldsymbol{A} \cup \boldsymbol{B}$, is the set of outcomes that are in $\boldsymbol{A}$, or $\boldsymbol{B}$, or both

- *If* the event $\boldsymbol{A} \cup \boldsymbol{B}$ *occurs, then either* $\boldsymbol{A}$ *or* $\boldsymbol{B}$ *or both occur*



Probability function

We now consider the probability that a certain event occurs.

Recall: An event occurs if any of the outcomes in that event occurs.

The probability of event $\boldsymbol{A}$ will be denoted by

$$P(A)$$

Probability

A number between 0 and 1, inclusive, that indicates **<u>how likely an event is to occur</u>**.

– An event with probability of 0 is a **<u>null event.</u>**

– An event with probability of 1 is a **<u>certain event.</u>**

- Probability of event $\boldsymbol{A}$ is denoted as $\boldsymbol{P(A)}$.
- The closer $\boldsymbol{P(A)}$ to 1, the more likely is $\boldsymbol{A}$ to happen.

Intuitive Notion of Probability

The probability of event $\boldsymbol{A}$ is

$$P(A) = \frac{\text{Possible outcomes associated with } A}{\text{Total possible outcomes}}$$

(Assumes each outcome is equally likely)

$$0 \leq P(A) \leq 1$$

Assigning Probability - Frequentist approach

- An experiment is repeated $\boldsymbol{n}$ times under essentially identical conditions

- if the event $\boldsymbol{A}$ occurs $\boldsymbol{m}$ times and $\boldsymbol{n}$ is large

$$P(A) \approx \frac{m}{n}$$

Dice example

Experiment: throwing a **fair** die once

Ω = {⚀, ⚁, ⚂, ⚃, ⚄, ⚅} $\Omega = \{1, 2, 3, 4, 5, 6\}$

- $P(\Omega) = 1$
- $P(1) = 1/6, \quad P(3) = 1/6, P(6) = 1/6$
- $P(\text{even number}) = 3/6 = 1/2$
- $P(\text{odd number}) = 3/6 = 1/2$

Example: poker

Example: In poker you are dealt 5 cards from a deck of 52

- What is the probability of being dealt four of a kind?

  – e.g. 4 aces or four kings, and so fourth?

$$P(\text{four of a kind}) = ?$$

Example: poker

Solution:

1. There are only 48 possible hands containing 4 aces, another 48 containing 4 kings, etc.

2. Thus, there are **13 x 48** possible "four of a kind" hands.

3. The possible number of hands is obtained from the combination formula for "52 things taken 5 at a time":

total possible outcomes: $$\binom{52}{5} = \frac{52!}{5!(52-5)!} = 2,598,960$$

4. Thus, $$P(\text{four of a kind}) = \frac{13 \times 48}{2,598,960} = 0.00024$$

Probability Space

The probability space is the triple:

$$(\Omega, \mathcal{S}, P)$$

Where

- Ω is the sample space
- $\mathcal{S}$ the set of all possible events
- $P : \mathcal{S} \to [0, 1]$ is the probability function

Probability function

Probability function: $P : \mathcal{S} \rightarrow [0, 1]$

Satisfies 3 axioms:

1. $P(S_i) \geq 0, \qquad \forall S_i \in \mathcal{S}$

2. $P(\Omega) = 1$

3. $P(S_i \cup S_j) = P(S_i) + P(S_j)$ if $S_i \cap S_j = \emptyset$

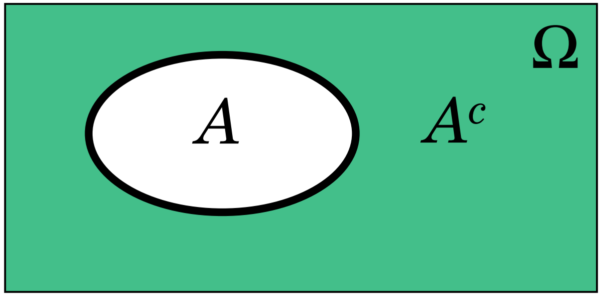
where $S_i, S_j \in \mathcal{S}$

Complement

- The **complement** of an event $\boldsymbol{A}$, denoted by $\boldsymbol{A^c}$, is the set of outcomes that are not in $\boldsymbol{A}$

- $\boldsymbol{A^c}$ occurring means that $\boldsymbol{A}$ *does not occur*

$$A^c = \{\omega \mid \omega \in \Omega \text{ and } \omega \notin A\}$$



$$\boxed{P(A^c) = 1 - P(A)}$$

Independent Events

- Two events are **<u>independent</u>** if

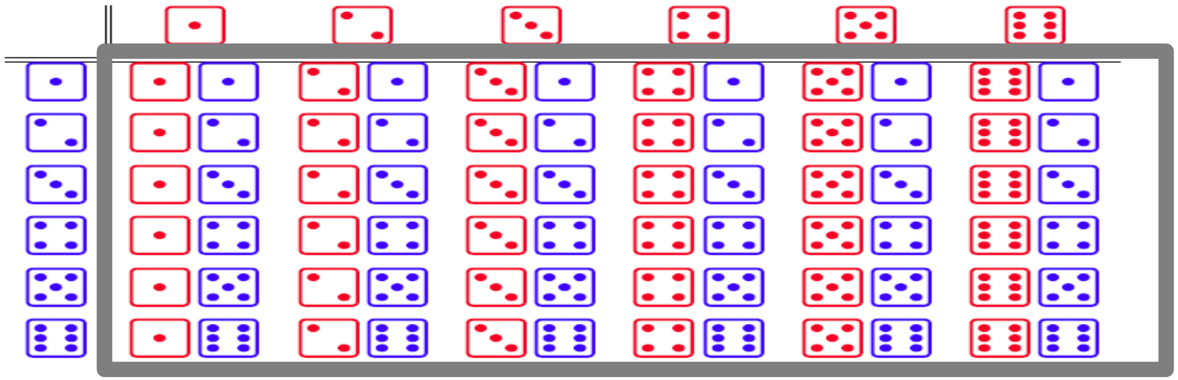
$$\boxed{P(A \cap B) = P(A) \times P(B)}$$

- Intuitively, two events are independent if the events do not influence each other:
  - Event A occurring does not affect the chances of B occurring, and vice versa.

Example of independence

Experiment: throwing a pair of dice (one red and one blue)

$\Omega =$



36 possible outcomes

The probability of throwing a red 1 and a blue 5 is

$$P(1 \cap 5) = 1/36$$

$$= 1/6 \times 1/6 = P(1) \times P(5)$$

Law of Union

- Recall: If A and B are mutually exclusive

$$\boxed{P(A \cup B) = P(A) + P(B)}$$

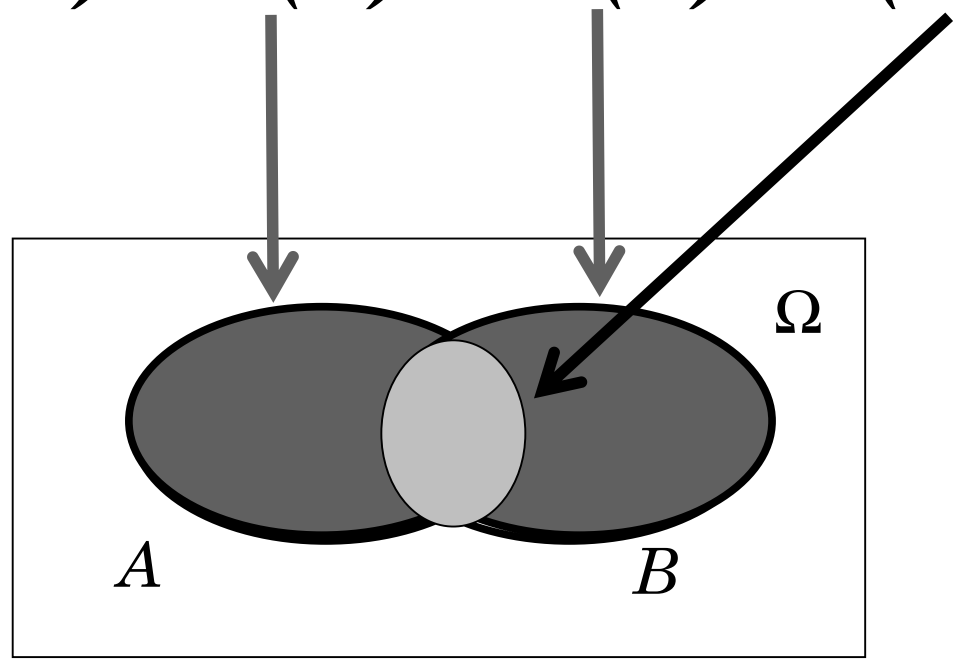
- If A and B are not mutually exclusive

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

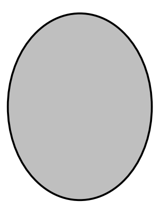
Law of Union

- If $\boldsymbol{A}$ and $\boldsymbol{B}$ are <u>not</u> mutually exclusive

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



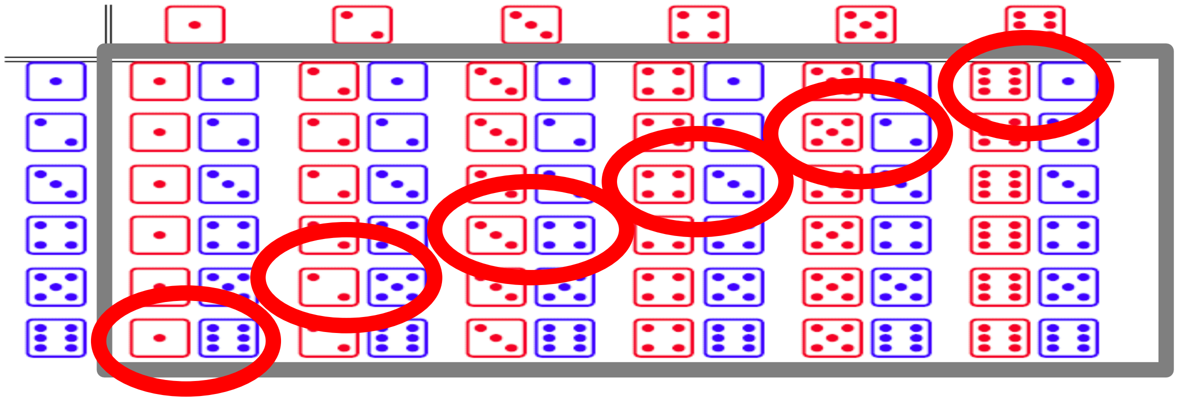
don't count



twice!

Example

Experiment: throwing a pair of dice (one red and one blue)

$\Omega =$ 

- $P(L)$ = *the probability of obtaining a* ***7***

$$L = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$$

$$P(L) = P(1,6) + P(2,5) + P(3,4) + P(4,3) + P(5,2) + P(6,1)$$

$$P(L) = 6/36 = 1/6$$

Joint Probability

Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be two events

$$P(A \cap B)$$

is often called the ***joint probability*** of A and B

$$P(A) \qquad\qquad P(B)$$

are often called the ***marginal probabilities*** of $\boldsymbol{A}$ and $\boldsymbol{B}$

Conditional Probability

Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be two events and $P(B) \neq 0$

The conditional probability of event $\boldsymbol{A}$ given that event $\boldsymbol{B}$ has occurred is

$$\boxed{P(A|B) = \frac{P(A \cap B)}{P(B)}}$$

Bayes' Rule

Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be two events

$$\boxed{\begin{aligned} P(A|B)P(B) &= P(B|A)P(A) \\ &= P(A \cap B) \end{aligned}}$$

Independence

The following are equivalent:

1. $\boldsymbol{A}$ and $\boldsymbol{B}$ *are* ***independent***
2. $P(A \cap B) = P(A)\, P(B)$
3. $P(A|B) \;=\; P(A)$
4. $P(B|A) \;=\; P(B)$

Array of Probabilities

Let C and D be two chance experiments.

Set of disjoint events associated with C

$$\mathcal{C} = \{C_1,\, C_2,\, \cdots C_m\}$$

Set of disjoint events associated with D

$$\mathcal{D} = \{D_1,\, D_2,\, \cdots D_n\}$$

Array of Probabilities

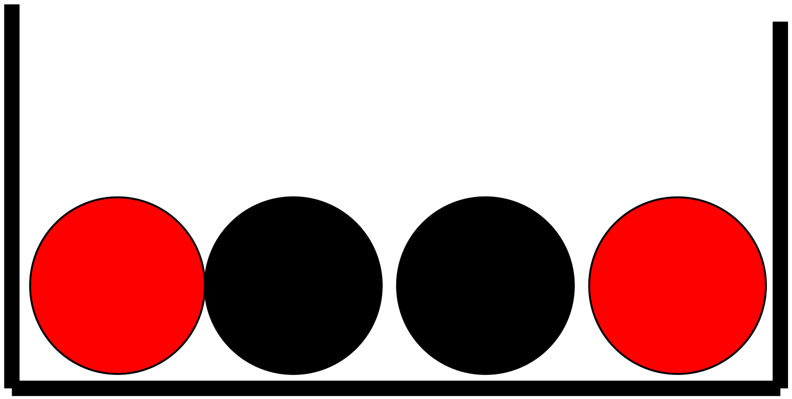
We can construct:

D / C	Event D_1	Event D_2	$\cdots$	Event D_n	Marginal Probabilities
Event C_1	$P(C_1 \cap D_1)$	$P(C_1 \cap D_2)$	$\cdots$	$P(C_1 \cap D_n)$	$P(C_1) = \sum_{i=1}^{n} P(C_1 \cap D_i)$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
Event C_m	$P(C_m \cap D_1)$	$P(C_m \cap D_1)$	$\cdots$	$P(C_m \cap D_n)$	$P(C_m) = \sum_{i=1}^{n} P(C_m \cap D_i)$
Marginal Probabilities	$P(D_1) = \sum_{i=1}^{m} P(C_i \cap D_1)$	$P(D_2) = \sum_{i=1}^{m} P(C_i \cap D_2)$	$\cdots$	$P(D_n) = \sum_{i=1}^{m} P(C_i \cap D_n)$	$Sum = 1$

Example:

There are 4 balls in one jar, 2 balls are red and two balls are black.

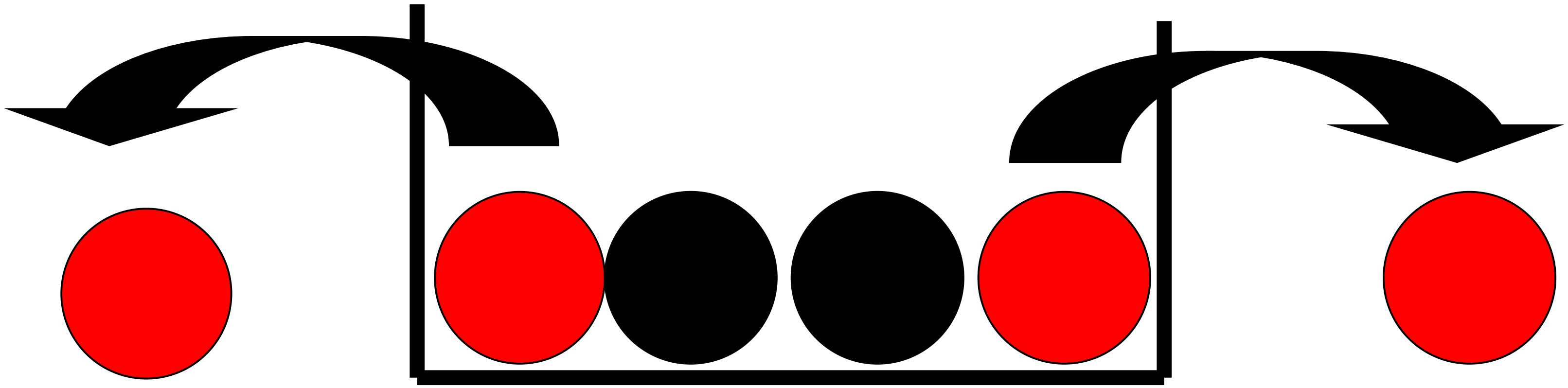
- A person can remove a ball from the jar two times, without seeing the balls inside the jar.



Example:

What is the probability of removing a red ball after having removing a red ball the first time?

To answer this question, lets build the table of probabilities.

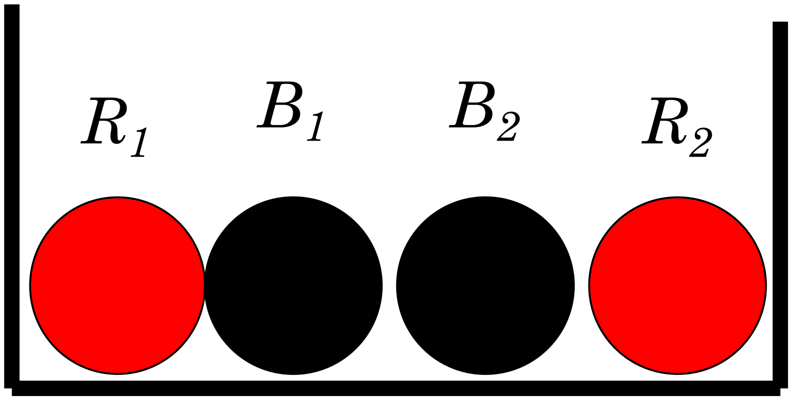


Example:

What is the probability of removing a red ball after having removing a red ball the first time?

To answer this question, lets build the table of probabilities.

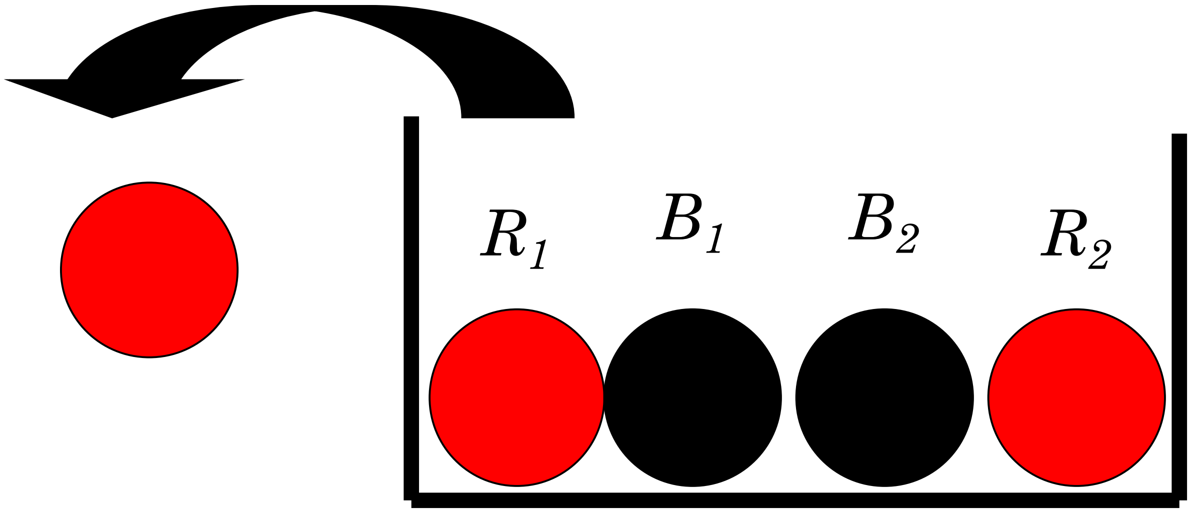
Labels:



Example:

Probability of picking R_1 the first time?

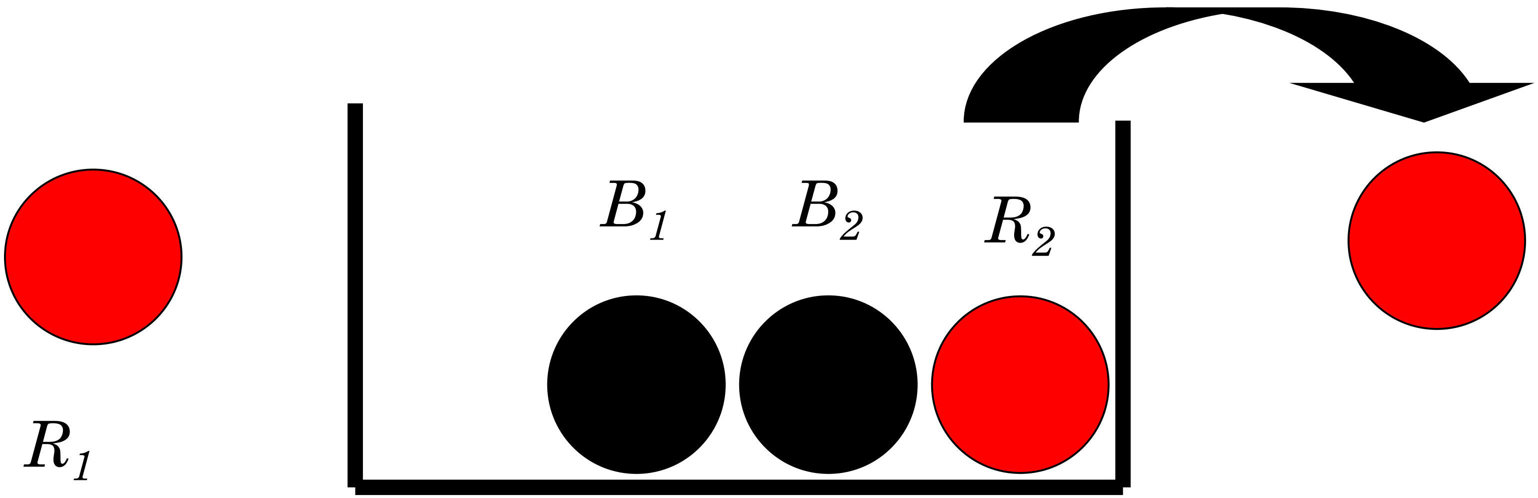
$$P(R_1) = 1/4$$



Example:

Probability of picking R_2 with only 3 balls left?

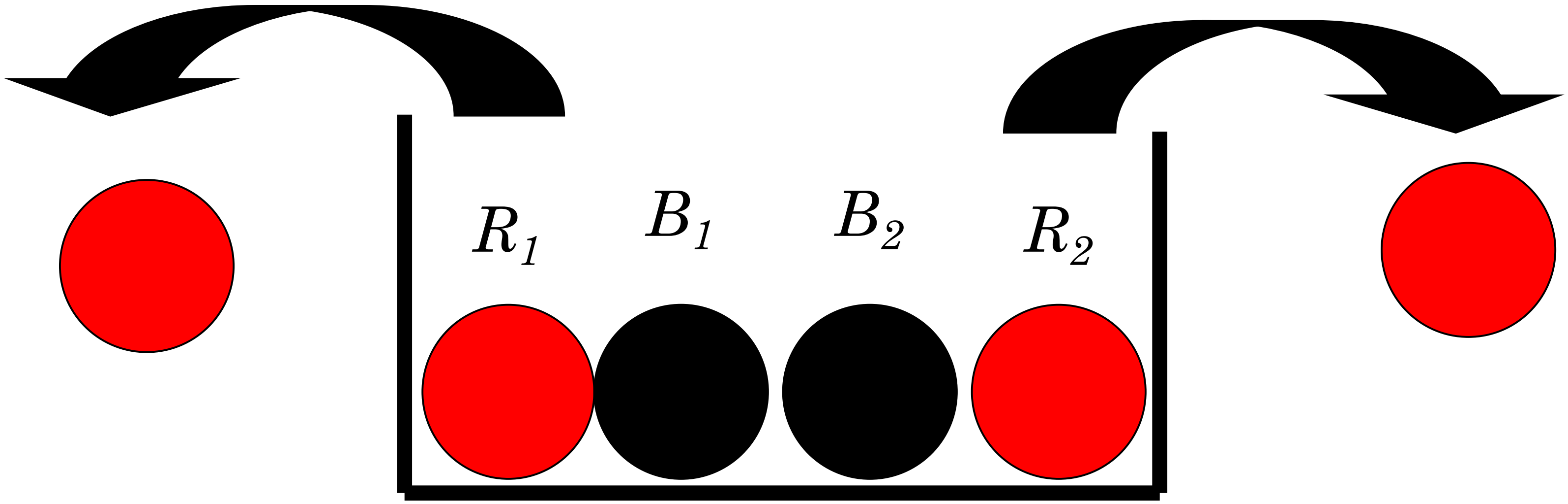
$$P(R_2) = 1/3$$ (second time)



Example:

Probability of picking R_1 the first time and R_2 the second time?

$$P(R_1 \cap R_2) = 1/4 \times 1/3 = 1/12$$



Example: Array of Probabilities

2 pick / *1 pick*	R_1	R_2	B_1	B_2	Marginal Probabilities
R_1	0	1/12	1/12	1/12	1/4
R_2	1/12	0	1/12	1/12	1/4
B_1	1/12	1/12	0	1/12	1/4
B_2	1/12	1/12	1/12	0	1/4
Marginal Probabilities	1/4	1/4	1/4	1/4	$Sum = 1$

Probability of pick red balls consecutively

Probability of event *A*: picking a red ball the first time and a red ball the second time?

- Event B: Picking R_1 first and R_2 second
- Event C: Picking R_2 first and R_1 second

Mutually exclusive *events*

$$P(A) = P(B \cup C)$$

$$= P(B) + P(C)$$

$$= \frac{1}{12} + \frac{1}{12} = \frac{1}{6}$$

Example: Array of Probabilities

$2\ pick$ / $1\ pick$	Red	$Black$	Marginal Probabilities
Red	$1/6$	$1/3$	1/2
$Black$	$1/3$	$1/6$	1/2
Marginal Probabilities	1/2	1/2	$Sum = 1$

Example:

What is the probability of picking a red ball the second time after having picked a red ball the first time?

$$P(Red_2|Red_1) = \frac{P(Red_2 \cap Red_1)}{P(Red_1)}$$

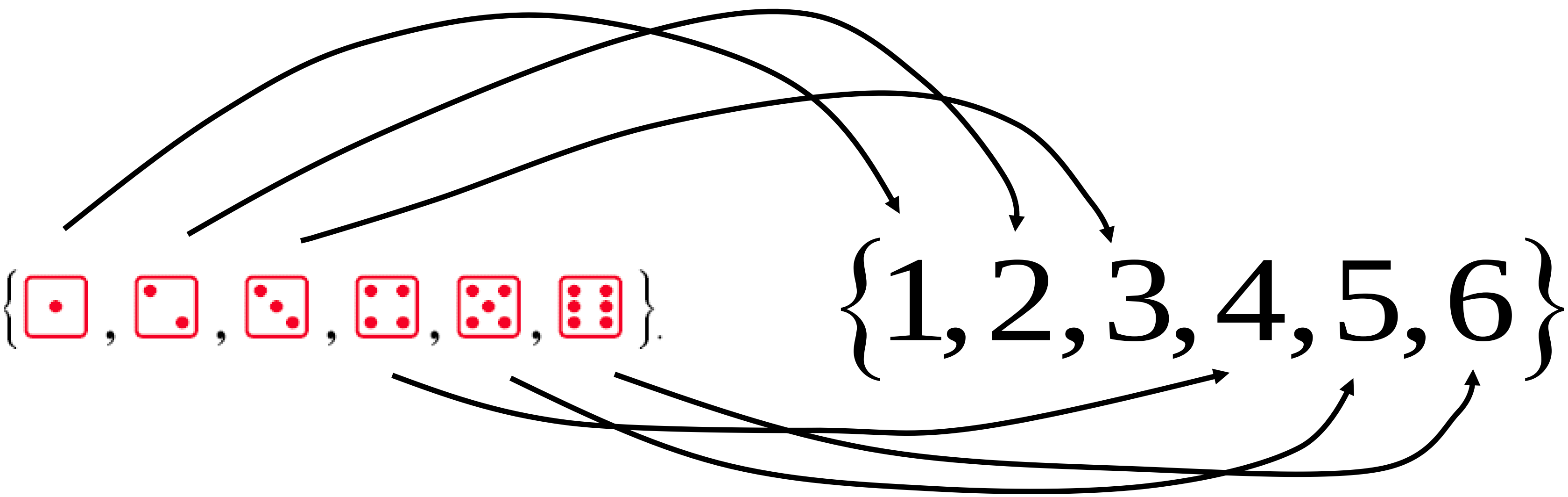
$$\boxed{P(Red_2|Red_1) = \frac{1/6}{1/2} = \frac{1}{3}}$$

Discrete random variable

Given a sample Space Ω, a random variable X is a function that assigns to each outcome a unique numerical value.

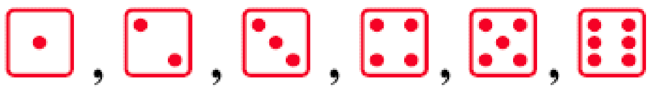
- Example: throwing of a die once

$\Omega =$ {⚀, ⚁, ⚂, ⚃, ⚄, ⚅}.



Discrete random variable

- <u>Example:</u> throwing of a die once

$\Omega = \{$  $\}$.

- In this case, the random variable X only takes discrete values

$$x_i \in \{1,2,3,4,5,6\}$$

- The discrete random variable X is defined by the **<u>probability mass function</u>**

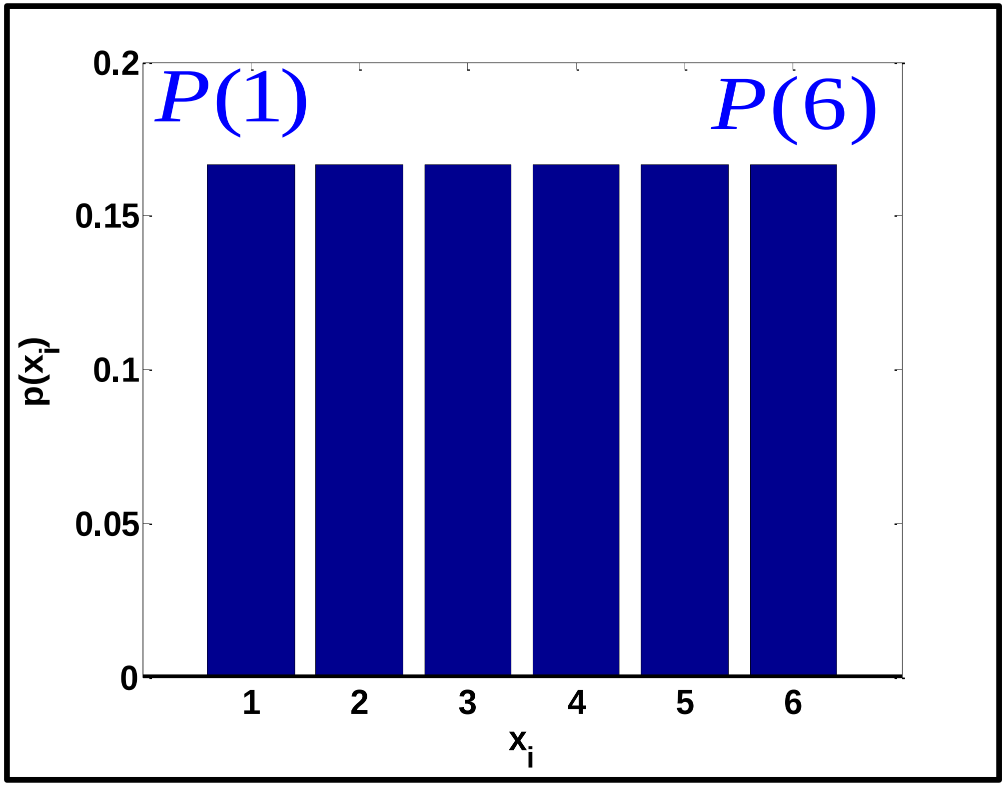
$$P(x_i) = P(X = x_i)$$

the probability that, after throwing a die, X will be equal to x_i

Discrete random variable

- For a **fair die,** the probability mass function of the random variable X is

$$P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = 1/6$$



the probability mass function satisfies:

$$\sum_{i=1}^{6} P(x_i) = 1$$

Expected value

- For a discrete random variable X taking on the N possible values

$$x_1, x_2, x_3, \dots, x_k, \dots, x_N$$

the **expected value** or **mean** of X is defined by

$$\boxed{E[X] = m_x = \hat{x} = \sum_{k=1}^{N} x_k P(x_k)}$$

$$E[X] = x_1 P(x_1) + x_2 P(x_2) + \cdots + x_N P(x_N)$$

Expected value of a function

- For a discrete random variable X taking on the N possible values

$$x_1, x_2, x_3, \ldots , x_k, \ldots, x_N$$

and the real-valued function f

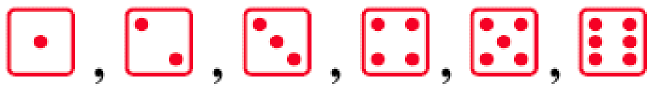
the **expected value** or **mean** of $Y=f(X)$ is defined by

$$\boxed{E[Y] = E[f(X)] = \sum_{k=1}^{N} f(x_k)P(x_k)}$$

$$E[Y] = f(x_1)P(x_1) + \cdots + f(x_N)P(x_N)$$

Expected value

Example: For a **fair dice,**

$\Omega = \{$  $\}$ $= \{1, 2, 3, 4, 5, 6\}$

- X takes 6 possible values $x_i = 1, 2, 3, 4, 5, 6$

- $P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = 1/6$

the **expected value** or **mean** of X

$$E(X) = m_x = \sum_{k=1}^{6} x_k P(x_k) = \frac{1}{6}\sum_{k=1}^{6} k = \frac{1}{6} 21 = 3.5$$

Variance and standard deviation

- For a discrete random variable X taking on the N possible values

  $x_1, x_2, x_3, \ldots, x_k, \ldots, x_N$ and a mean $m_X = \hat{x}$

the **variance** of X is defined by

$$\boxed{E[(X - m_X)^2] = \sigma_X{}^2 = \sum_{k=1}^{N} (x_k - m_X)^2 P(x_k)}$$

where $\boldsymbol{\sigma_X}$ is the standard deviation of X

Variance and standard deviation

Example: For a **fair dice,** where $x_i = 1, 2, 3, 4, 5, 6$

has mean $m_x = 3.5$ and $P(x_i) = 1/6$

the variance and standard deviation of X are

$$E[(x - m_x)^2] = \sum_{k=1}^{6} (x_k - 3.5)^2 P(x_k) = \frac{1}{6} \sum_{k=1}^{6} (k - 3.5)^2$$

$$= \frac{1}{6}\left[(1-3.5)^2 + (2-3.5)^2 + \ldots + (6-3.5)^2\right] = 2.9167$$

$$\sigma_x = \sqrt{E[(X - m_x)^2]} = \sqrt{2.9167} = 1.7078$$

Cumulative Distribution Function

- The cumulative distribution function (CDF) for a discrete random variable X is

$$F_X(x) = P(X \le x)$$

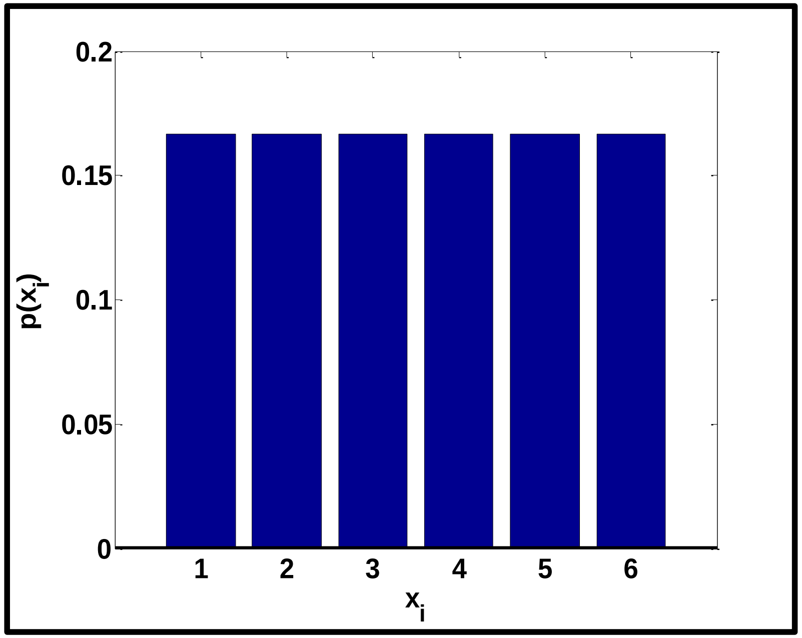
Find index $\boldsymbol{k}$ *such that* $$x_k \le x < x_{k+1}$$

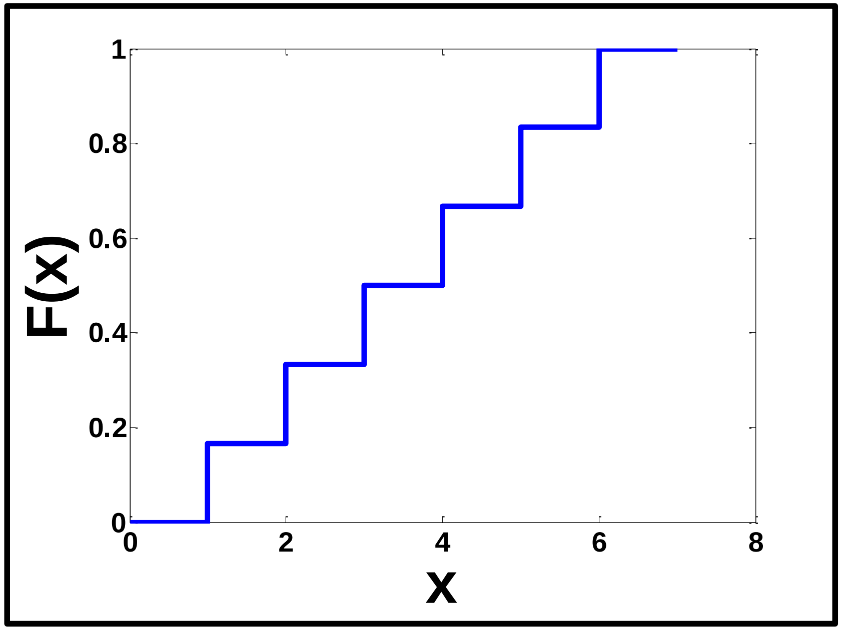
$$F_X(x) = \sum_{j=1}^{k} P(x_j)$$

Cumulative Distribution Function

- The <u>cumulative distribution function</u> (CDF) for a discrete random variable X is

$$F_X(x) = \sum_{j=1}^{k} P(x_j) \qquad x_k \leq x < x_{k+1}$$





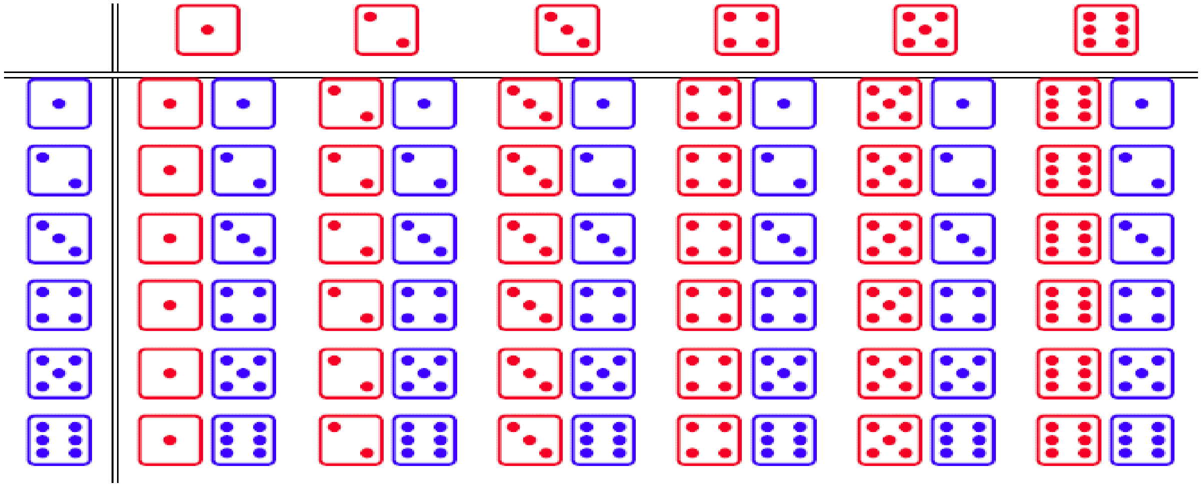
Sum of two uniform independent random variables

- Let X and Y be 2 independent random variables with constant probability mass function
- Let $Z = X + Y$
- The probability mass function of Z will not be constant

Throwing two fair dice

Experiment: throwing a pair of **<u>fair</u>** dice (red and blue)

$\Omega =$

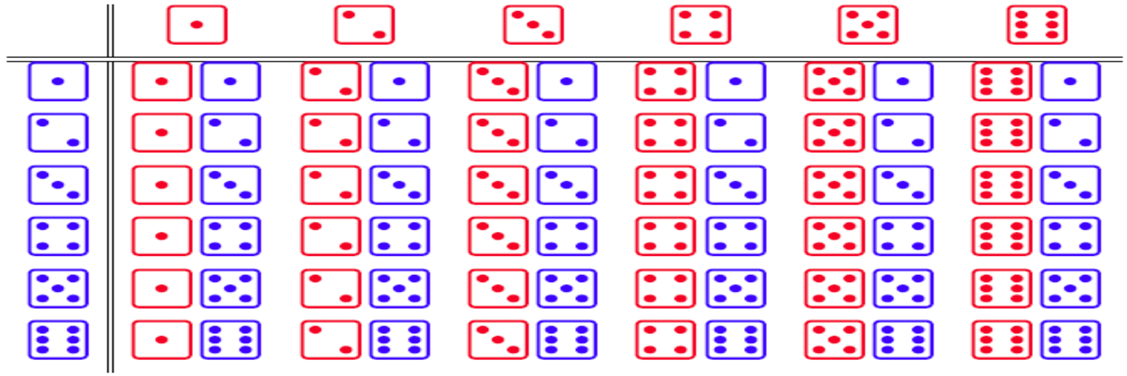


– the sample space has ***36*** outcomes:

– each outcome has a ***1/36*** probability of occurring

Throwing two fair dice

$\Omega =$

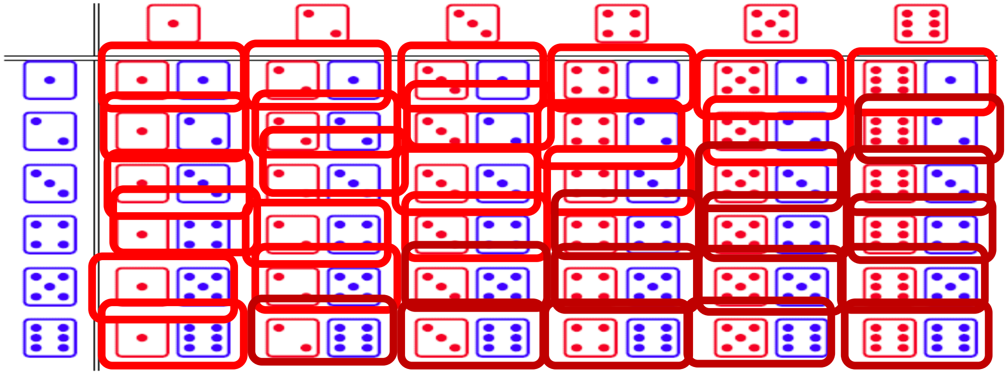


- Define the random variable Z associated with the **event** of observing the total number of dots on both dice after each throw

$Z = k$ when the throw results in the number k

Throwing two fair dice

$\Omega =$



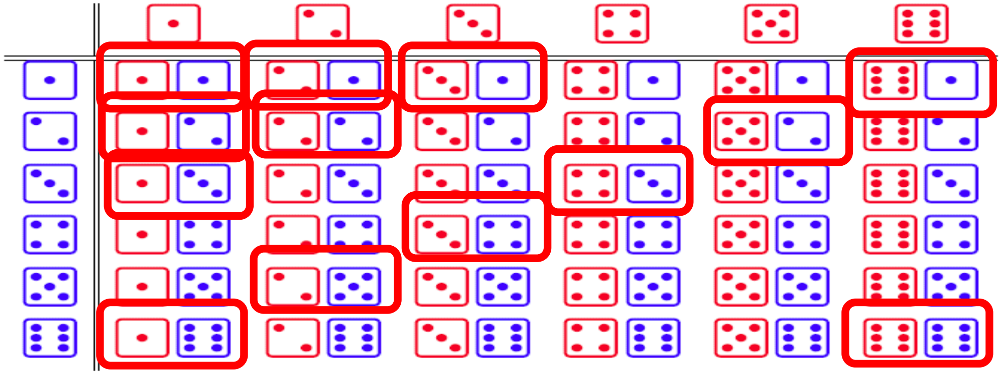
number of outcomes ***36***

Z only takes discrete values

$$z_i \in \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$$

Throwing two fair dice

$\Omega =$



probability of each outcome ***1/36***

we now estimate:

$$Z = 2 \rightarrow \quad P(2) = 1/36$$

$$Z = 3 \rightarrow \quad P(3) = 2/36$$

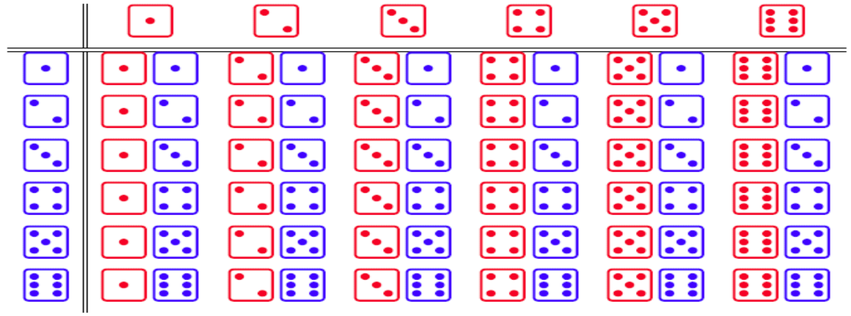
$$Z = 4 \rightarrow \quad P(4) = 3/36$$

$$Z = 7 \rightarrow \quad P(7) = 6/36$$

$$Z = 12 \rightarrow \quad P(12) = 1/36$$

Throwing two fair dice

$\Omega =$



The **probability mass function** is

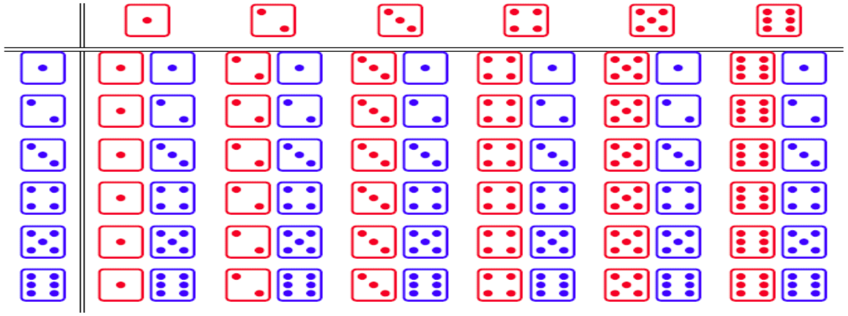
$P(2) = 1/36$ $P(5) = 4/36$ $P(8) = 5/36$ $P(11) = 2/36$

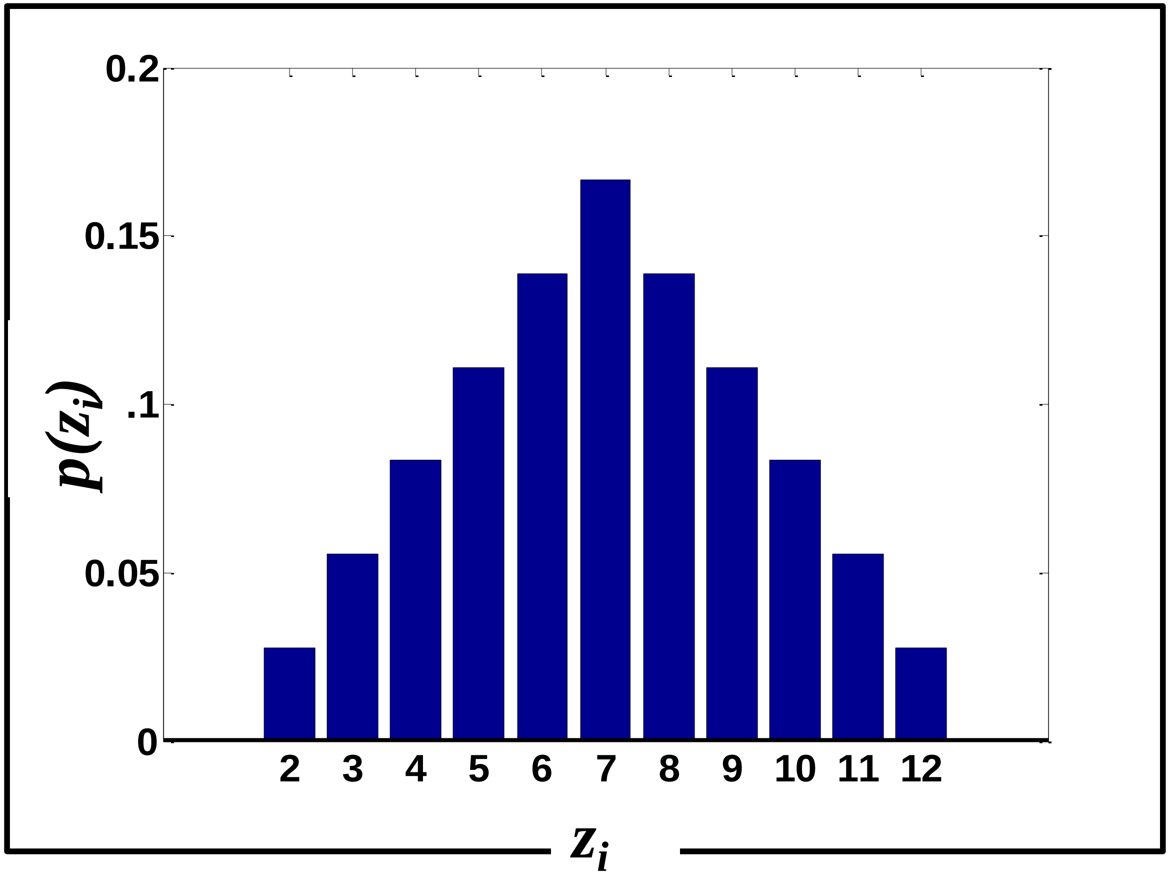
$P(3) = 2/36$ $P(6) = 5/36$ $P(9) = 4/36$ $P(12) = 1/36$

$P(4) = 3/36$ $P(7) = 6/36$ $P(10) = 3/36$

Throwing two fair dice

$\Omega =$





the probability mass function satisfies:

$$\sum_{k=2}^{12} P(k) = 1$$