Review of Probability

Liang Liang

Set

- A set is an unordered collection of distinct objects
- The objects in a set are called elements
- Examples:

$$Set1 = \{3, 5, 6, 7, 1\} = \{1, 3, 5, 6, 7\}$$

$$Set2 = \{a, b, c, d\} = \{c, b, a, d\}$$

Set

- Three basic set operations: **union**, **intersection**, **difference**

union: The union of sets A and B is denoted by $A \cup B$.

$$A \cup B = \{ x \mid x \textbf{ in } A \textbf{ or } x \textbf{ in } B \}$$

x is an element of the combined set

Example: If $A = \{1, 2, 3\}$ and $B = \{4, 5\}$, then $A \cup B = \{1, 2, 3, 4, 5\}$

Set

- Three basic set operations: **union**, **intersection**, **difference**

intersection: The intersection of two sets A and B is denoted by $A \cap B$.

$$A \cap B = \{\, x \mid x \textbf{ in } A \textbf{ and } x \textbf{ in } B \,\}$$

x is an element of the intersection

Example:

$A = \{1, 2, 3\}$ and $B = \{1, 2, 5\}$, then $A \cap B = \{1, 2\}$

Set

- Three basic set operations: **union**, **intersection**, **difference**

difference: The difference of two sets A and B is denoted by $A - B$ or $A \setminus B$

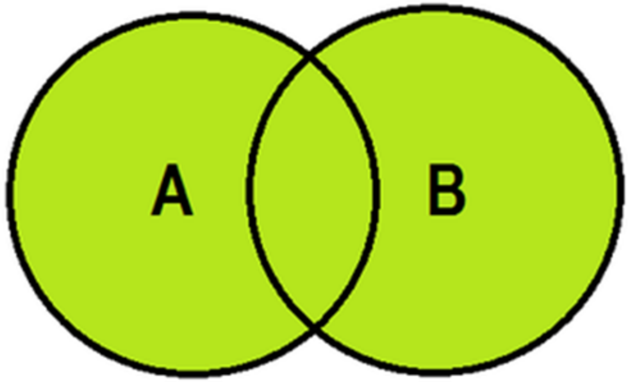
$$A - B = A \setminus B = \{ x \mid x \textbf{ in } A \textbf{ and } x \textbf{ not in } B \}$$

x is an element of the difference set

Example:

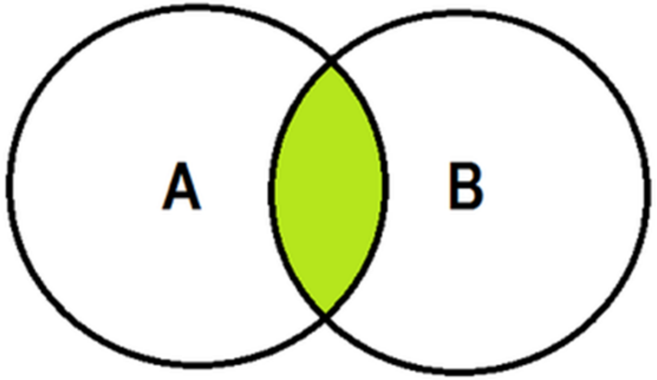
$A = \{1, 2, 3\}$ and $B = \{1, 2, 5\}$, then $A - B = \{3\}$.

$A \cup B = \{\, x \mid x$ **in** A **or** x **in** $B \,\}$



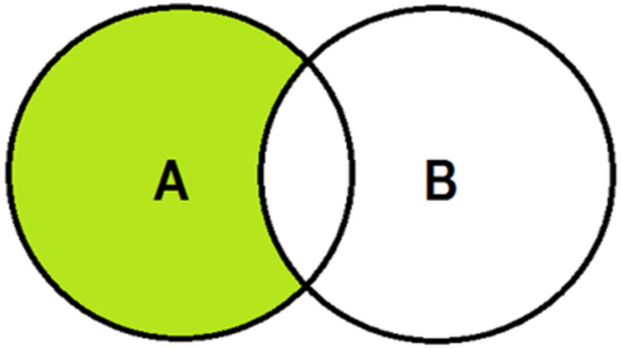
A union B
union of A and B

$A \cap B = \{\, x \mid x$ **in** A **and** x **in** $B \,\}$



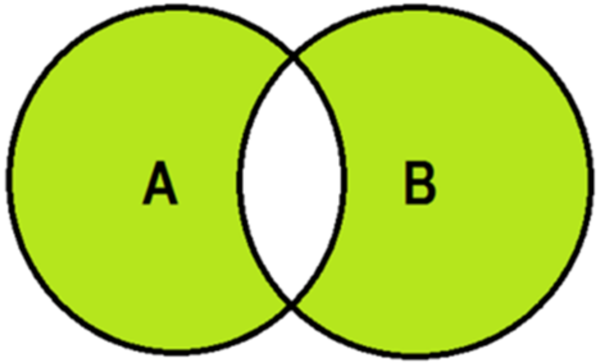
A intersect B
intersection of A and B

$A - B = A \setminus B = \{\, x \mid x$ **in** A **and** x **not in** $B \,\}$



A minus B
A complement B
difference of A and B

symmetric difference of A and B



$$A \Delta B = (A - B) \cup (B - A)$$

Set is a data type in Python

https://docs.python.org/3/tutorial/datastructures.html

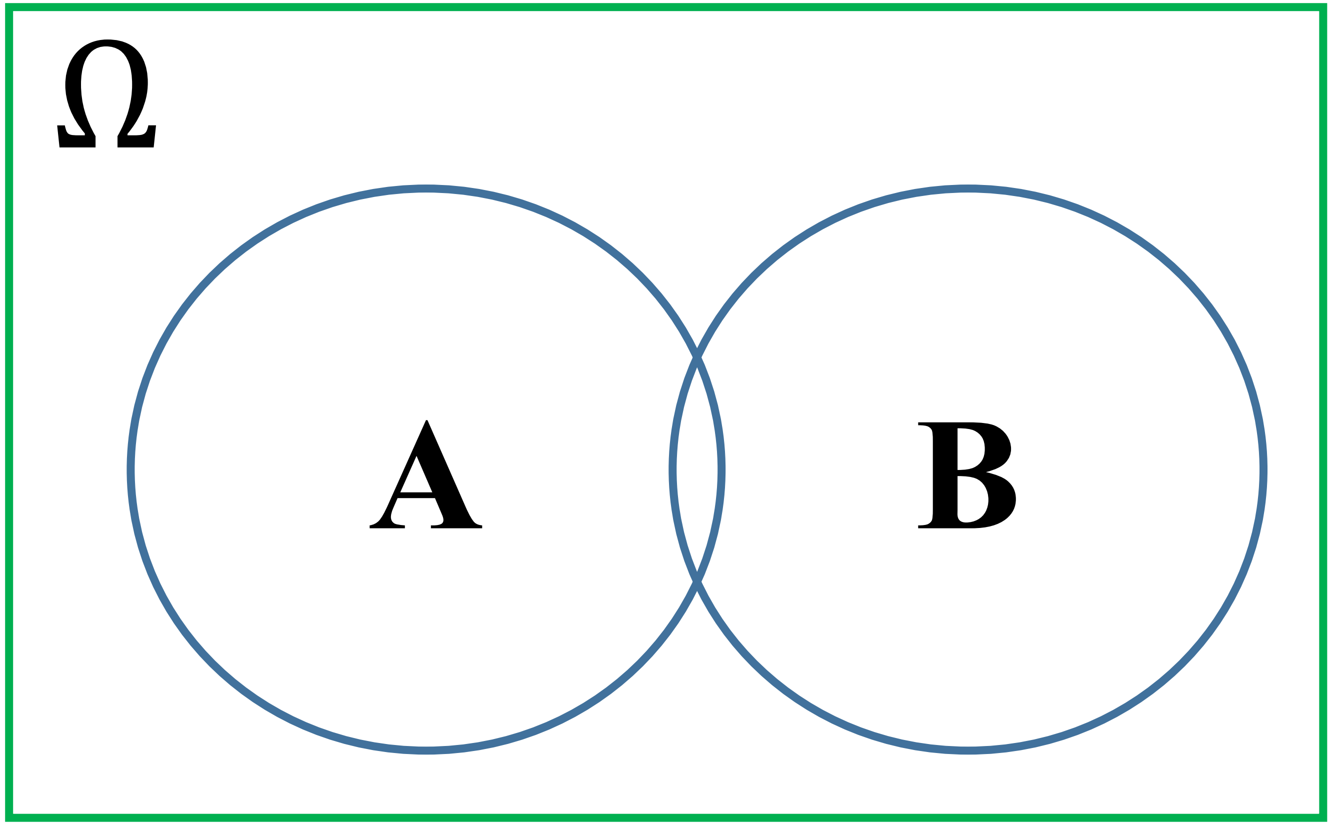
5.4. Sets

Python also includes a data type for *sets*. A set is an unordered collection with no duplicate elements. Basic uses include membership testing and eliminating duplicate entries. Set objects also support mathematical operations like union, intersection, difference, and symmetric difference.

Curly braces or the `set()` function can be used to create sets. Note: to create an empty set you have to use `set()`, not `{}`; the latter creates an empty dictionary, a data structure that we discuss in the next section.

Here is a brief demonstration:

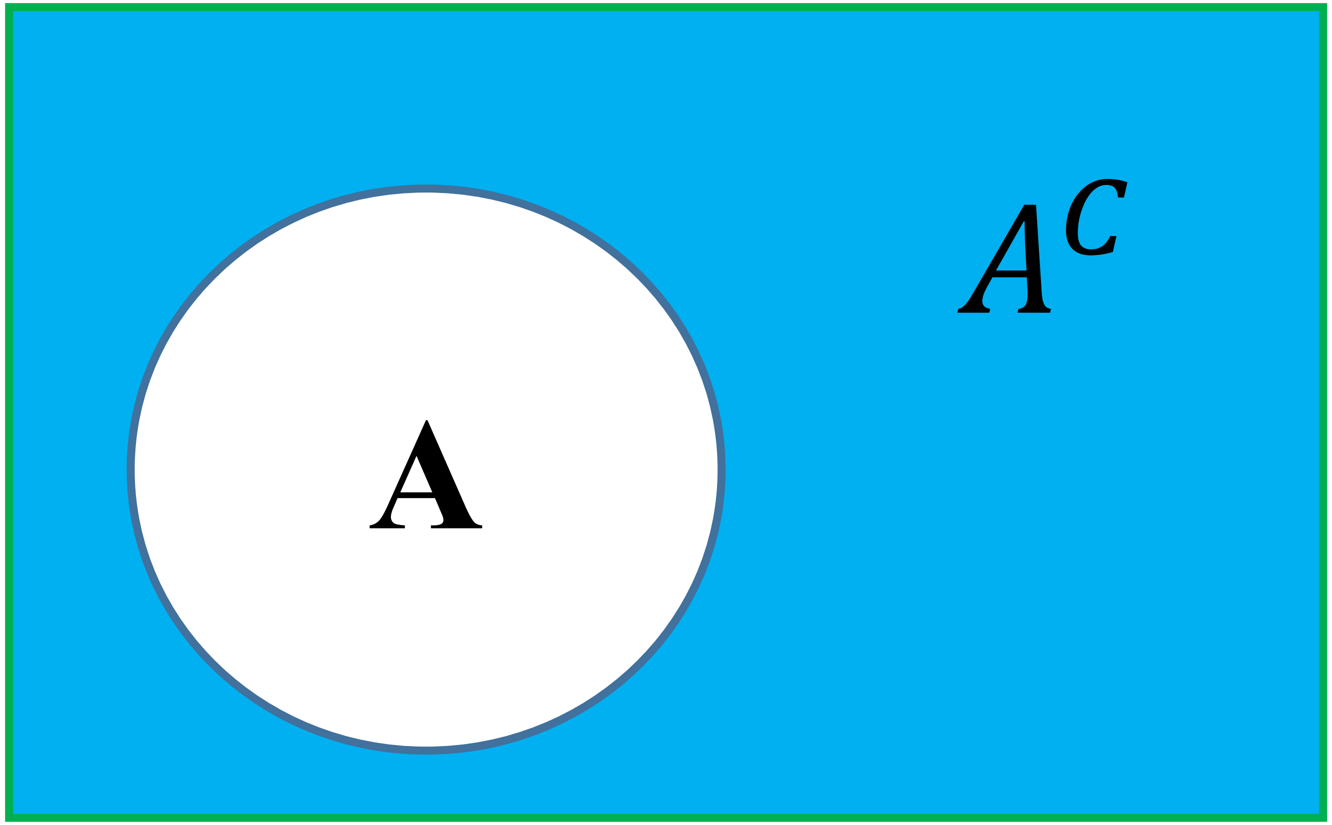
```
>>> basket = {'apple', 'orange', 'apple', 'pear', 'orange', 'banana'}
>>> print(basket)                      # show that duplicates have been removed
{'orange', 'banana', 'pear', 'apple'}
>>> 'orange' in basket                 # fast membership testing
True
>>> 'crabgrass' in basket
False
```



Ω : the whole space

$A \subseteq \Omega$: A is a subset of Ω

$B \subseteq \Omega$: B is a subset of Ω



Ω : the whole space

$A^C = \Omega - A$: the complement of A

$\bar{A}$ is another notation: $\bar{A} = A^C$

$\bar{\bar{A}} = A$

$\emptyset$ the Empty Set

nothing inside the empty set

$$\emptyset = \{\}$$

Ω denotes a sample space, it is a set

$\omega \in \Omega$: ω is a sample in Ω

$A \subseteq \Omega$: A is a subset of Ω

Example:

Ω is the set of real numbers

A is the set of real positive numbers

$A=\{a \mid a \in \Omega, a>0\}$

$A=\{\omega \mid \omega \in \Omega, \omega>0\}$

Elements of Probability

- Ω denotes a sample space, it is a set
- $\omega \in \Omega$: ω is a sample in Ω
- $A \subseteq \Omega$: A is a subset of Ω
- $\mathcal{F}$: a class/set of subsets of Ω , and it has the following properties:
  - $\emptyset \in \mathcal{F}$: the empty set $\emptyset$ belongs to $\mathcal{F}$
  - $\Omega \in \mathcal{F}$: the space Ω belongs to $\mathcal{F}$
  - if $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$
  - if $A_1, A_2, A_3 \ldots$ is a countable collection of sets in $\mathcal{F}$ then both the union $\cup_i A_i$ and the intersection $\cap_i A_i$ are also in $\mathcal{F}$

    $\cup_i A_i = A_1 \cup A_2 \cup A_3 \cup \cdots$

    $\cap_i A_i = A_1 \cap A_2 \cap A_3 \cap \cdots$

Elements of Probability

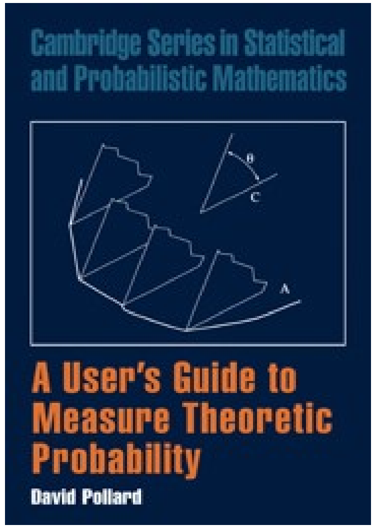
- **sample space** Ω, sample $\omega \in \Omega$,
- **sigma-field** $\mathcal{F}$, set $A \in \mathcal{F}$ and set $A \subseteq \Omega$
- **probability measure** is a function $P : \mathcal{F} \rightarrow \mathcal{R}$ and has the following properties:
  - $P(A) \geq 0$ for all $A \in \mathcal{F}$
  - $P(\Omega) = 1$ and $P(\emptyset) = 0$
  - if $A_1, A_2, A_3 \ldots$ is a countable collection of pairwise disjoint sets in $\mathcal{F}$, then
    $P(\cup_i A_i) = \sum_i P(A_i)$, pairwise disjoint $A_i \cap A_j = \emptyset,\ for\ i \neq j$
    for example, $P(A_1 \cup A_2) = P(A_1) + P(A_2)$, disjoint $A_1 \cap A_2 = \emptyset$

Elements of Probability

- **probability measure** is a function $P : \mathcal{F} \rightarrow \mathcal{R}$
- More properties:
- if $A \subseteq B$, then $\mathrm{P}(A) \leq P(B)$
- $\mathrm{P}(A \cap B) \leq \min(\mathrm{P}(A), P(B))$, for any $A, B \in \mathcal{F}$
- $\mathrm{P}(A \cup B) = \mathrm{P}(A) + P(B) - \mathrm{P}(A \cap B) \leq \mathrm{P}(A) + P(B)$, for any $A, B \in \mathcal{F}$
- $\mathrm{P}(\Omega - A) = 1 - \mathrm{P}(A)$
- if $A_1, A_2, A_3 \ldots$ is a countable collection of pairwise disjoint sets in $\mathcal{F}$, such that $\cup_i A_i = \Omega$, then $\sum_i P(A_i) = 1$

  for example, $A \cup A^C = \Omega$, then $P(A) + P\left(A^C\right) = 1$

- Probability and Measure theory
  - $\mathcal{F}$ is called a sigma-field
  - set $A \in \mathcal{F}$ and $A \subseteq \Omega$



- Elementary Probability Theory
  - $\mathcal{F}$ is called an event space
  - A is an event in $\mathcal{F}$, $A \in \mathcal{F}$

- Since we are not theoretical mathematicians, let's do this:

  $\mathcal{F}$ is a sigma-field and an event space

  A is a set and an event in $\mathcal{F}$, $A \in \mathcal{F}$

Conditional Probability And Independence

Let B be an event with non-zero probability, $P(B) > 0$.
The **conditional probability** of any event A given B is defined as,

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

In other words, $P(A|B)$ is the probability measure of the event A after observing the occurrence of event B.

Two events are called **independent** if and only if

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

which is equivalent to

$$P(A|B) = P(A)$$

Mutual Independence

Two events are called **independent** if and only if

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

Mutual Independence of three events

Three events A, B and C are be mutually independent if and only if the following two conditions are met:

(1) $P(A \cap B \cap C) = P(A)P(B)P(C)$

(2) A, B and C are pairwise independent

A and B are independent

A and C are independent

B and C are independent

Random Variable

A random variable X is function: $\Omega \to \mathcal{R}$,

$\omega \in \Omega$, $X(\omega)$ is a random variable

$X(\omega)$ is a mapping from a sample ω to a number

continuous random variable: $X(\omega)$ is a continues function

discrete random variable: $X(\omega)$ takes discrete values

Cumulative Distribution Function (CDF)

A random variable X is function: $\Omega \to \mathcal{R}$,
$\omega \in \Omega$, $X(\omega)$ is a random variable

The **cumulative distribution function**, $F_X(x)$, is defined by

$$F_X(x) = P(X \leq x) = P(\{\omega \mid \omega \in \Omega, X(\omega) \leq x\})$$

x is a number, e.g., $F_X(100)$

set $A = \{\omega \mid \omega \in \Omega, X(\omega) \leq x\})$

$A \in \mathcal{F}$ and $A \subseteq \Omega$

Example: coin toss

- You toss a coin **once** :
  you get 10$ if you get a head
  you lose 1$ if you get a tail

Head Tail

- sample space $\Omega = \{H, T\}$
- event space $\mathcal{F} = \{\emptyset, \Omega, \{H\}, \{T\}\}$
- Ω is set of all outcomes of a random experiment
- The money you get/lose is a discrete random variable $\mathrm{X}(\omega), \omega \in \Omega$

$\mathrm{X}(H) = 10, \ \mathrm{X}(T) = -1$

- What is the probability that you get some money ?

$P(\mathrm{X}>0)=\mathrm{P}(\{\omega \mid \omega \in \Omega, \mathrm{X}(\omega)>0\})$

$\mathrm{A}=\{\omega \mid \omega \in \Omega, \mathrm{X}(\omega)>0\}=\{\mathrm{H}\} \in \mathcal{F}$

Example: roll a dice

- You roll a dice **once**:
  you get **x** \$ if you get a number **x** on the upper face
- sample space $\Omega = \{1, 2, 3, 4, 5, 6\}$
- event space $\mathcal{F} = \{\emptyset, \Omega, \{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}, \{1,2\}, \{1,2,3\}, \ldots\}$
- The money you get is a discrete random variable $X(\omega), \omega \in \Omega$

$$X(\omega) = \omega$$

- What is the probability that you get more than 3\$?

$$P(X>3)=P(\{\omega \mid \omega \in \Omega, X(\omega)>3\})$$

$$A=\{\omega \mid \omega \in \Omega, X(\omega)>3\}=\{4,5,6\} \in \mathcal{F}$$

Example: roll a dice

- The term "event" is the source of confusion
- event space $\mathcal{F} = \{\emptyset, \Omega, \{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}, \{1,2\}, \{1,2,3\}, \ldots\}$
  - A_1=$\{1\}$, the event is "get the number 1 after rolling the dice once"
  - A_4=$\{4\}$, the event is "get the number 4 after rolling the dice once"
  - A_5=$\{5\}$, the event is "get the number 5 after rolling the dice once"
  - A_6=$\{6\}$, the event is "get the number 6 after rolling the dice once"
  - A_{456}=$\{4,5,6\}$, the event is "get the number 4 or 5 or 6 after rolling the dice once"

  $A_{456} = A_4 \cup A_5 \cup A_6$
- Assuming it is a fair dice, then

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

- $P(A_{456})=P(A_4 \cup A_5 \cup A_6)=P(A_4)+P(A_5)+P(A_6)=1/2$

  because the sets A_4, A_5, A_6 are pairwise disjoint.

Example: roll a dice

- The term "event" is the source of confusion
- event space $\mathcal{F} = \{\emptyset, \Omega, \{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}, \{1,2\}, \{1,2,3\}, \ldots\}$
  - A={1,2,3,4,5,6}, the event is "get some money after rolling the dice once" and "get the number 1 or 2 or 3 or 4 or 5 or 6 after rolling the dice once"
- Assuming it is a fair dice, then

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

- Then, $P(A)=P(\{1\} \cup \{2\} \cup \{3\} \cup \{4\} \cup \{5\} \cup \{6\})=\sum_{n=1}^{6} P(\{n\}) = 1$ which means you will surely get some money after rolling the dice once

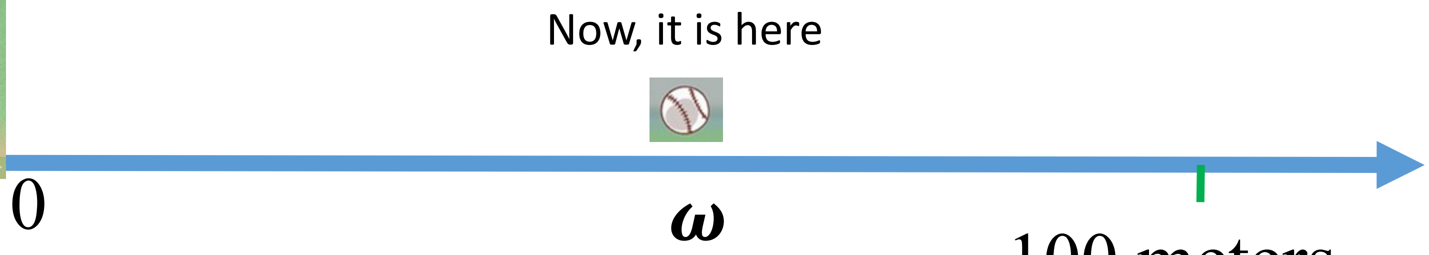
Example: throw a ball

https://www.wikihow.fitness/Throw-a-Ball-Faster

You throw a ball randomly

It lands on the ground $\boldsymbol{\omega}$ meters away

You get money $10\boldsymbol{\omega}$ \$, assume $\boldsymbol{\omega} \leq 100$

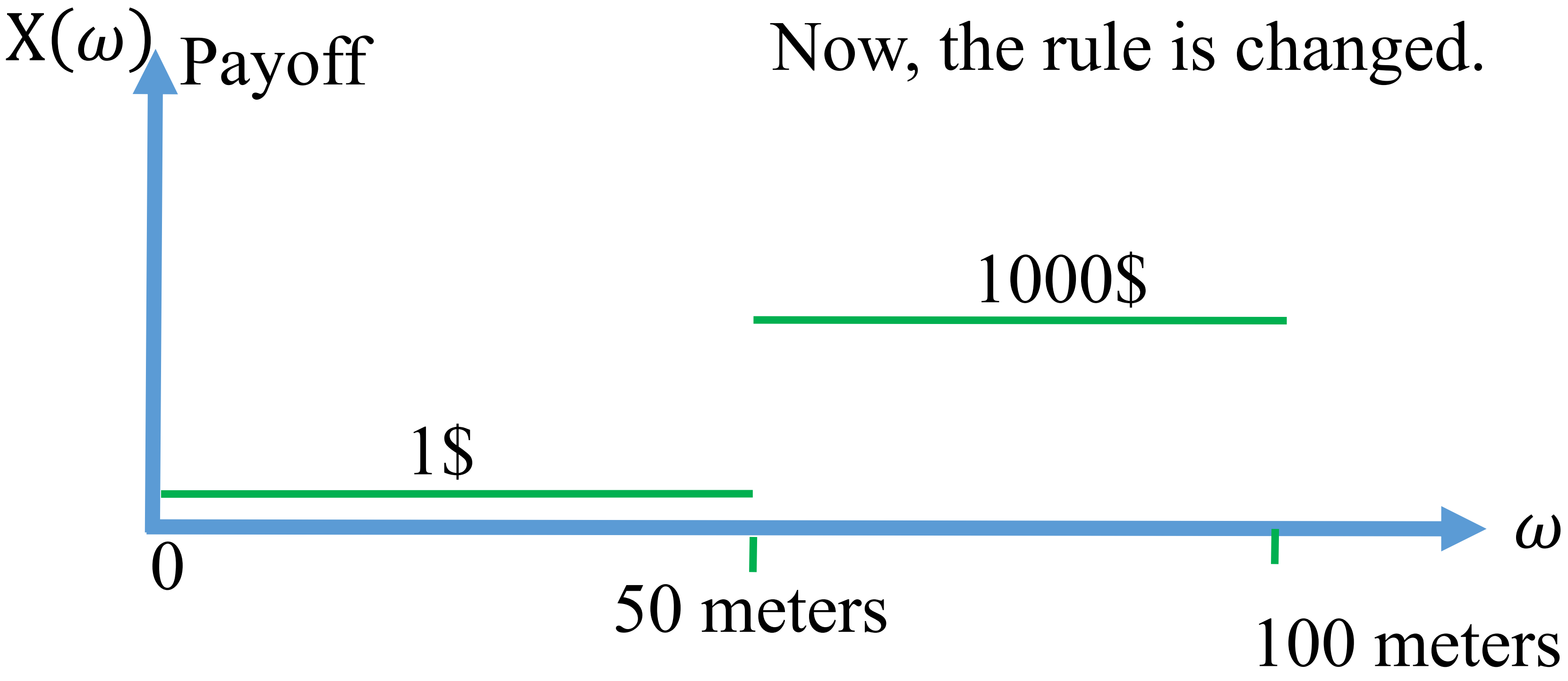


100 meters

- sample space $\Omega = [0, 100]$, a continuous interval
- event space $\mathcal{F} = \{\emptyset, \Omega, [0,1], [1,2], \dots\}$
- Your payoff is a continuous random variable $X(\omega) = 10\omega, \omega \in \Omega$
- What is the probability that you get more than 900\$?

$P(X>900)=P(\{\omega \mid \omega \in \Omega, X(\omega)>900\})=P(\{\omega \mid \omega \in \Omega, \omega>90\})$

Example: throw a ball



- sample space $\Omega = [0, 100]$, a **continuous** interval
- event space $\mathcal{F} = \{\emptyset, \Omega, [0,1], [1,2], \dots\}$
- Your payoff is a **discrete** random variable

  $\mathrm{X}(\omega) = 1$ if $0 \leq \omega < 50$; $\mathrm{X}(\omega) = 1000$ if $\omega > 100$

Example: coin toss - many times

- You toss a coin **once** :
  you get 10$ if you get a head
  you lose 1$ if you get a tail
- **Now, you toss the coin 2 times**
- What is the probability that you get more than 10$?
- What is the sample space ?
- **Solution-1:**

sample space $\Omega = \{HH, HT, TH, TT\}$, $\mathrm{X}(\omega)$, $\omega \in \Omega$

$\mathrm{X}(HH) = 10 + 10 = 20,\quad \mathrm{X}(HT) = 10 - 1 = 9$

$\mathrm{X}(TH) = -1 + 10 = 9,\quad \mathrm{X}(TT) = -1 - 1 = -2$

$P(\mathrm{X}>10)=\mathrm{P}(\{\omega \mid \omega \in \Omega, \mathrm{X}(\omega)>10\})$,

$\mathrm{A}=\{\omega \mid \omega \in \Omega, \mathrm{X}(\omega)>10\}=\{HH\} \in \mathcal{F}$

$P(\mathrm{X}>10)=\mathrm{P}(\{HH\})=0.25$

Example: coin toss - many times

- You toss a coin once : get 10$ if head; lose 1$ if tail
- **Now, you toss the coin 100 times**
- **Solution-2:**

sample space $\Omega = \{H, T\}$, $X_n(\omega)$, $\omega \in \Omega$
$X_n(H) = 10,\ X_n(T) = -1$
toss the coin, the payoff is X_1
toss the coin again, the payoff is X_2
...
You have a sequence of payoffs, $X_1, X_2, \ldots, X_n, \ldots X_{100}$
The total payoff after tossing the coin 100 times: $Y = \sum_{n=1}^{100} X_n$

The probability that you get more than 10$ is
$P(Y>10)=\mathrm{P}(\{\omega \mid \omega \in \Omega, \sum_{n=1}^{100} X_n(\omega)>10\})$
Then, you can write a Python program to calculate the result

The distribution function of X

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

$$P(X > x) = 1 - F_X(x)$$

Joint Distribution Function of Random Variables

If $X_1, X_2, \ldots, X_N$ are random variables, their joint distribution function is defined by

$$F_{X_1,\, X_2,\, \ldots,\, X_n}(x_1, x_2, \ldots, x_N)$$
$$= P(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_N \leq x_N)$$

The random variables are independent if for all $x_1, x_2, \ldots, x_N$

$$F_{X_1,\, X_2,\, \ldots,\, X_N}(x_1, x_2, \ldots, x_N) = \prod_{n=1}^{N} F_{X_n}(x_n)$$

where $F_{X_n}(x_n) = P(X_n \leq x_n)$

Probability Mass Function (PMF) of Discrete Random Variables

- Probability Mass Function $f_X(x) = P(X = x)$
- another the notation $p_X(x) = P(X = x)$
- Examples

tossing a coin, 0~head, 1~tail

$$p_X(0) = 0.5, p_X(1) = 0.5$$

Bernoulli distribution (tossing a rigged coin)

$$p_X(0) = p, p_X(1) = 1 - p,\ \text{e.g.,}\ p = 0.1$$

Binomial Distribution

(the probability of getting k heads after tossing a rigged coin n times)

$$p_X(k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

Probability Mass Function (PMF) of Discrete Random Variables

- Joint Probability Mass Function of random variables, $X_1, X_2, \ldots, X_N$ is defined by

$$p_{X_1,\, X_2,\, \ldots,\, X_n}(x_1, x_2, \ldots, x_N)$$
$$= P(X_1 = x_1, X_2 = x_2, \ldots, X_N = x_N)$$

Probability Density Function (PDF) of Continuous Random Variables

- X is a continuous random variable
- $F_X(x)$ is the cumulative distribution function (CDF) $F_X(x) = P(X \le x)$
- The probability density function, $f_X(x)$ or $p_X(x)$, (if it exists), is defined by

$$f_X(x) = \frac{\partial F_X(x)}{\partial x} = \lim_{\Delta \to 0} \frac{P(x \le X \le x + \Delta)}{\Delta}$$

- Thus, $F_X(x) = \int_{-\infty}^{x} f_X(x)dx$
- Properties of PDF:

$$f_X(x) \ge 0$$

$$\int_{-\infty}^{\infty} f_X(x)dx = 1$$

Probability Density Function (PDF) of Continuous Random Variables

- Joint Probability Density Function of random variables, $X_1, X_2, \ldots, X_N$, (if it exits) is defined by

$$f_{X_1,\ldots,X_N}(x_1,\ldots,x_N) = \frac{\partial^N F_{X_1,\ldots,X_N}(x_1,\ldots,x_N)}{\partial x_1 \ldots \partial x_N}$$

- $F_{X_1,\, X_2,\, \ldots,\, X_n}(x_1, x_2, \ldots, x_N) = \int_{-\infty}^{x_1} \ldots \int_{-\infty}^{x_N} f_{X_1,\ldots,X_N}(x_1,\ldots,x_N) dx_1 \ldots dx_N$
- Properties of PDF:

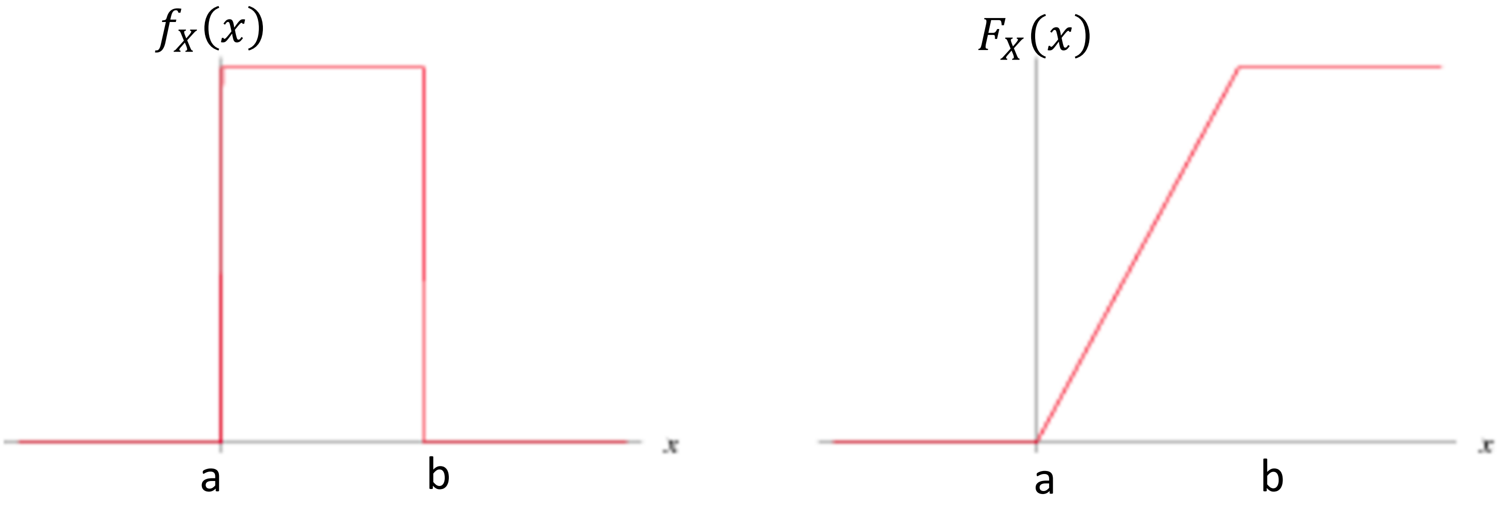
$$f_{X_1,\ldots,X_N}(x_1,\ldots,x_N) \geq 0$$

$$\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} f_{X_1,\ldots,X_N}(x_1,\ldots,x_N) dx_1 \ldots dx_N = 1$$

Probability Density Function (PDF) of Continuous Random Variables

- Examples:
  - Uniform distribution

$$f_X(x) = \begin{cases} \dfrac{1}{b-a} & for\ a \le x \le b \\ 0 & otherwise \end{cases}$$



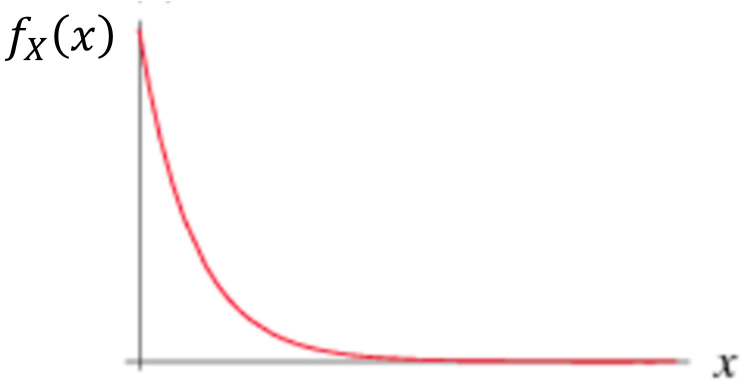
http://mathworld.wolfram.com/ExponentialDistribution.html

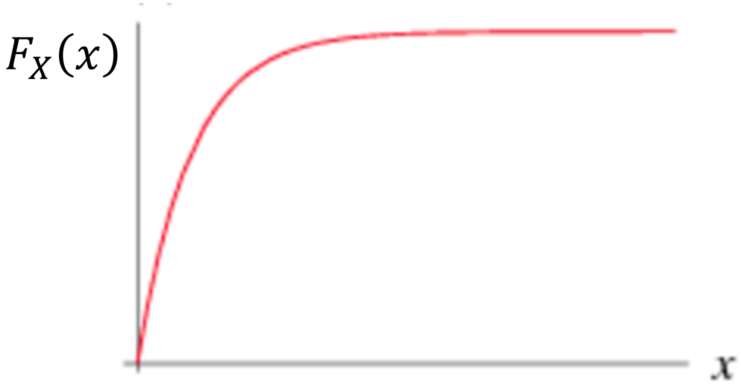
Probability Density Function (PDF) of Continuous Random Variables

- Examples:
  - Exponential distribution

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & for\ x \geq 0 \\ 0 & otherwise \end{cases}$$

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x} & for\ x \geq 0 \\ 0 & otherwise \end{cases}$$



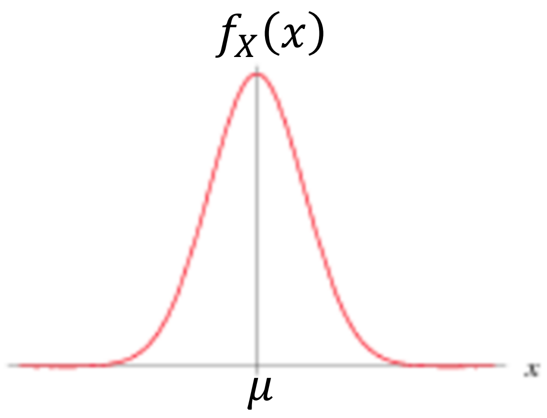


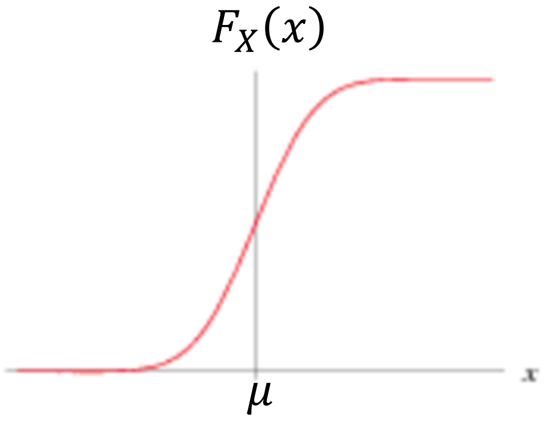
http://mathworld.wolfram.com/ExponentialDistribution.html

Probability Density Function (PDF) of Continuous Random Variables

- Examples:
  - Gaussian/Normal distribution

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$





http://mathworld.wolfram.com/ExponentialDistribution.html

Random Variable, CDF, PDF in Python

- Generate random numbers:

  https://docs.scipy.org/doc/numpy-1.15.1/reference/generated/numpy.random.rand.html

  https://docs.scipy.org/doc/numpy-1.15.1/reference/generated/numpy.random.randn.html

  https://docs.scipy.org/doc/numpy-1.15.1/reference/generated/numpy.random.randint.html

- Compute PDF/CDF:

  https://docs.scipy.org/doc/scipy/reference/tutorial/stats.html