Chapter 3. Probability (A)

January 19, 2011

1 Sample space and events

A random experiment is an action or process that leads to one of several possible outcomes.

For example: tossing a coin and the up face is observed, or drawing a card the number is recorded.

More examples

Experiment	Outcomes
Flip a coin	Heads, Tails
Exam Marks	Numbers: 0, 1, 2, ..., 100
Assembly Time	t > 0 seconds
Course Grades	F, D, C, B, A, A+

The collection of ALL possible outcomes is called the Sample Space, denoted by Ω or S. Each outcome is called a simple Events denoted by $\{w_1\}, \{w_2\}, \ldots$ or an element denoted by $w_1, w_2, \ldots$.

The following requirements are imposed on the sample space

- The sample space must be exhaustive, i.e. It includes ALL possible outcomes. e.g. roll a die Ω ={1,2,3,4,5,6} $\checkmark$, $\Omega =$ {1,2,3,4,5} $\times$
- The sample space must be mutually exclusive, i.e. no two outcomes can occur at the same time.

Example A roll a die twice and their sum is recorded, the sample space

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

The difference between the sample space and population by considering the following examples

Example B Suppose in a class, there are 20 students with heights 1.692, 1.79, 1.718, 1.729, 1.711, 1.744, 1.71, 1.979, 1.583, 1.515, 1.586, 1.591, 1.657, 1.683, 1.678, 1.754, 1.739, 1.775, 1.878, 1.822, which is the population. An experiment is to select one person and measure his/her height. In this example

$$\text{sample space } \Omega = \text{Population.}$$

Example C For the village in Chapter 1, the experiment is to measure a person's weight,

$$\Omega = \{75.6, 75.7, 75.8, 75.9, 76.0, 76.1, 76.2, 76.3, 76.4\}$$

In this example, population $\neq$ sample space!

However, their distributions are the same (to be discussed later)

In population Jargon *In experiment Jargon*	value outcome	frequency	relative frquency p_i probability p_i
	75.6	6	0.075
	75.7	8	0.100
	75.8	8	0.100
	75.9	14	0.175
	76.0	18	0.225
	76.1	11	0.1375
	76.2	8	0.100
	76.3	5	0.0625
	76.4	2	0.025
	total	80	1.00

Conclusion:

- sample space together with its probability (to be discussed later) is equivalent to the population
- To investigate the population, one can turn to investigate the random experiment together with probability! Or later a random variable with probability function.

Sample space can be classified to

- (according to its number of elements). Finite sample space and infinite sample space. tossing a coin has finite sample space; measuring a robe has infinite sample space.

- (according to variable's type). Discrete sample space and continuous sample space: the former has finitely many or a countable infinity of elements; if the elements (points) of a sample space constitute a continuum, the space is called continuous sample space. For example, all the points on a line segment, or all the points in a plane

Events. Any subset of Ω is called an event, denoted by $A, B, \ldots.$ Thus, $A \subset \Omega, B \subset \Omega, \ldots.$

3 special events

- whole set Ω, which includes all elements
- empty set, denoted by $\emptyset$ or ϕ, which has no elements at all.
- simple event, which include only one element, denoted by $E_1, E_2, \ldots$

Example Roll of a die: $\Omega = \{1, 2, 3, 4, 5, 6\}$. An even number (one of 2, 4, or 6) will be rolled is an event

$$A = \{2, 4, 6\}$$

Example Roll of a die twice and their numbers are recorded

$$\begin{aligned}\Omega \;=\; \{&(1,1),\; (1,2),\; (1,3),\; (1,4),\; (1,5),\; (1,6),\\ &(2,1)\;\; (2,2),\; (2,3),\; (2,4),\; (2,5),\; (2,6),\\ &(3,1)\;\; (3,2)\;\; (3,3),\; (3,4),\; (3,5),\; (3,6),\\ &(4,1)\;\; (4,2)\;\; (4,3)\;\; (4,4),\; (4,5),\; (4,6),\\ &(5,1)\;\; (5,2)\;\; (5,3)\;\; (5,4)\;\; (5,5),\; (5,6),\\ &(6,1)\;\; (6,2)\;\; (6,3)\;\; (6,4)\;\; (6,5)\;\; (6,6)\}\end{aligned}$$

The event of “rolling a four” would be

$$A = \{(1,3),(2,2),(3,1)\}$$

The event of “first role is 2” is

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

Example Observe daily traffic accidents in a city, then the sample space is $\Omega = \{0, 1, 2, ...\}$. Then

$$A = \{10, 11, ...\}$$

is an event.

Example Observing a person's height in a city.

$$A = \text{the observed height is above 1.8m}$$

is an event.

Sometimes, it is not easy to convert a *daily life event* to the event defined here!?

1.1 Relationship between Events

- **Comparability**. Two sets/events A and B are said to be comparable if

$$A \subset B \text{ or } B \subset A$$

i.e. if one of the sets/events is a subset of the other. Two sets/events A and B are said to be not comparable if

$$A \not\subset B \text{ and } B \not\subset A$$

We thus note that if two sets/events A and B are not comparable there is necessarily an element in A that is not in B and an element in B that is not in A.

- **Disjoint/exclusive sets/events**. Two sets/events are (mutually) disjoint/exclusive if they have no elements in common i.e. the intersection of the sets/events is the empty set.

- **Union of sets/events**. The union of two sets/events A and B is the set/event consisting of all elements in A plus all elements in B and is denoted by A$\cup$B or A $+$ B.

  Example. If A $=$ $\{$a, b, c, d$\}$ and B $=$ $\{$b, c, e, f, g$\}$ then $A \cup B = \{a, b, c, d, e, f, g\}$.

- **Intersection of sets/events**. The intersection of two sets/events A and B is the set/event consisting of all elements that occur in both A and B (i.e. all elements common to both), denoted by $A \cap B$, or AB.

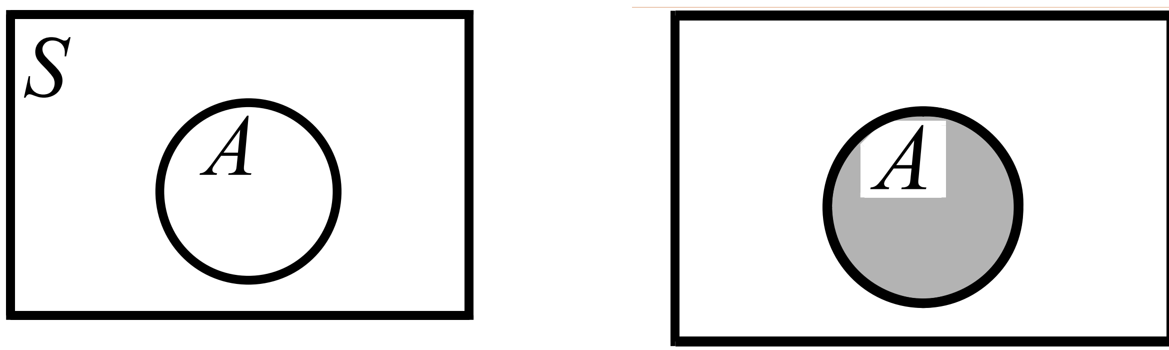
Example. If A = {a, b, c, d} and B = {b, c, e, f, g} then A∩B = {b, c}.

- **Difference of two sets/events**. The set/event consisting of all elements of a set/event A that do not belong to a set/event B is called the difference of A and B and denoted by A - B.

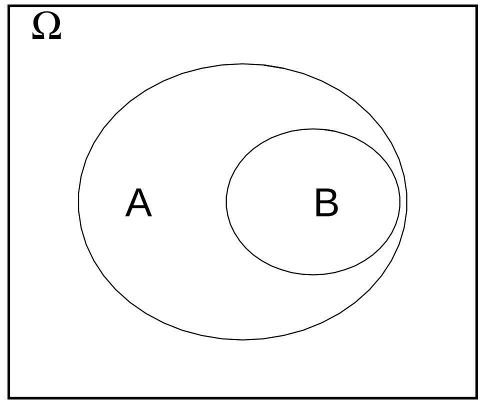
  Example. If $A = \{a, b, c, d\}$ and $B = \{b, c, e, f, g\}$ then $A - B = \{a, b\}$.

- **Complement of a set/event**. The complement of a set/event A with respect to a given universal set/event Ω is the set/event of elements in Ω that are not in A. Denoted by $\bar{A}$ or A^c.

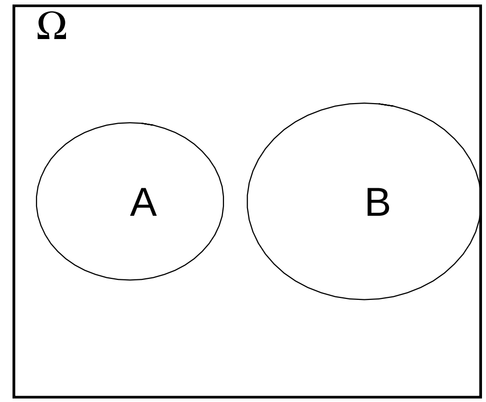
Venn diagrams for the relations: using a rectangle (with area 1) to represent the sample space, any sets inside to represent an event.



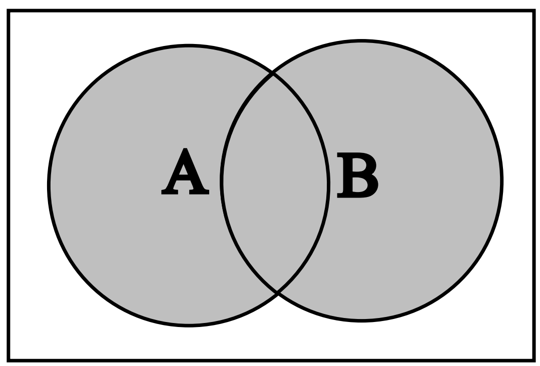
(a) Sample space containing event A (b) Event A shaded



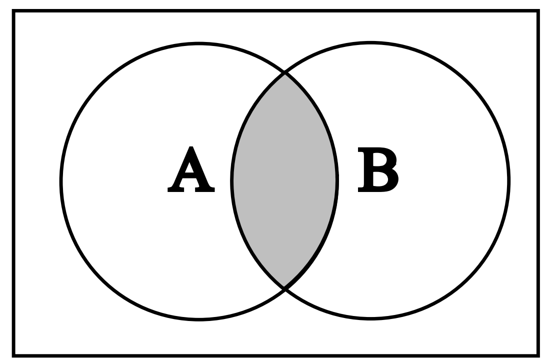
$A \supset B$



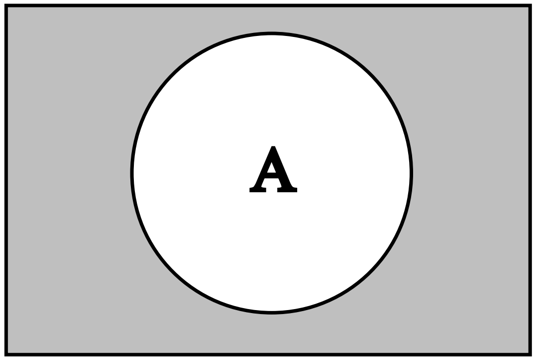
$A \cap B = \phi$



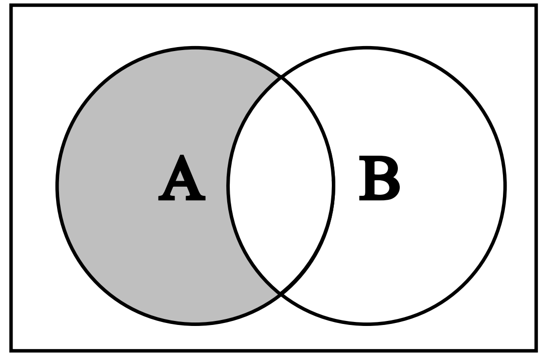
$\mathbf{A} \cup \mathbf{B}$



$\mathbf{A} \cap \mathbf{B}$



$\tilde{\mathbf{A}}$



$\mathbf{A-B}$

Example Based on the Venn diagram, prove that

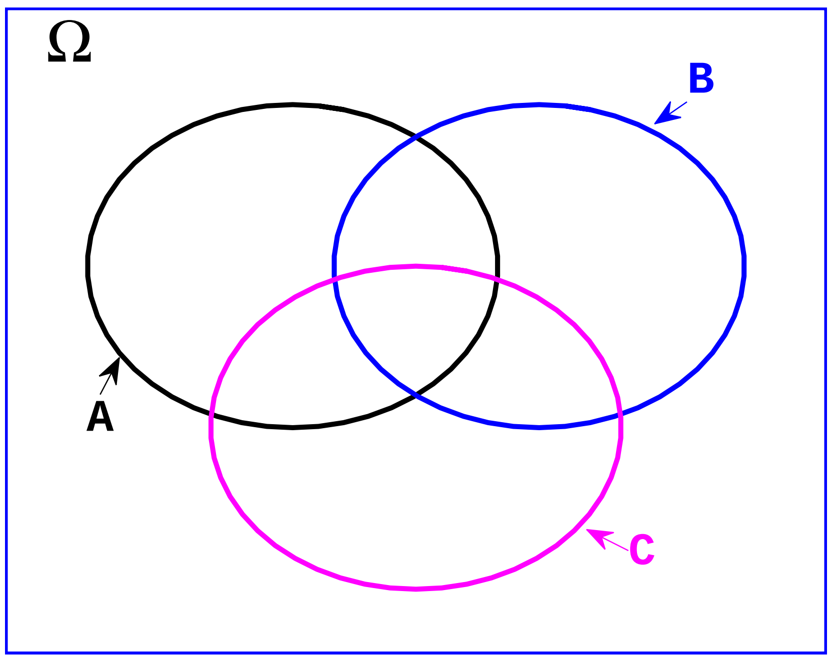
1. $A-(B\cup C)=(A-B)\cap(A-B)$

2. $\overline{A\cup B}=\bar{A}\cap\bar{E}$

3. $A\cup(B\cap C)=(A\cup B)\cap(A\cup C)$

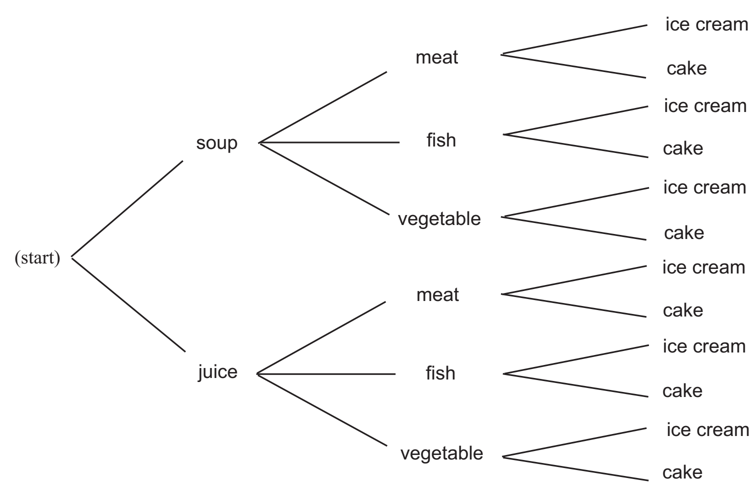
4. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

5. Using the operators to denote the smallest regions?



2 Counting Techniques

Example You are eating at a restaurant and the waiter informs you that you have (a) two choices for appetizers: soup or juice; (b) three for the main course: a meat, fish, or vegetable dish; and (c) two for dessert: ice cream or cake. How many possible choices do you have for your complete meal?



Theorem A task is to be carried out in a sequence of r stages. There are n_1 ways to carry out the first stage; for each of these n_1 ways, there are n_2 ways to carry out the second stage; for each of these n_2 ways, there are n_3 ways to carry out the third stage, and so forth. Then the total number of ways in which the entire task can be accomplished is

$$N = n_1 \times n_2 \times ... \times n_r.$$

Factorial If there are n objects, the number of different ways of arranging or ordering them is

$$n! = n \times (n-1) \times (n-2) \times ... \times 2 \times 1$$

Example Suppose that you have a deck of card and you are only interested in shuffling the suit of hearts. How many different possible orders can there be for shuffling the hearts?

Example Suppose you wish to take 5 cards from the suit of hearts. How many different possible combinations that you can take 5 cards? Do I care about the order in which I take the 5 cards? what if I dont, what if I do?

Permutation (also called "arrangement number" or "order") When chosen r objects from n objects, there are

$$P_n^r = \frac{n!}{(n-r)!}$$

different permutation. The 'order' of the items IS considered.

Example For $\{1, 2, 3, 4\}$, choose 3 numbers from it, we have the following 24 permutations:

{1,2,3}, {1,3, 2}, {2,1, 3}, {2, 3, 1}, {3,2,1}, {3,1,2},

{1,2,4}, {1,4, 2}, {2,1, 4}, {2, 4, 1}, {4,2,1}, {4,1,2},

{1,3,4}, {1,4, 3}, {3,1, 4}, {3, 4, 1}, {4,3,1}, {4,1,3},

{2,3,4}, {2,4, 3}, {3,2, 4}, {3, 4, 2}, {4,3,2}, {4,2,3}

Combinations When chosen r objects from n objects, there are

$$\binom{n}{r} = C_n^r = \frac{n!}{r!(n-r)!}$$

different combinations. Note that when we count the number of combinations, the 'order' of the items is NOT considered.

Example (The same example above) For $\{1, 2, 3, 4\}$, choose 3 numbers from it, we have the following 4 combinations:

~~{1,2,3}, {1,3, 2}, {2,1, 3}, {2, 3, 1}, {3,2,1}, {3,1,2}~~, $\Rightarrow \{3, 1, 2\}$

~~{1,2,4}, {1,4, 2}, {2,1, 4}, {2, 4, 1}, {4,2,1}, {4,1,2}~~, $\Rightarrow \{4, 1, 2\}$

~~{1,3,4}, {1,4, 3}, {3,1, 4}, {3, 4, 1}, {4,3,1}, {4,1,3}~~, $\Rightarrow \{3, 1, 4\}$

~~{2,3,4}, {2,4, 3}, {3,2, 4}, {3, 4, 2}, {4,3,2}, {4,2,3}~~, $\Rightarrow \{3, 4, 2\}$

3 Probability

3.1 Classical Approach:

Classical probability is predicated on the assumption that the outcomes of an experiment are equally likely to happen.

$$P(A) = \frac{\text{Number of outcomes/elements in } A}{\text{Total number of outcomes/elements in } \Omega}.$$

The classical probability is used when

- the events have the same chance of occurring (called equally likely events)
- and the set of events are mutually exclusive and collectively exhaustive.

Experiment: Rolling 2 die [***dice***] and summing 2 numbers on top.

Sample Space: S = {2, 3, …, 12}

What are the underlying, unstated assumptions??

Probability Examples:

P(2) = 1/36

P(7) = 6/36

P(10) = 3/36

	1	**2**	**3**	**4**	**5**	**6**
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Example: A box contains three light bulbs, and one of them is bad. If two light bulbs are chosen at random from the box, what is the probability that one of the two bulbs is bad?

method 1: Assume that the bulbs are labeled A, B, and C and that light bulb A is bad. If the two bulbs are selected one at a time, there are $P_3^2 = 6$ possible outcomes of the experiment: {AB, BA,AC, CA,CB, BC}. Note that 4 of the 6 outcomes result in choosing A, the bad light bulb. If these 6 outcomes are equally likely, then the probability of selecting the bad bulb is $\frac{4}{6} = \frac{2}{3}$.

method 2: ignore the order, we have outcomes with one being bad {AB, AC} , the total outcomes {AB, AC, BC}. The probability is also $\frac{2}{3}$.

Example (birthday problem or birthday paradox) For n randomly chosen people, what is the probability that there is at least two of them have the same birth date?

All possible outcomes

$$365^n$$

The events that no one has the same birthdate as anyone of others

$$365 \times (365-1) ... \times (365-n+1)$$

The probability for B = (all birthdates are different)

$$P(B) = \frac{365 \times (365-1) ... \times (365-n+1)}{365^n}$$

There are at least 2 people who have the same birthdate is

$$P(A) = 1 - P(B) = 1 - \frac{365 \times (365-1) ... \times (365-n+1)}{365^n}$$

n	P(at least 2 persons have the same birthdate)
10	0.117
20	0.411
21	0.444
22	0.476
23	0.507
24	0.538
30	0.706
50	0.970
57	0.990
100	0.9999997

3.2 Relative Frequency Approach:

Relative probability is based on cumulated historical data. It assigns the probability:

$$P(A) = \frac{\text{Number of times } A \text{ occurred in the past}}{\text{Total number of opportunities for the event to occur}}$$

Note that relative probability is not based on rules or laws but on what has happened in the past.

Bits & Bytes Computer Shop tracks the number of desktop computer systems it sells over a month (30 days):

Desktops Sold	# of Days
0	1
1	2
2	10
3	12
4	5

For example,
10 days out of 30
2 desktops were sold.

From this we can construct
the "estimated" probabilities of an event
(i.e. the # of desktop sold on a given day)…

Desktops Sold [X]	# of Days	Desktops Sold
0	1	1/30 = .03 =P(X=0)
1	2	2/30 = .07 = P(X=1)
2	10	10/30 = .33 = P(X=2)
3	12	12/30 = .40 = P(X=3)
4	5	5/30 = .17 = P(X=4)
		Σ = 1.00

"There is a 40% chance Bits & Bytes will sell 3 desktops on any given day" [Based on estimates obtained from sample of 30 days]

$$P(\text{selling 3 desktops a day}) = 0.4$$

or denoted as $P(X = 3) = 0.4$.

3.3 Subjective Approach:

The subjective probability is based on personal judgment, accumulation of knowledge, and experience. For example, medical doctors sometimes assign subjective probabilities to the length of life expectancy for people having cancer. Weather forecasting is another example of subjective probability.

P(NASA will successfully land a man on the moon)

P(girlfriend says yes when you ask her to marry you)

3.4 The Axioms of Probability

Any function of events $P(.)$ is a probability (function) if

Axiom 1. $0 \leq P(A) \leq 1$ for any event A.

Axiom 2. $P(\Omega) = 1$.

Axiom 3. If A and B are mutually exclusive events in Ω, then

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

Note that all the approaches of probability satisfy the 3 axioms.

Axiom probability is not uniquely defined, e.g. tossing a coin, we can assign $P(H) = 0.5, P(T) = 0.5$ or $P(H) = 0.1, P(T) = 0.9$.

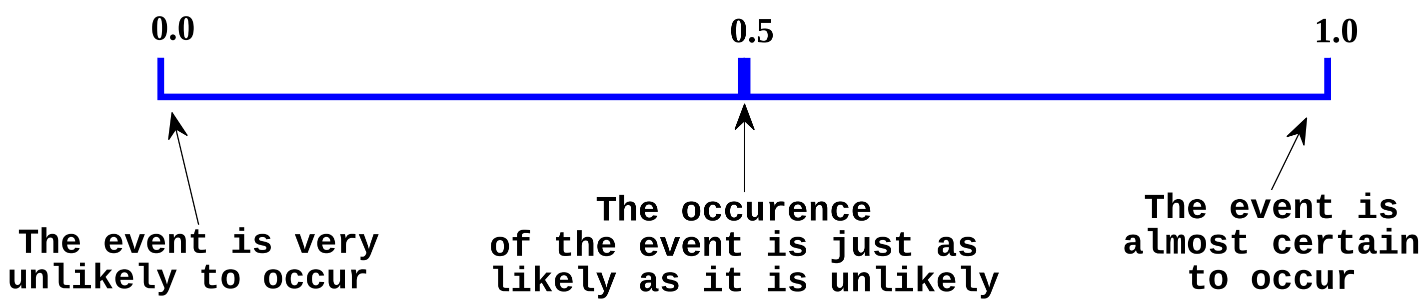
Which approach to be used: according to the following priority

(1) Classical approach (2) Frequency approach (3) Subjective approach

Interpretation of probability:

- If a random experiment is repeated an infinite number of times, the relative frequency for any given outcome is the probability of this outcome.

- The probability is between 0 and 1 with



4 Elementary Theorem

All below can be proved by the Venn diagrams.

1. Monotonicity, if $A \subset B$, then

$$P(A) \leq P(B)$$

2. If A is any event in Ω, then

$$P(A) = 1 - P(\bar{A}).$$

   Thus $P(\emptyset) = 0$.

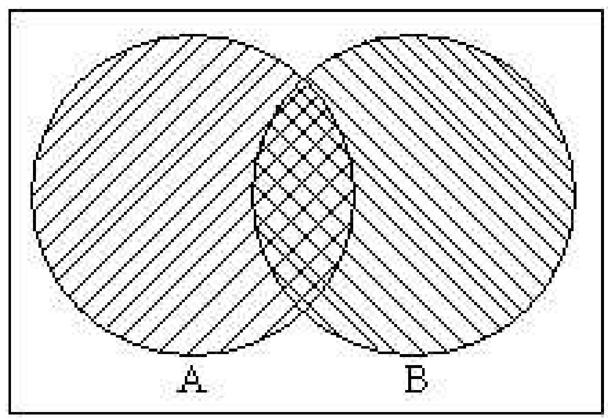
3. $A_1, A_2, ..., A_n$ are mutually exclusive events in a sample space, then

$$P(A_1 \cup A_2 \cup ... \cup A_n) = P(A_1) + P(A_2) + ... + P(A_n)$$

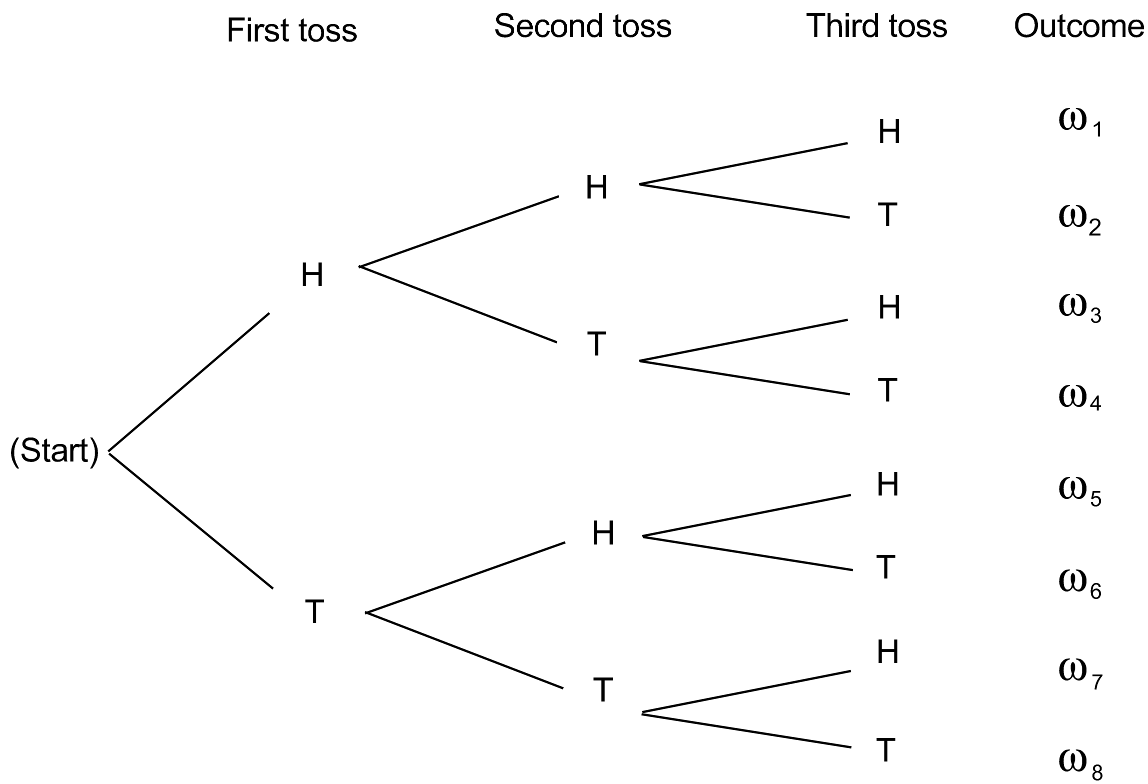
4. If A and B are any two events in Ω, then

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

The intersection must be subtracted out intersection because it would be counted twice:



Example Toss an even coin 3 times. It is convenient to represent the outcomes by a tree diagram as below.



Let A be the event "the first outcome is a head" and B the event "the second outcome is a tail." By looking at the paths in the above tree diagram, we see that

$$P(A) = P(B) = 1/2$$

Moreover, $A \cap B = \{w_3, w_4\}$, and so $P(A \cap B) = 1/4$. Using the elementary Theorem, we obtain

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{1}{2} + \frac{1}{2} - \frac{1}{4} = \frac{3}{4}$$

Another method: Since $A \cup B = \{$HHH,HHT,HTH,HTT,TTH,TTT$\}$; and

$$P(A \cup B) = \frac{\#(A \cup B)}{\#\Omega} = \frac{6}{8}$$

we see that we obtain the same result by direct enumeration.