Chapter 2 Probability

2.1. Sample Space, Events and Probability

Sample Space
A collection of all *outcomes* of an experiment is called a *sample space*.

Event
An *event* is a set of outcomes, and simultaneously, a subset of the sample space.

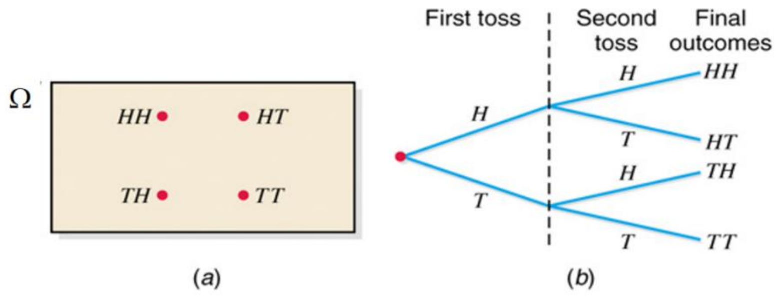
Example:

Experiment	*Outcomes*	*Sample Space*
Toss a coin once	Head, Tail	Ω = { Head, Tail }
Roll a die once	1,2,3,4,5,6,	Ω = {1,2,3,4,5,6}
Select a worker	Male, Female	Ω = { Male, Female}
Take a test	Pass, Fail	Ω = { Pass, Fail }

A ***Venn diagram*** is a picture that depicts all the possible outcomes for an experiment.

A ***tree diagram*** is a picture that represents each outcome by a branch of the tree.

Example 2.1.
Draw the Venn and tree diagrams for the experiment of tossing a coin twice.
Solution:
Experiment: Tossing a coin twice
Let H = head obtained from a toss of coin
T = tail obtained from a toss of coin
Sample space, Ω = {HH, HT, TH, TT}



Example 2.2.
List out the outcome(s) for each of the following event.

Experiment	Possible events
(a) Roll a die once $\Omega=\{1,2,3,4,5,6\}$	Event A = roll a 4 = Event B = roll an even number =
(b) Toss two coins $\Omega=\{HH,HT,TH,TT\}$	Event C = at least one head = Event D = exactly one head =

2.2. Calculating Probability

Probability is a numerical measure of the likelihood that a specific event will occur.

Two properties of probability.

1. The probability of an event always lie in the range 0 to 1.
   i.e. $0 \le P(A) \le 1$ for event A

2. The sum of the probabilities of all individual events (or final outcomes) is always 1.
   i.e. $\sum P(A_i) = P(A_1) + P(A_2) + \ldots = 1$

Extreme cases.

A sample space Ω consists of all possible outcomes, therefore, it occurs for sure. On the contrary, an empty event $\varnothing$ never occurs. Hence,

$$P\{\Omega\} = 1 \text{ and } P\{\varnothing\} = 0.$$

Three conceptual approaches to probability.

1. **Classical Probability**
   Each outcome in the sample space has the **equally likely** probability of occurrence.
   If an experiment can result in any one of N different equally likely outcomes, and if exactly n of these outcomes correspond to event A, then the probability of event A is

$$P(A) = \frac{n}{N} = \frac{n(A)}{n(\Omega)} .$$

Example 2.3.

A fair die is thrown. Let A be the event "the number is odd" and B be the event "the number is greater than 4".

(a) State the sample space.

(b) Find $P(A)$ and $P(B)$.

Solution:

2. **Relative Frequency Concept of Probability**
   Outcomes for the corresponding experiments are **NOT** equally likely.
   To calculate such probabilities, perform the experiment again and again to generate data to obtain the relative frequency. If an experiment is repeated n times and an event A is observed f times, then, according to the relative frequency concept of probability,

$$P(A) \simeq \frac{f}{n} .$$

> ***Law of Large Numbers***
> If an experiment is repeated again and again, the probability of an event obtained from the relative frequency approaches the actual or theoretical probability.

Example 2.4.

A survey of 350 families gave the following data on the number of children under 16 years old in each family.

Number of children under 16 years old	Frequency
0	185
1	51
2	90
3 or more	24
Total	350

Find the probability of the following events

(a) *A*, a household selected has no children under 16 years old;

(b) *B*, a household selected has at least one child under 16 years old.

Solution:

3. **Subjective Probability**

   Subjective probability is the probability assigned to an event based on subjective judgment, experience, information and belief.

   Examples:

   (a) The probability that Carol, who is taking statistics, will earn an *A* in this course.

   (b) The probability that the Dow Jones Industrial Average will be higher at the end of the next trading day.

   (c) The probability that Joe will lose the lawsuit he has filed against his landlord.

Axioms of Probability

Suppose Ω is a sample space associated with an experiment. To every event A in Ω, a number $P(A)$, called the probability of event *A*, the following axioms hold:

Axiom 1 : $0 \le P(A) \le 1$

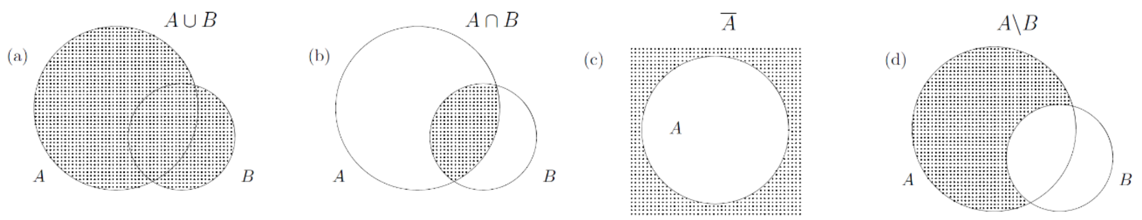
Axiom 2 : $P(\Omega) = 1$

Axiom 3 : If $A_1, A_2, A_3, \ldots$ form a sequence of pairwise mutually exclusive events in Ω, that is, $A_i \cap A_j = \varnothing$ for $i \ne j$, then $P(A_1 \cup A_2 \cup A_3 \cup \ldots) = \sum_{i=1}^{\infty} P(A_i)$.

2.3. <u>Set Operations</u>

Events are *sets* of outcomes. Therefore, to learn how to compute probabilities of events, we shall discuss some *set theory*.

<u>Venn diagrams for (a) union, (b) intersection, (c) complement and (d) difference of events.</u>



Union
A *union* of events $A, B, C, \ldots$ is an event consisting of *all* outcomes in all these events. It occurs if *any* of $A, B, C, \ldots$ occurs, and therefore, corresponds to the word "OR": A or B or C or $\ldots$.

$$A \cup B \text{ = union}$$

Intersection
An *intersection* of events $A, B, C, \ldots$ is an event consisting of outcomes that are *common* in all these events. It occurs if *each* $A, B, C, \ldots$ occurs, and therefore, corresponds to the word "AND": A and B and C and $\ldots$.

$$A \cap B \text{ = intersection}$$

Complement
A *complement* of an event A is an event that occurs every time when A does not occur. It consists of outcomes excluded from A, and therefore, corresponds to the word "NOT": not A.

$$\overline{A} \text{ = complement}$$

Difference
A *difference* of events A and B consists of all outcomes included in A but excluded from B. It occurs when A occurs and B does not, and corresponds to "BUT NOT": A but not B.

$$A \backslash B \text{ = difference}$$

Mutually Exclusive
Events A and B are *disjoint* if their intersection is empty.

$$A \cap B = \varnothing$$

Events $A_1, A_2, A_3, \ldots$ are *mutually exclusive* or *pairwise disjoint* if any two of these events are disjoint.

$$A_i \cap A_j = \varnothing \text{ for any } i \neq j$$

Exhaustive
Events $A, B, C, \ldots$ are *exhaustive* if their union equals the whole sample space.

$$A \cup B \cup C \cup \ldots = \Omega$$

2.4. Rules of Probability

Intersection

The Joint Probability

The probability of the intersection of two events is called their *joint probability* and written as

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

Example 2.5.

The table below classifies 140 students by their gender and favorite indoor leisure activities. A student is chosen at random from the sample. Find the probability that the student is a female and prefers reading.

Gender	Puzzles (A)	Reading (B)	Painting (C)
Male (M)	35	20	15
Female (F)	22	35	13

Solution:

Gender	Puzzles (A)	Reading (B)	Painting (C)	*Total*
Male (M)	35	20	15	
Female (F)	22	35	13	
Total				

The Joint Probability of Mutually Exclusive Events.

Events that cannot occur together are said to be *mutually exclusive events*. The joint probability of two *mutually exclusive events* is always zero.

If A and B are two mutually exclusive events, then

$$P(A \cap B) = 0.$$

Example 2.6.

Consider the following events for one roll of a die.

A = an even number is observed
B = an odd number is observed
C = a number less than 5 is observed

Are events A and B mutually exclusive? Are events A and C mutually exclusive?

Solution:

Union

<u>The Additive Law of Probability.</u>

The probability of the union of two events *A* and *B* is

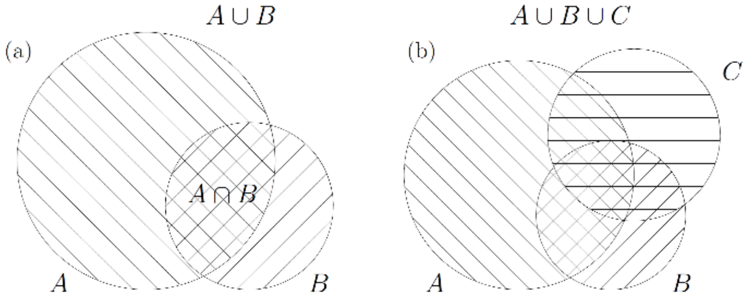
$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

If *A* and *B* are two mutually exclusive events, that is, $P(A \cap B) = 0$, then

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

<u>Example 2.7.</u>

Generalize the formula for probability of a union of three events A, B and C as shown in the following Venn diagram (b).



Solution:

<u>Example 2.8.</u>

The table below classifies 7955 workers in a company by their gender and marital status. Suppose a worker is randomly selected, find the probability that the worker is male or single.

Gender	Single (A)	Married (B)	*Total*
Male (M)	1562	2675	4237
Female (F)	1960	1758	3718
Total	3522	4433	7955

Solution:

Complimentary

Recall that events A and $\overline{A}$ are exhaustive, hence $A \cup \overline{A} = \Omega$. Also, they are disjoint, hence

$$P(A) + P(\overline{A}) = P(A \cup \overline{A}) = P(\Omega) = 1.$$

Solving this for $P(\overline{A})$ and obtain

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

Example 2.9.

If a system appears protected against a new computer virus with probability 0.7, find the probability that the system is exposed to it.

Solution:

Example 2.10.

Suppose a computer code has no errors with probability 0.45. Find the probability that it has at last one error.

Solution:

2.5. Conditional Probability and Independence

Conditional Probability

Suppose you are meeting someone at an airport. The flight is likely to arrive on time; the probability of that is 0.8. Suddenly it is announced that the flight departed one hour behind the schedule. Now it has the probability of only 0.05 to arrive on time. New information affected the probability of meeting this flight on time. The new probability is called conditional probability, where the new information, that the flight departed late, is a condition.

Conditional probability of event A given event B is the probability that A occurs when B is known to occur.

$$P(A|B) = \text{conditional probability of } A \text{ given } B$$

If A and B are two events, then

$$P(B|A) = \frac{P(A \cap B)}{P(A)}, \; P(A) \neq 0$$

and

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \; P(B) \neq 0.$$

Multiplicative Law of Probability

The probability of the intersection of two events A and B is

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

Example 2.11.

The following is a two way classification of the responses of 100 researchers whether they are in favor of or against genetic engineering.

Gender	*In Favor (V)*	*Against (A)*	Total
Male (M)	15	45	60
Female (F)	4	36	40
Total	19	81	100

Suppose a researcher is selected at random, find the probability that the researcher selected is

(a) a male.

(b) in favor of genetic engineering.

(c) against to genetic engineering given that this researcher is a female.

(d) a male given that this researcher is in favor of genetic engineering.

Solution:

Independence
Two events are said to be independent if the occurrence of one does not affect the probability of the occurrence of the other. In other words, A and B are *independent events* if either

1. $P(A \cap B) = P(A)P(B)$,
2. $P(A \mid B) = P(A)$,
3. $P(B \mid A) = P(B)$.

If the occurrence of one event ***affects*** the probability of the occurrence of the other event, then the two events are said to be *dependent events*.

Example 2.12.
Ninety percent of flights depart on time. Eighty percent of flights arrive on time. Seventy-five percent of flights depart on time and arrive on time.
(a) You are meeting a flight that departed on time. What is the probability that it will arrive on time?
(b) You have met a flight, and it arrived on time. What is the probability that it departed on time?
(c) Are the events, departing on time and arriving on time, independent?

Solution:
Let A = flights arriving on time
D = flights departing on time

Example 2.13.
Given $P(X) = 0.28$, $P(Y) = 0.36$ and $P(Y \mid X) = 0.6$. Find
(a) $P(X \mid Y)$ [0.4667]
(b) $P(X \mid \overline{Y})$ [0.175]
(c) $P(X \cup \overline{Y})$ [0.808]

2.6. Applications in Reliability

Formulas of the previous section are widely used in *reliability*, when one computes the probability for a system of several components to be functional.

Example 2.14.

There is a 1% probability for a hard drive to crash. Therefore, it has two backups, each having a 2% probability to crash, and all three components are independent of each other. The stored information is lost only in an unfortunate situation when all three devices crash. What is the probability that the information is saved?

Solution:

Let H = hard drive crashes

B_1 = first backup crashes

B_2 = second backup crashes

Example 2.15.

Suppose that a shuttle's launch depends on three key devices that operate independently of each other and malfunction with probabilities 0.01, 0.02, and 0.02, respectively. If any of the key devices malfunctions, the launch will be postponed. Compute the probability for the shuttle to be launched on time, according to its schedule.

Solution:

Let A = device 1 malfunction

B = device 2 malfunction

C = device 3 malfunction

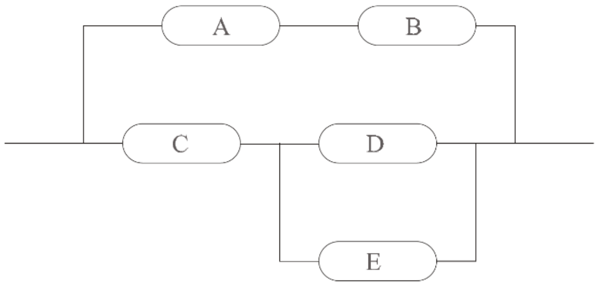
When the system's components are connected in parallel, it is sufficient for at least one component to work in order for the whole system to function. Reliability of such a system is computed as in *Example 2.14.*. Backups can always be considered as devices connected in parallel.

At the other end, consider a system whose components are connected in sequel. Failure of one component inevitably causes the whole system to fail. Such a system is more "vulnerable". In order to function with a high probability, it needs each component to be reliable, as in *Example 2.15.*.

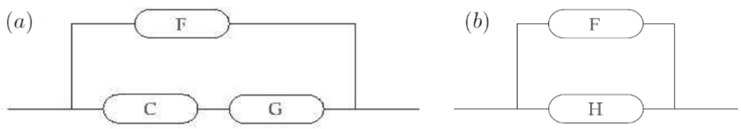
Many modern systems consist of a great number of devices connected in sequel and in parallel.

Example 2.16.

Calculate reliability of the system in the following figure if each component is operable with probability 0.92 independently of the other components.



Solution:



1. The upper link A-B works if both A and B work, which has probability

   Represent this link as one component F that operates with probability ___________.

2. By the same token, components D and E, connected in parallel, can be replaced by component G, operable with probability

3. Components C and G, connected sequentially, can be replaced by component H, operable with probability

4. Lastly, the system operates with probability

In fact, the event "the system is operable" can be represented as $(A \cap B) \cup (C \cap (D \cup E))$.

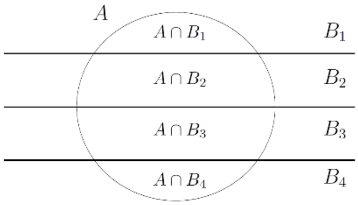
2.7. Bayes' Rule

For some positive integer *k*, let the events $B_1, B_2, ..., B_k$ be such that

1. $\Omega = B_1 \cup B_2 \cup ... \cup B_k$, ***Exhaustive***
2. $B_i \cap B_j = \varnothing$ for any $i \neq j$. ***Mutually Exclusive***

Then the collection of sets $\{B_1, B_2, ..., B_k\}$ is said to be a ***partition*** of the sample space Ω.

If the events $B_1, B_2, ..., B_k$ constitute a partition of the sample space Ω such that $P(B_i) \neq 0$ for $i = 1, 2, ..., k$, then for any event *A* of Ω,



$$A = (A \cap B_1) \cup (A \cap B_2) \cup \ldots (A \cap B_k),$$

and this is also a union of mutually exclusive events. Hence

$$P(A) = \sum_{i=1}^{k} P(A \cap B_i) = \sum_{i=1}^{k} P(A \mid B_i) P(B_i).$$ ***Law of Total Probability***

If $P(A) > 0$, then with ***Bayes' Rule***

$$P(B_r \mid A) = \frac{P(B_r \cap A)}{P(A)} = \frac{P(A \mid B_r) P(B_r)}{\sum_{i=1}^{k} P(A \mid B_i) P(B_i)}$$ ***Posterior Probability***

for $r = 1, 2, ..., k$.

Example 2.17.
According to a report, 7.0% of the population has lung disease. Of those having lung disease, 90.0% are smokers; of those not having lung disease, 25.3% are smokers. Determine the probability that a randomly selected smoker has lung disease.
Solution:
Let L = lung disease
M = smokers
$P(L)=0.7$, $P(M\,|\,L)=0.9$, $P(M\,|\,\bar{L})=0.253$
Find $P(L\,|\,M)=?$

Example 2.18.
There exists a test for a certain viral infection (including a virus attack on a computer network). It is 95% reliable for infected patients and 99% reliable for the healthy ones. That is, if a patient has the virus (event V), the test shows that (event S) with probability $P(S\,|\,V)=0.95$, and if the patient does not have the virus, the test shows that with probability $P(\bar{S}\,|\,\bar{V})=0.99$.
Consider a patient whose test result is positive (i.e., the test shows that the patient has the virus). Knowing that sometimes the test is wrong, naturally, the patient is eager to know the probability that he or she indeed has the virus. Suppose that 4% of all the patients are infected with the virus, $P(V)=0.04$. Find the probability that a patient has the virus if the test shows positive results.
Solution:

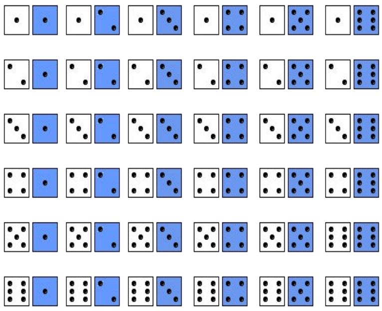
2.8. Counting Sample Points

Multiplicative Rule

If an operation can be performed in n_1 ways, and if for each of these a second operation can be performed in n_2 ways, and for each of the first two a third operation can be performed in n_3 ways, and so forth, then the sequence of k operations can be performed in $n_1 n_2 \ldots n_k$ ways.

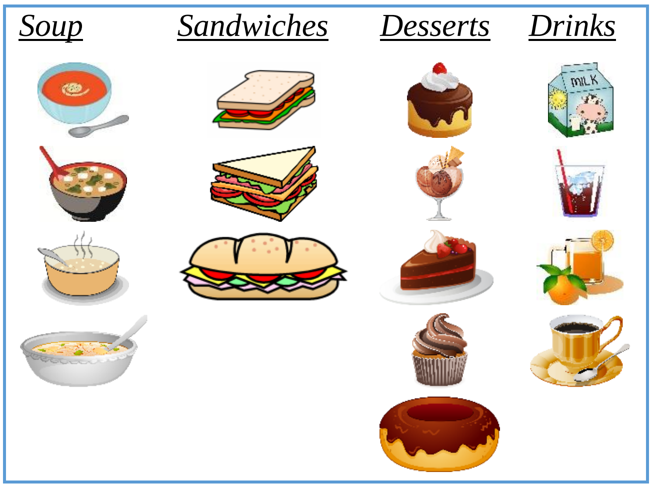
Example 2.19.

How many sample points are in the sample space when a pair of dice is thrown once?



Example 2.20.

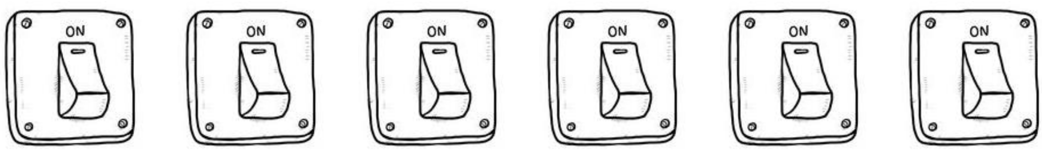
How many lunches consisting of a soup, sandwich, dessert and a drink are possible if we can select from 4 soups, 3 kinds of sandwiches, 5 desserts and 4 drinks?



Repetitions are Allowed

Example 2.21.

There are 6 light switches on the wall, how many different orders of on/off are there?



Repetitions are Not Allowed

Example 2.22.

How many four digit integers (greater than 2000) can be formed using digits 1, 2, 3, 4 if no digit is repeated?

Solution:

Example 2.23.

How many three digit numbers can be made from the integers 2, 3, 4, 5, and 6 if

(a) each integer is used only once;

(b) there is no restriction on the number of times each integer can be used;

(c) the first digit must be 5 and repetition is not allowed?

(d) the first digit must be 5 and repetition is allowed?

Solution:

Example 2.24.

(a) From an alphabet consisting of 10 digits, 26 lower-case and 26 upper-case letters, how many different 8-character passwords can be formed?

(b) Suppose at a speed of 1 million passwords per second, how long will it take a spy program to try all of them? Find the probability that it guesses your password in a week.

(c) If capital letters are not used, how long will it take a spy program to try all of them? Find the probability that it guesses your password in a week.

Solution:

2.9. Permutation

An arrangement of all or part of a set of objects for which the order of the objects is important. The number of permutations of *n* distinct objects is $n!$ (read as "n factorial"), namely,

$$n! = n(n-1)(n-2)\ldots(3)(2)(1) \text{ and } 0! = 1.$$

The number of permutations of *n* distinct objects taken *r* at a time is ${}^{n}P_{r} = \dfrac{n!}{(n-r)!}$.

Permutation: Arranging a Subset
Given a bunch of objects but only arrange a few.

Example 2.25.
A club has 20 members. In how many ways can a committee of three comprises of President, Vice-President and Treasurer be formed?
Solution:

Permutation: Specific Positions
A particular position must be occupied by a particular item.

Example 2.26.
How many ways can Alee, Abu, Bakar and Muthu be seated in a row if Abu must be in the second chair?
Solution: 3 1 2 1
Abu

Permutation: Adding Permutations
More than one case in arrangement of objects.

Example 2.27.
How many words (of any number of letters) can be formed from "P, E, N"?
Solution:

Permutation: Items Always Together
Certain items must always be kept together.
Steps.

1. Treat together elements as 1 unit
2. { permutation of *together*-unit } × { permutation of *total*-unit }

Example 2.28.
How many ways can 3 calculus books, 4 statistics books and 6 finance books be arranged on a shelf if the books of each subject must be kept together?
Solution:

Example 2.29.
How many ways can all letters “W, I, N, T, E, R” be arranged if the vowels must be together?
Solution:

Permutation: Items Never Together
Certain items must be kept apart.
Methods.

3. { permutation of *all* items } – { permutation of *together* items }
4. { permutation of *no restriction* items} × { possible space for items *to be separated*}

**Method 1 does not work for more than 2 items to be separated.

Example 2.30.
How many ways can all the letters “A, C, U, T, E” be arranged if C and T must not be together?
Solution:

Example 2.31.
Five boys and four girls must stand in line for a picture.
(a) How many line-up’s will have the boys separated from each other?
(b) How many line-up’s will have the girls separated from each other?
Solution:

Permutation: Identical Items
Some of the items to be arranged are identical.
Steps. Dividing out repetitions.

Example 2.32.
How many ways can all of the letters “S, T, A, T, I, S, T, I, C, S” be arranged in a row?
Solution:

Example 2.33.
Solve the following equation for n:

(a) $\dfrac{(n-2)!}{n!} = \dfrac{1}{12}$ [n = 4]

(b) $2\left({}^{n}P_2\right) + 50 = {}^{2n}P_2$ [n = 5]

2.10. Combination

A combination is actually a partition with two cells, the one cell containing the *r* objects selected and the other cell containing the (*n* – *r*) objects that are left.

The number of combinations of n distinct objects taken r at a time is ${}^nC_r = \binom{n}{r} = \dfrac{n!}{r!(n-r)!}$.

Example: Arrangements for letters CAT (*order is important*)
CAT, CTA, ACT, ATC, TAC, TCA

Grouping of CAT (*order is not important*)
CAT, there is only one way/combination

Combination: Basic Combinations
The order is not important.

Example 2.34.
An antivirus software reports that 3 folders out of 10 are infected. How many possibilities are there?
Solution:

Combination: Specific Items
To include/exclude particular items.

Example 2.35.
A committee of 4 person is to be formed from a group of 9 people. How many committees can be formed if
(a) Andy must be in the committee?
(b) Andy must not be in the committee?
Solution:

Combination: Multiple Selection Pools
To make selection from multiple groups of items.

Example 2.36.
A committee of 2 boys and 3 girls is to be formed from a group of 6 boys and 7 girls. How many committees are possible?
Solution:

Combination: At least/ At most

Could be more than one case.

Example 2.37.

A committee of 5 people is to be formed from a group of 6 boys and 7 girls. How many possible committees can be formed if at least 4 boys are on the committee?

Solution:

Example 2.38.

(a) $\binom{n+2}{6} = \binom{n}{4}$ [n = 4]

(b) $\binom{n}{2} - n = \binom{8}{2} - 1$ [n = 9]

Probability and Combination Analysis

In most of the probability problems, we need to determine the number of possible outcomes in the sample space as well as the event. This is usually involves permutation and combination.

Example 2.39.

There are 20 computers in a store. Among them, 15 are brand new and 5 are refurbished. Six computers are purchased for a student lab. From the first look, they are indistinguishable, so the six computers are selected at random. Compute the probability that among the chosen computer, two are refurbished.

Solution:

Example 2.40.

Five cards are drawn from a pack of 52 well-shuffled cards. Find the probability that

(a) 4 are aces

(b) 4 are aces and 1 is king

(c) 3 are tens and 2 are jacks

(d) a 9, 10, jack, queen, king are obtained in any order

Solution: