Statistics &Data Analysis Concepts for Data Science and ML 4

Probability Fundamentals & Probability Theory

Learning Objectives - What you will learn…

- **Learn the basic concepts of probability theory and understand the importance of probability in statistics and data analysis**
- **Describe the related terms including experiments, sample space, events, equally likely events, mutually exclusive events and other terms associated with probability theory**
- **Describe different ways of calculating and assigning probabilities —classical, relative frequency, and subjective probability approaches**
- **Understand the concepts of sets in probability and difference between mutually exclusive and independent events**
- **Determine probabilities for mutually exclusive and non-mutually exclusive events using the addition laws of probabilities**

Learning Objectives…cont.

- **Calculate probabilities for statistically independent and dependent events**
- **Calculate joint probabilities for both independent and dependent events using laws of multiplication**
- **Understand the concept of conditional probabilities**

Probability Defined

- Probability is what determines the likelihood or, the chance that something will happen.

- Probability is the chance that a particular ***event*** will occur when an ***experiment*** is performed.

- Probability is expressed as a fraction, decimal, or percentage and is between 0 and 1 or, 0% to 100%.

A probability of 0 indicates there is no chance of occurrence and a probability of 1 indicates a 100% chance of occurrence of an event.

The probability of an event A is denoted as P (A), which means "the probability that event A occurs" is between 0 and 1. That is,

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

Important Terms in Probability

- ***Event***: An event is one or more possible outcome of an experiment.
- ***Experiment:*** An experiment is any process that produces an outcome or observation. For example, throwing a die is a simple experiment while the number we get (1 or 2 …., or 6) on the top face is an event. Similarly, tossing a coin is an experiment: getting a head (H) or a tail (T) is an event.
- In probability theory, we use the term ***experiment*** in a very broad sense.
- ***Sample Space***: The set of all possible outcomes of an experiment is called the ***sample space*** and is denoted by S.

Examples: Experiment and Sample Space

- Consider the experiment of tossing a single coin. The outcome of this experiment is either a head (H) or a tail (T) and the sample space S is

$$S = \{H, T\}$$

- An experiment consists of tossing two coins and noting whether they land heads or tails. There are four outcomes of this experiment (H,H), (H,T), (T, H), and (T,T). The sample space S is

$$S = \{(H, H), (H, T), (T, H), (T, T)\}$$

Examples: Experiments and Sample Space

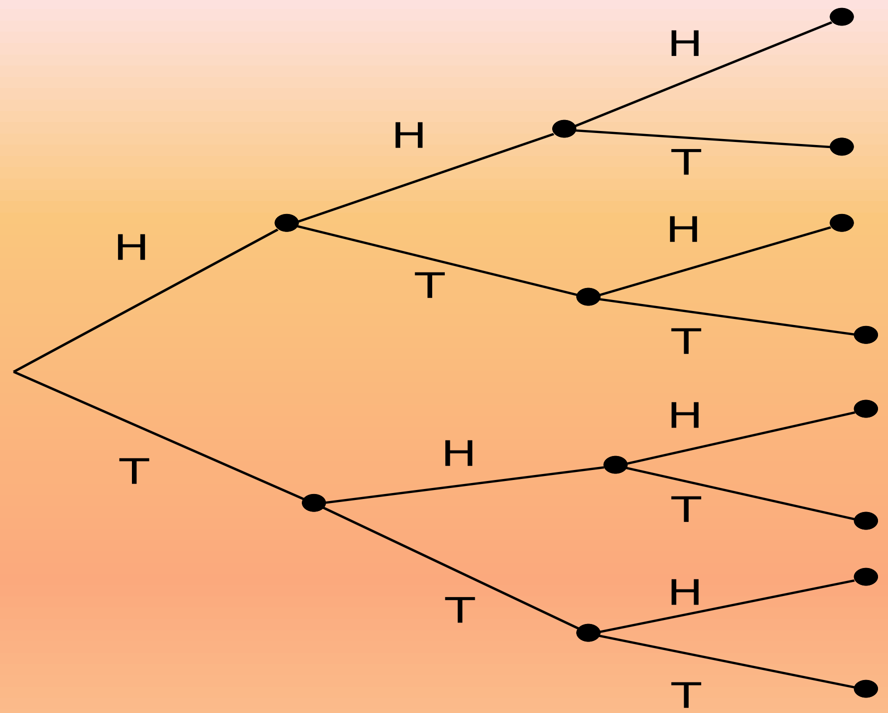
- Consider the experiment of rolling two six-sided dice (one green and the other red) and observing the sum of the numbers on the top faces. If we let (i, j) denote the outcome in which the green die has value i and the red has value j, then the list of all possible outcomes or the sample space is

  S = {(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) }

- Two parts produced by a manufacturing process are being inspected for defects. The parts can be either defective (D) or non defective (ND). The sample space S is:

  S = {(D, D), (D, ND), (ND, D), (ND, ND)}

Sample Space of Tossing a Coin Three Times



S = {HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}

More Probability Terms…

- ***Event:*** An event is one or more possible outcomes of an experiment. An event is a subset of the sample space and is denoted by upper case letters A, B, C, and so on.

[a] Consider the sample space of rolling two six-sided dice. If A is the event that the sum of the numbers on the top faces is 5. Then

A = {(1, 4), (2, 3), (3, 2), (4, 1)}

[b] Suppose we define an event B, which denotes that one of the numbers on the top faces is a “1". Then

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

- ***Mutually Exclusive Events:*** When the occurrence of one event excludes the possibility of occurrence of another event then we say the events are mutually exclusive. In other words, only one event can take place at a time.

Example – Mutually Exclusive Events In tossing a coin, the events Head (H) and Tail (T) are mutually exclusive. In the roll of a six-sided die, the numbers 1 through 6 are mutually exclusive since only one of the numbers 1 through 6 can appear on the top face..

- ***Exhaustive Events***: The total number of possible outcomes in any trial is known as exhaustive events.

  Example: In a roll of two dice, the exhaustive number of events or the total number of outcomes is 36. If three coins are tossed at the same time, the total number of outcomes is 8 (try to list these outcomes).

- ***Equally Likely Events***: A situation where all the events have an equal chance of occurrence or when there is no reason to expect one in preference to the other.

  Example: In tossing a coin, the head or the tail is equally likely. In rolling a single die, all the six faces are equally likely (provided the coin and the die are unbiased).

Counting Rules, Permutations and Combinations

The counting rules determine the number of possible outcomes for a particular experiment

(1) Multiple-Step Experiment or Filling Slots

(2) Permutations

(3) Combinations

(1) Multiple-step Experiment or Filling Slots

Suppose an experiment can be described as a sequence of k steps in which n_1 = the number of possible outcomes for the first step

n_2 = the number of possible outcomes for the second step

:

n_k = the number of possible outcomes for the kth step

Then the total number of possible outcomes is given by

$$(n_1)(n_2)(n_3)\ldots\ldots(n_k)$$

- The above counting rule can also be seen as filling slots. Suppose we want to fill k different slots in which

n_1 = the number of ways for filling the first slot

n_2 = the number of ways for filling the second slot after the first slot is filled

:

n_k = the number of ways for filling kth slot, assuming the slots 1 through (k-1) are filled then the total number of ways for filling k slots can be given by

$$(n_1)(n_2)(n_3)\ldots\ldots(n_k)$$

(2) Permutations

- Permutation is another counting rule that allows us to select ***n objects*** from a set of ***N objects when the order of selection is important***. If the same n objects are selected in a different order, a different outcome results. Determining the number of permutations (arrangements) is a special case of filling slots.

The number of ways of selecting n distinct objects from a group of N objects where the order of selection is important is known as the number of permutations on N objects, using n at a time, and is written as

$$P_n^N = \frac{N!}{(N-n)!} = (n)(n-1)\ldots\ldots(n-k+1)$$

The symbol N! is read as "N factorial" and its value is determined by multiplying N by all positive integers smaller than N. That is,

N! = (N) (N-1) (N-2)..........(2) (1)

For example,

6! = (6) (5) (4) (3) (2) (1) = 720

Note that 0! = 1 not 0 by definition.

Example: Permutation

How many two digit numbers can be constructed using the digits 2, 3, 4, and 5 without repeating any digit?

All possible two digit numbers can be determined by

$$P_n^N = \frac{N!}{(N-n)!} = \frac{4!}{(4-2)!} = \frac{(4)(3)(2)(1)}{(2)(1)} = 12$$

Thus, 12 two digit numbers can be formed. Note that the order of selection is important; that is, 34 is not the same as 43. The 12 permutations in this case can be written as

23	24	25
32	34	35
42	43	45
52	53	54

(3) Combinations

The number of combinations of N objects taken n at a time is given by

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

The order of selection is not important in combination and this disregard of arrangement makes the combination different from the permutation rule. In general, an experiment will have more permutations than combinations.

Example: Combination

- How many ways can a team of eight players be formed from a group of ten players?

  This problem involves selection, not arrangement. Therefore, we can apply the combination formula. The number of ways this can happen can be calculated using the combination formula below where N=10 and n=8

$$C_n^N = \binom{N}{n} = \frac{N!}{n!(N-n)!} = \frac{10!}{8!(10-8)!} = 45$$

- How many combinations of four parts can a quality control inspector select from a batch of 12 parts?

  The number of possible combinations is given by (note that N=12, n=4)

$$C_n^N = \binom{N}{n} = \frac{N!}{n!(N-n)!} = \frac{12!}{4!(12-4)!} = 495$$

Methods of Assigning Probabilities

There are two basic rules of probability assignment.

1. The probability of an event A is written as P(A) and it must be between 0 and 1. That is,

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

2. If an experiment results in *n* number of outcomes say, A_1, A_2..., An then the sum of the probabilities for all the experimental outcomes must equal 1. That is, $P(A_1) + P(A_2) + P(A_3) + \ldots\ldots + P(A_n) = 1$

There are three methods for assigning probabilities

1. ***Classical Method***
2. ***Relative Frequency Approach***
3. ***Subjective Approach***

1. Classical Method

- The classical method of probability is defined as the favorable number of outcomes divided by the total number of possible outcomes. Suppose an experiment has n number of possible outcomes and the event A occurs in *m* of the *n* outcomes. Then the probability that event A will occur is

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

The classical definition of probability equation assumes that ***n possible*** outcomes are equally likely or have the equal chance of occurrence.

Note that P (A) denotes the probability of occurrence of event A. The probability that the event A will not occur is given by P ($\overline{A}$) which is read as P (not A) or 'A complement.'

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

This means that the probability that event A will occur, plus the probability that event A will not occur, must be equal to 1.

Examples on Classical Probability

(a) Suppose that a person is equally likely to be born on any day of the week. What is the probability that a baby is born

[i] On a Sunday

P (S) = 1/7

[ii] On a day beginning with the letter T?

P (Tuesday or Thursday) = 2/7

(b) What is the probability of getting a sum of four when two dice are thrown?

P (sum=4)= 3/36

- ***These are examples of classical probability.***

(2) Relative Frequency Approach

1. **We calculate the relative frequency of an event in a very large number of trials (note that relative frequency is calculated by dividing the frequency by the total number of observations).**

Example: ***suppose that a researcher has determined that 70 out of 20,000 males in the age group 70 to 80 years have a chance of getting a rare type of blood disease. Then the probability or chance of getting this type of disease is***

70/20,000 = 0.0035 or 0.35%

2. In a relative frequency approach, we calculate the proportion of times an event has occurred in the long run when conditions are stable.

Examples on Relative Frequency Approach

Table below provides the frequency distribution for times between failures of 500 electronic components (note that the 1st class interval is 0 but less than 100 hours, and so on).

Time (Hours)	Frequency (f) Number of Failures	Relative Frequency
0-100	130	0.26
100-200	120	0.24
200-300	75	0.15
300-400	50	0.10
400-500	40	0.08
500-600	35	0.07
600-700	20	0.04
700-800	15	0.03
800-900	10	0.02
900-1000	5	0.01

[a] Based on this information, what is the probability that a component will fail between 200 and 300 hours of operation?

P (200 - 300 hours) = 75/500 or 15%

[b] What is the probability that the component will fail in less than 500 hours of operation?

P (less than 500 hours) = 0.26+0.24+0.15+0.10+0.08 = 0.83 or 83%

(c) What is the probability the component will last 700 hours or more?

P (700 hours or more) = 0.03+0.02+0.01 = 0.06 or 6%

3. Subjective Probability

Subjective probability is used when the events occur only once or very few times and when little or no relevant data are available. Subjective probability is a measure of our belief that a particular event will occur.

Example:

For example, suppose a decision is to be made regarding the construction of a nuclear plant at a location where there is some evidence of geological fault. The decision to locate the nuclear plant at this site will depend upon how high the probability is for a nuclear accident at this location. In such a case, there may not be any past data available. The decision or the likelihood of a nuclear accident must be determined based on the judgment or expert opinion.

Basic Concepts of Probability through Sets

Define a set, complement of a set, union and intersection of sets, and venn diagrams. These concepts are important in evaluating different probabilities.

- ***Set:*** ***A set is an aggregate or collection of objects and is denoted by using upper case letters A, B, C, etc. Members of set A are called the elements of A.***

Example: Set A, whose elements are all even numbers between 0 and 10, is written as : A = {2, 4, 6, 8, 10}

- ***Universal Set:*** ***The universal set is a set of all objects under consideration. A universal set is denoted by U. A set is contained within the universal set.***

Concepts of Sets ...cont.

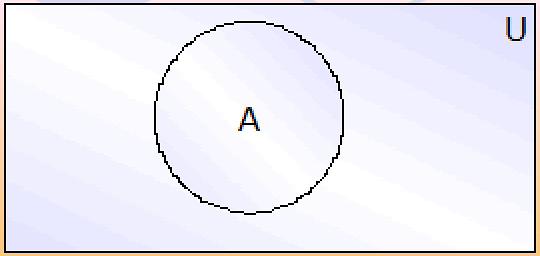
- ***Null Set or Empty Set:*** ***A Null set is a set that contains no elements and is denoted by***

- ***Equality of Sets:*** ***If the set A = {x, y, z} and the set B = {z, x, y}, then A = B (order is immaterial).***

- ***Venn Diagrams:*** ***A Venn diagram is often used to represent a universal set and the sets contained within that universal set.***

In a Venn diagram, the universal set is represented by a rectangle and the set or the event is represented by a circle.

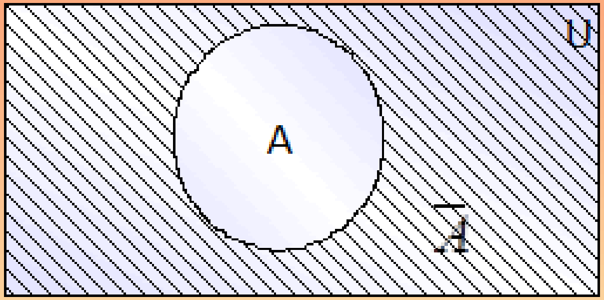
Venn Diagrams and Sets



Example of a Venn diagram

COMPLEMENT OF A SET A

The complement of set A is denoted by $\overline{A}$ (read as A-bar) or A^c and is the set made up of the elements of U that do not belong to A. In other words, the complement of set A is everything but A (with respect to the universal set).

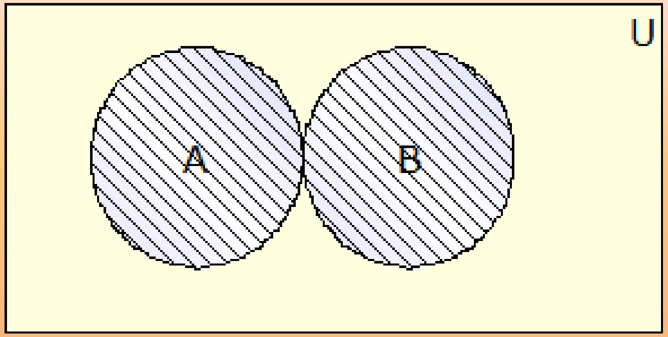


$$P(A) + P(\overline{A}) = 1$$

In the figure above, $\overline{A}$ (complement of A) is the shaded area.

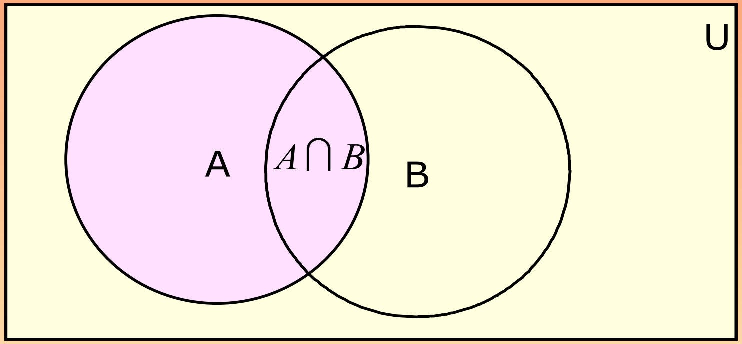
Union of Sets A and B

- ***Union of Sets A and B is denoted by $A \cup B$ (read as "A union B") and is a set of elements that belong to at least one of the sets A or B. In other words, $A \cup B$ contains all the elements of A and all the elements of B but is not repeated. (see the figure)***



Intersection of two Sets A and B

Intersection of two sets A and B is denoted by $A \cap B$(read as "A intersection B").This is the set of elements that belong to both A and B or the elements which are common to A and B. Figure shows the intersection of A and B.



Example

Suppose the universal set U = {1, 2, 3, 4, 5, 6, 7, 8, 9} and three sets A, B, and C are defined as

A = {1, 2, 4, 5, 8} B = {4, 6, 8, 9} C = {1, 3, 5, 7, 8}

Find the following $\overline{A}$ or $A^c, A \cup B, A \cup C, B \cup C, A \cap B, A \cap C$

$$A^c = \{3,6,7,9\}$$

$$A \cup B = \{1,2,4,5,6,8,9\}$$

$$A \cup C = \{1,2,3,4,5,7,8\}$$

$$B \cup C = \{1,3,4,5,6,7,8,9\}$$

$$A \cap B = \{4,8\}$$

$$A \cap C = \{1,5,8\}$$

Addition Law for Mutually Exclusive Events

If we have two events A and B that are mutually exclusive, then the probability that A or B will occur is given by

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

If three events A, B, and C are mutually exclusive then the probability that A or B or C will happen can be given by

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

Addition law for non mutually exclusive events

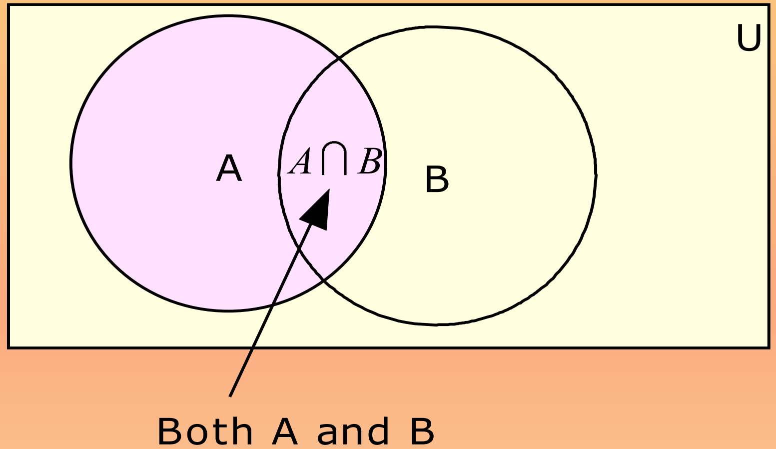
If two events A and B are non-mutually exclusive then they can occur together. If the events A and B are non-mutually exclusive, the probability that A or B will occur is given by

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

$$or, P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad \text{(B)}$$

Note that P(A or B) is same as $P(A \cup B)$ and P(A and B) is same as $P(A \cap B)$ or P(AB).

The difference between equations (A) and (B) : The equations for calculating 'or' probabilities are different when the events are ***mutually exclusive*** and ***non-mutually exclusive***.



The above venn diagram is a situation when A and B both occur. In this case, there is some common area between A and B (intersection of A and B). The events A and B are non-mutually exclusive. And if we want to find P (A or B), we cannot find it by adding P(A) + P (B) because we have counted P (A and B) twice. Therefore we need to subtract P (A and B) to obtain the actual area corresponding to P (A or B).

If events A, B, and C are non-mutually exclusive, then the probability that A or B or C will occur,

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AandB) - P(AandC) - P(BandC) + P(AandBandC)$$

or,

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

Example

Two dice, one green and the other red, are rolled. Let A be the event where the sum of the numbers on the top faces is odd, and B is the event where at least one of the faces is a “1."

[a] Describe the sample space.

The sample space of rolling two dice is shown below

S = {(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) }

[b] Describe the events A, B, $\overline{B}$ (B-complement), $A \cup B$, and find their probabilities, assuming all 36 sample points have equal probabilities.

A = the event that the sum of the numbers shown by the two dice is odd. Therefore,

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

The probability that A will occur

$$P(A) = \frac{18}{36} = \frac{1}{2} \; or, 50\%$$

B = the event that at least one face is "1." Therefore,

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

$$P(B) = \frac{11}{36}$$

$\overline{B}$ = the event that each of face obtained is not an ace or number “1"

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

$$P(\overline{B}) = \frac{25}{36}$$

$A \cup B$ = the elements common to A and B or the event that the sum is odd and least one face is a “1"

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

and,

$$P(A \cup B) = \frac{23}{36}$$

[c] Describe the events $(A \cap B), (A \cap \overline{B})$. What are their probabilities?

$$(A \cap B) = \{(1,2),(2,1),(1,4),(4,1),(1,6),(6,1)\}$$

$$P(A \cap B) = \frac{6}{36} = \frac{1}{6}$$

Example

A survey of 10,000 business professionals found that 20% of all professionals use the internet, 40% use cell phones, and 12% use both the internet and cell phone.

[a] What is the probability that a randomly selected business professional uses the internet?

Let I = event that the professional uses the internet then

P (I) =0.20

[b] What is the probability that a randomly selected business professional uses cell phone?

C= event that the professional uses cell phone then

P(C) = 0.40

[c] What is the probability that a randomly selected business professional uses the internet or cell phone?

In this case, we need to calculate ‘or’ probability. Note that there are two events — I and C —these are not mutually exclusive because there are professionals who use both the internet and cell phone. Recall that if two events A and B are non-mutually exclusive, then the probability that A or B will occur is given by

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

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

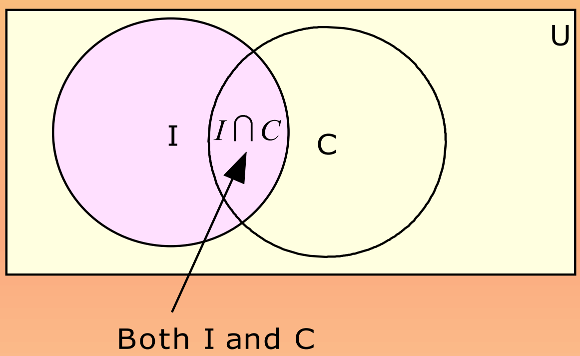
In this case, we have events I and C where, P (I) = 0.20, P(C) =0.40, and P (I and C) =0.12. Therefore, required probability,

$$P(I \cup C) = P(I) + P(C) - P(I \cap C) = 0.20 + 0.40 - 0.12 = 0.48$$

Note that $\cup$ is "or" and $\cap$ is "and." Also, given

$$P(I \cap C) = 0.12$$

[d] What is the probability that a randomly selected business professional uses the internet or the cell phone but not both?



From the above figure,

$P(I \cup C)$ but not both = $P(I \cup C) - P(I \cap C) = 0.48 - 0.12 = 0.36$

[e] What is the probability that a randomly selected business professional uses neither the internet nor the cell phone?

$$P(\overline{I} \cap \overline{C}) = 1 - P(I \cup C) = 1 - 0.48 = 0.52$$

Probabilities of Equally Likely Events and Basic Laws of Probabilities

Equally Likely Events are when all the events have an equal chance of occurrence or when there is no reason to expect one in preference to the other. In tossing a coin, the two possible outcomes head and the tail are equally likely. In rolling a single die, all the six faces are equally likely (provided the coin and the die are unbiased).

Fundamental properties (laws) of assigning probabilities

1. ***The probability of an event A is written as P(A) and it must be between 0 and 1. That is,***

$$0 \leq P(A) \leq 1 \qquad \text{(C)}$$

2. ***If an experiment results in n number of outcomes A_1, A_2, ..., An (to the nth), then the sum of the probabilities for all the experimental outcomes must equal 1. That is,***

$$P(A_1) + P(A_2) + P(A_3) + \ldots\ldots + P(A_n) = 1.0 \qquad \text{(D)}$$

Example

An experiment has four equally likely outcomes: E1, E2, E3, and E4. Assign probabilities to each outcome and show that the basic rules of probability are met; that is, equations (C) and (D) on the previous slide are satisfied.

Since each of the four outcomes is equally likely, each has equal probability of occurrence; that is,

$$P(E1) = P(E2) = P(E3) = P(E4) = 0.25$$

Each probability is within 0 and 1, and the sum of the probabilities for four outcomes is 1.0. Therefore, both rules (C) and (D) are satisfied.

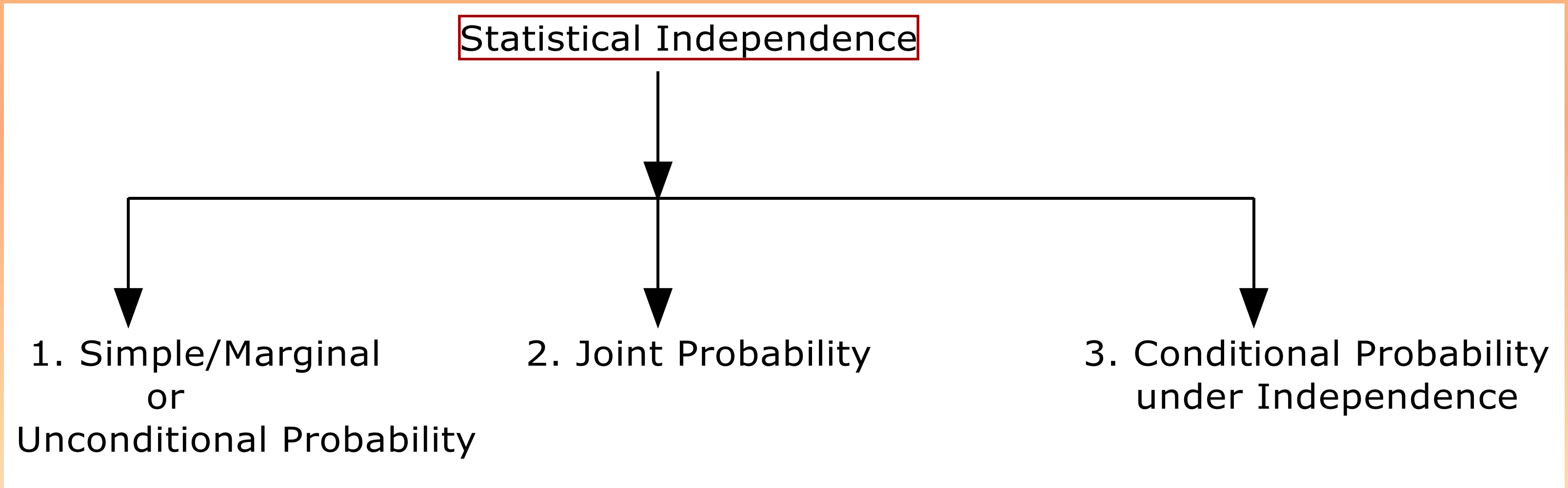
In the previous sections, we calculated probabilities using the classical and relative frequency approach. We also calculated probabilities when the events were mutually exclusive, non-mutually exclusive, and equally likely. There are instances where the events we are interested in are dependent or independent of each other. In the next section, we will first define what is meant by independent and dependent events, and then describe ways of calculating probabilities for independent and dependent cases.

Probabilities under Statistical Independence

When two events occur, the occurrence of the first event may or may not have an effect on the occurrence of the second one. More simply put, the events may be either dependent or independent.

***Statistical Independence:* *When two or more events occur, the occurrence of one event has no effect on the probability of occurrence for any other event. In this case, the events are considered independent*.**

Probabilities under statistical independence



1. ***Simple probability*** is also known as marginal or unconditional, and is the probability of occurrence for a single event; say A, and is denoted by P (A).

2. ***Joint Probability under Statistical Independence:*** Joint probability is the probability of occurrence of two or more events together or in succession. It is also known as “and” probability.

Suppose we have two events A and B which are independent. Then the joint probability, P (AB) is the probability of occurrence of both A “and” B and is given by

$$P(AB) = P(A).P(B)$$

The probability of two independent events occurring together or in succession is ***the product of their marginal or simple probabilities***.

.

$$P(AB) = P(A).P(B)$$

P (AB) = probability of event A and B occurring together.
This is known as ***joint probability***
P(AB) is the same as P (A and B) or,
P(A) = marginal probability of event A
P(B) = marginal probability of event B
$P(AB) = P(A).P(B)$ is also written as:

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

can be extended to more than two events

$$P(ABC) = P(A).P(B).P(C)$$

P(ABC) is same as P(A and B and C) and is also written as $P(A \cap B \cap C)$

Example: Toss a coin twice. What is the probability of getting a head on the first toss, and a head on the second toss?

Suppose: H_1 = probability of getting a head on the first toss, and
H_2 = probability of getting a head on the second toss, then
the probability of a head on the first toss ***and*** the probability of a head on the second toss is

$$P(H_1 H_2) = P(H_1).P(H_2) = (0.5)(0.5) = 0.25$$

Here the events are statistically independent because the probability of any outcome is not affected by any preceding outcome

Example: Note that in tossing a coin three times, the sample space is

S = {HHH, HHH, HTH, HTT, THH, THT, TTH, TTT}

There are eight outcomes; and these are independent of each other. Suppose we want to calculate the following probabilities:

[a] the probability of getting a tail on the first toss, a head on the second toss, and a tail on the third toss.

$$P(THT) = P(T)P(H)P(T) = (0.5)(0.5)(0.5) = 0.125$$

3. *Conditional Probability under Statistical Independence*

The conditional probability is written as

$$P(A|B)$$

and is read as the probability of event A, given that B has already occurred, or the probability of A given B. If two events A and B are ***independent***, then

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

This means that ***if the events are independent, the probabilities are not affected by occurrence of each other***. The probability of occurrence of B has no affect on the occurrence of A. That is, the condition has no meaning if the events are independent.

Example: ***Toss a coin twice. What is the probability of getting a head on the second toss if the first toss resulted in a head?***

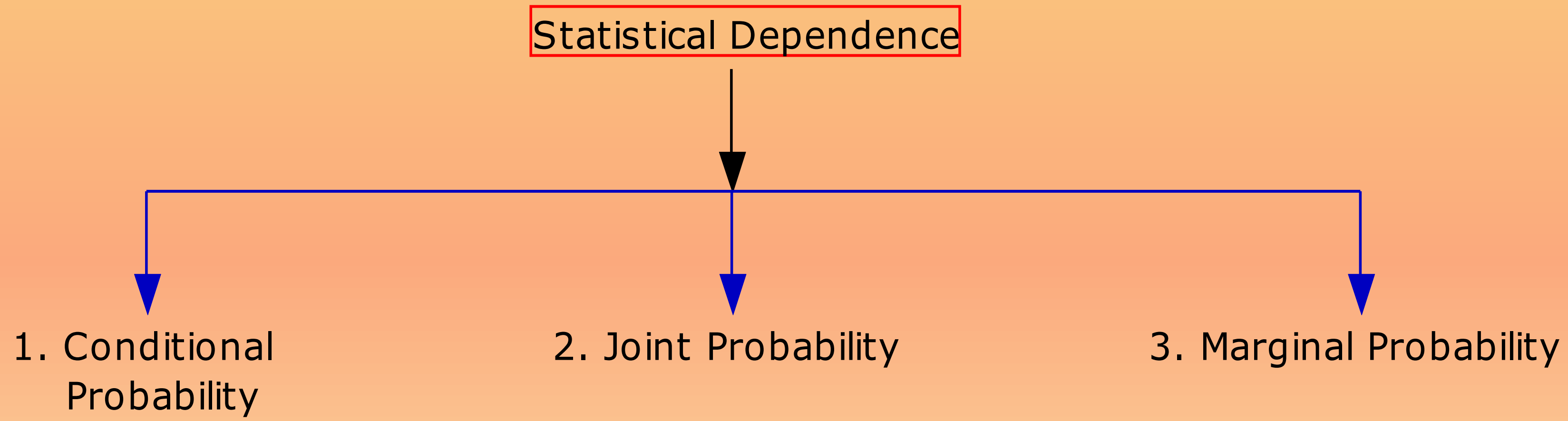
Let H_2 = probability of a head on the second toss,

H_1 = probability of a head on the first toss

The probability of a head on the second toss, given that the first toss resulted in a head, can be written as $P(H_2|H_1) = P(H_2) = 0.5$

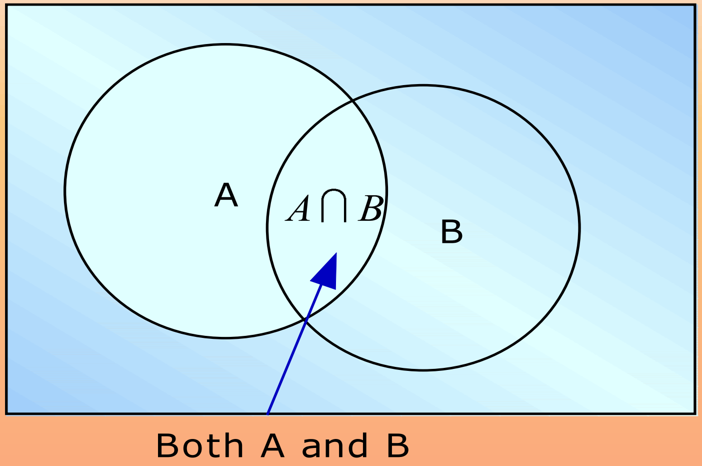
Probabilities under Statistical Dependence

When two or more events occur, the occurrence of one event has an effect on the probability of the occurrence of any other event. In this case, the events are considered to be dependent. The probabilities when the events are dependent are shown below.



1. Conditional Probability under Statistical Dependence

The probabilities discussed so far relate to the entire sample space. Sometimes we are interested in evaluating the probability of events where the event is conditioned on some part or subset of the sample space. Consider the following example:



Suppose there is a group of 100 people, out of which 40 are college graduates, 30 are businessmen, and 15 are both college graduates and businessmen.

If we define,

B = set of college graduates

A = set of businessmen

Then $A \cap B$ = set of college graduates and businessmen.

This is shown using the venn diagram. From the diagram the probabilities on the right can be calculated.

$$P(A) = \frac{30}{100} = 0.30 \qquad P(B) = \frac{40}{100} = 0.40$$

$$P(A \cap B) = \frac{15}{100} = 0.15$$

The above probabilities are calculated using the entire sample space.

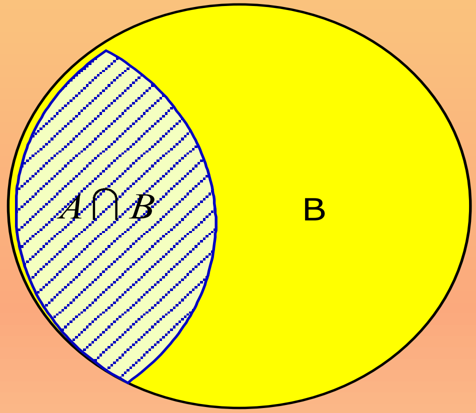
Now, suppose we select a person from the group who we know is a college graduate. What is the probability that the person is a businessman?

The probability that we want to calculate is a conditional probability and is written as

$$P(A|B)$$

The above probability statement is read as – probability of A given that B is known.

Since we know that the person belongs to the college graduate group, so we need to look only in the college graduate group. In other words, the sample space is now reduced (see Figure on the right). Note that A is the set of businessmen and B is the set of college graduates.



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

The sample space is reduced in which only college graduates are considered and the probability is given by

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.15}{0.40} = 0.375$$

Example: ***A bag contains 50 balls, of which 15 are red and dotted, 5 are red and striped, 10 are green and dotted, and 20 are green and striped. Suppose a red ball is drawn from the box. What is the probability that the ball is dotted? What is the probability that it is striped?***

For this type of problem, it is easier to calculate the probabilities if we construct a ***joint probability table***. We will first show how to use the given information to construct a joint probability table and then calculate the required probabilities. First, put the information in a joint probability table.

	Dotted (D)	Striped (S)	Marginal Probability Totals
Red (R)	15	5	20
Green (G)	10	20	30
Marginal Probability Totals	25	25	50

From the above table, we can calculate the following relative frequency probabilities:

P (R and D) = 15/50 = 0.30; P (R and S) = 5/50 =0.10

P (G and D) = 10/50 = 0.20; P (G and S) = 20/50 = 0.40

Note that the above probabilities are ***joint*** probabilities. Joint probabilities are ***"and"*** probabilities. Similarly, we can calculate the simple or marginal probabilities as

P (R) = 20/50 = 0.40; P (G) = 30/50 = 0.60

P(D) = 25/50 = 0.50; P (S) = 25/50 = 0.50

Now, we can use the above probabilities to construct a ***joint probability table***. The table is shown below.

	Dotted (D)	Striped (S)	Total
Red (R)	0.30	0.10	0.40
Green (G)	0.20	0.40	0.60
Total	0.50	0.50	**1.00**

This table contains the joint ("and") probabilities and also the marginal or simple probabilities. We want to find the following probabilities:

[a] Suppose a red ball is drawn from the box, what is the probability that the ball is dotted?

This is a conditional probability because we know that the ball drawn is a red one. The required probability can be written as P(D|R).

$$P((D|R) = \frac{P(DandR)}{P(R)} = \frac{0.30}{0.40} = 0.75$$

[b] Suppose a red ball is drawn from the box. What is the probability the ball is striped?

$$P(S|R) = \frac{P(SandR)}{P(R)} = \frac{0.10}{0.40} = 0.25$$

[c] Try calculating the following probabilities, using the information in the joint probability table .

$P(G|D)$ $P(R|D)$ $P(R|S)$

Example: Let A and B be two events, such that P(A) = 0.7, P(B)= 0.3 , and P (A and B) = 0.1. Find the following probabilities.

$P(A|B)$ $P(B|A)$.

$$P(A|B) = \frac{P(AandB)}{P(B)} = \frac{0.1}{0.3} = 0.333$$

$$P(B|A) = \frac{P(BandA)}{P(A)} = \frac{0.1}{0.7} = 0.1429$$

2. Joint Probability under Statistical Dependence

The joint probability under statistical dependence can be calculated using the formula of conditional probability. Refer to the following conditional probability formulas, discussed under conditional probability

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

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

The first equation is ***the probability of event A when B has already occurred*** and the second equation is ***the probability of event B when A has already occurred***.

The joint probabilities can be calculated from the above equations as:

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

or

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

and

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

or,

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

Example:

In a certain manufacturing plant, 40% of the workers are skilled, 70% of the workers are full time, and 90% of the skilled workers are full time. If a full time worker is selected at random, what is the probability that he or she is a skilled full time employee?

Let S = skilled, F = full time, then P (S) =0.40, P(F) = 0.70, $P(F|S) = 0.90$

We want to find P (S and F) or P (S and F). This probability can be calculated as

$$P(F \cap S) = P(F|S)P(S) = (0.90)(0.40) = 0.36$$

3. Marginal Probability under Statistical Dependence

The marginal probability is the probability of occurrence for a single event. Under statistical dependence, the calculation of marginal probability can be demonstrated using the joint probability table shown below. This problem was discussed under conditional probability. The problem and the joint probability table are reproduced below.

	Dotted (D)	Striped (S)	Marginal Probability Totals
Red (R)	15	5	20
Green (G)	10	20	30
Marginal Probability Totals	25	25	50

Joint- probability table for the above problem is shown on the next slide

	Dotted (D)	Striped (S)	Total
Red (R)	0.30	0.10	0.40
Green (G)	0.20	0.40	0.60
Total	0.50	0.50	**1.00**

Suppose we want to calculate the probability of a ball being red; that is, P (R). From the joint probability table, this probability is 0.40 and can be calculated by adding the two joint probabilities; P (D and R) plus P (S and R). Therefore, we can write

P(R) = P (D and R) + P (S and R)

The expressions on the right hand side of the above equation are "and" or "joint" probabilities under statistical dependence,

P(R) = P (D and R) + P (S and R) which is equal to

$$P(R) = P(D|R)P(R) + P(S|R)P(R)$$

The above equation is the expression for marginal probability under statistical dependence. This formula is useful in revising probabilities.

Probability Theory

Probability Terms

Counting Rules in Probability

Some Important Terms in Probability

Probability: *Probability is the chance that a particular event will occur when an experiment is performed. The probability of an event A is denoted as P (A), which means "the probability that event A occurs" is between 0 and 1. That is,*

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

Event: *An event is one or more possible outcomes of an experiment.*

Experiment: *An experiment is any process that produces an outcome or observation.*

Sample Space: *The set of all possible outcomes of an experiment is called thesample space and is denoted by S.*

Mutually Exclusive Events: *When the occurrence of one event excludes the possibility of another event occurring, then we say the events are mutually exclusive. In other words, only one event can take place at a time.*

Exhaustive Events: *The total number of possible outcomes in any trial is known as exhaustive events.*

Equally Likely Events: *A situation where all the events have an equal chance of occurrence or when there is no reason to expect one in preference to the other.*

(1) Multiple-Step Experiment or Filling Slots

Suppose an experiment can be described as a sequence of k steps in which

n1 = the number of possible outcomes on the first step

n2 = the number of possible outcomes on the second step

:

n_k = the number of possible outcomes on the k^{th} step, then the total number of possible outcomes is given by

$$(n_1)(n_2)(n_3)\ldots\ldots(n_k)$$

(2) Permutations

The number of ways of selecting n distinct objects from a group of N objects—where the order of selection is important—is known as the number of permutations on N objects, using n at a time and is written as

$$P_n^N = \frac{N!}{(N-n)!} = (n)(n-1)\ldots\ldots(n-k+1)$$

(3) Combinations

Combination is selecting n objects from a total of N objects. The order of selection is not important in combination. This disregard of arrangement makes the combination different from the permutation. In general, an experiment will have more permutations than combinations.

The number of combinations of N objects taken n at a time is given by

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

Note $0! = 1$

Chapter 4: Probability Theory - Flow Diagram (1)

Probability Theory...continued

Ways of Assigning Probabilities

Assigning Probabilities: *There are two basic rules for probability assignment:*

The probability of an event A is written as P(A) and it must be between 0 and 1. That is,

$$0 \le P(A) \le 1.0$$

If an experiment results in n number of outcomes $A_1, A_2 \ldots A_n$; then the sum of the probabilities for all the experimental outcomes must equal 1. That is,

$$P(A_1) + P(A_2) + P(A_3) + \ldots\ldots + P(A_n) = 1$$

Methods of Calculating Probabilities: *There are three methods for assigning probabilities*

1. ***Classical Method***
2. ***Relative Frequency Approach***
3. ***Subjective Approach***

Probabilities for Mutually and Non-mutually Exclusive Events

Addition Law for Mutually Exclusive Events

If we have two events A and B, that are mutually exclusive, then the probability that A or B will occur is given by

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

Note that the "union" sign is used for "or" probability; that is, $P(A \cup B)$. This is same as P (A or B). This rule can be extended to three or more mutually exclusive events. If three events A, B, and C are mutually exclusive then the probability that A or B or C will occur

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

Different Ways of Calculating Probability

1. Classical Method

The classical approach of probability is defined as the favorable number of outcomes divided by the total number of possible outcomes. Suppose an experiment has n number of possible outcomes and the event A occurs in m of the n outcomes, then the probability that event A will occur is

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

Note that P(A) denotes the probability of occurrence for event A. The probability that the event A will not occur is given by $P(\bar{A})$, which is read as P (not A) or 'A complement.' Thus,

$$P(A) + P(\bar{A}) = 1$$

which means that the probability that event A will occur, plus the probability that event A will not occur, must be equal to 1.

2. Relative Frequency Approach

Probabilities are also calculated using the relative frequency. In many problems, we define probability by relative frequency.

3. Subjective Probability

Subjective probability is used when the events occur only once or very few times and when little or no relevant data are available. In assigning subjective probability, we may use any information available, such as our experience, intuition, or expert opinion.

Addition Law for Non-Mutually Exclusive Events

The occurrence of two events that are ***non-mutually exclusive*** *means that they can occur together. If the events A and B are non-mutually exclusive, the probability that A or B will occur is given by*

$$P(A \cup B) = P(A) + P(B) - P(A \text{ and } B)$$
$$\text{or, } P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

If events A, B, and C are non-mutually exclusive, then the probability that A or B or C will occur:

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \text{ and } B) - P(A \text{ and } C) - P(B \text{ and } C) + P(A \text{ and } B \text{ and } C)$$

or,

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

Chapter 4: Probability Theory - Flow Diagram (2)

Probability Theory...continued

Probabilities when the Events are Independent

Simple/Marginal or Unconditional Probability

Simple Probability- *is also known as marginal or unconditional and is the probability of occurrence for a single event, say A, and is denoted by P (A).*

P(A) = marginal probability of event A; P(B) = marginal probability of event B

Joint Probability

Joint Probability under Statistical Independence:

Joint probability is the probability of occurrence for two or more events together or in succession. It is also known as 'and' probability. Suppose we have two events, A and B, which are independent. Then the joint probability, P (AB), which is the probability of occurrence of both A 'and' B, is given by

$$P(AB) = P(A).P(B)$$

or, $$P(A \cap B) = P(A).P(B)$$

Note that P (AB) = probability of event A and B occurring together — is known as joint probability. P(AB) is the same as P (A and B) or $P(A \cap B)$

Conditional Probability under independence

Conditional Probability under Statistical Independence

The conditional probability is written as

$$P(A|B)$$

and is read as the probability of event A, given that B has occurred, or the probability of A, given B. If the two events A and B are independent, then

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

This means that if the events are independent, the probabilities are not affected by the occurrence of each other.

Chapter 4: Probability Theory - Flow Diagram (3)

Probability Theory...continued

Probabilities when the Events are Dependent

Conditional Probability

Conditional probability under Statistical Dependence

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

or,

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

Joint Probability

Joint probability under Statistical Dependence

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

or

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

or,

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

or

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

Marginal Probability

Marginal probability under Statistical Dependence

The marginal probability under statistical dependence can explained using the joint probability table below. Consider the four events: D, S, R, and G and the probabilities below.

	(D)	(S)	Total
(R)	0.30	0.10	0.40
(G)	0.20	0.40	0.60
Total	0.50	0.50	1.00

P(R) = P(D and R) + P(S and R)

$$P(R) = P(D|R)P(R) + P(S|R)P(R)$$

Bay's Theorem

$$P(A_i|D) = \frac{P(A_i)P(D|A_i)}{P(A_1)P(D|A_1) + P(A_2)P(D|A_2) + \ldots + P(A_n)P(D|A_n)}$$

This equation can be used to compute any ***posterior probability*** $P(A_i|D)$ *when prior probabilities* $P(A_1), P(A_2)...., P(A_n)$ *and conditional probabilities* $P(D|A_1), P(D|A_2)...., P(D|A_n)$ *are known.*

Chapter 4: Probability Theory - Flow Diagram (4)