Probability

Permutation → order of items is important

- In mathematics, permutation relates to the act of arranging all the members of a set into some sequence or order
- Different arrangements of a given number of elements taken one by one, or some, or all at a time
- E.g.
  - if we have two elements A and B, then there are two possible arrangements, AB and BA

$n = 2$
$r = 2$

n = total no of items
r = no of items selected

A B C D

[AB CD AC AD BA DC . . . . .]

arrangements

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

$^n_rP = {}^nP_r$

$1! = 1$
$2! = 2 \times 1 = 2$
$3! = 3 \times 2 \times 1 = 6$
$4! = 4 \times \underbrace{3 \times 2 \times 1}_{3!} = 24$
$5! = 5 \times 4!$
$= 5 \times 24$
$5! = 120$

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

Permutation Example

DEL LED

- How many words can be formed by using 3 letters from the word "DELHI" ?

total no of letters = 5, no of selected letters = 3

$n = 5$ $r = 3$

$$^{n}P_{r} = \frac{n!}{(n-r)!} = \frac{5!}{(5-3)!} = \frac{5 \times 4 \times 3 \times \cancel{2!}}{\cancel{2!}}$$

$$\boxed{^{n}P_{r} = 60}$$

Combination : order is NOT important

- The combination is a way of selecting items from a collection, such that (unlike permutations) the order of selection does not matter
- Different selections of a given number of elements taken one by one, or some, or all at a time
- E.g.
  - if we have two elements A and B, then there is only one way select two items, we select both of them

$$ {}^{n}_{r}C = \frac{n!}{r!\,(n-r)!} $$

AB = BA

1

$$ {}^{n}_{r}C = {}^{n}C_{r} = \binom{n}{r} $$

Combination Example

- In how many ways, can we select a team of 4 students from a given choice of 15 ?

total no of students $(n) = 15$

no of students to be selected $(r) = 4$

$$ {}^{n}_{r}C = \frac{n!}{r!\,(n-r)!} = \frac{15!}{4! \times 11!} = \frac{15 \times \overset{7}{\cancel{14}} \times 13 \times \overset{3}{\cancel{12}} \times \cancel{11!}}{\cancel{4} \times \cancel{3} \times \cancel{2} \times 1 \times \cancel{11!}} $$

$$ \boxed{\text{combinations} = 1365} $$

Introduction

probability score [p-value]

- Probability has become the basis for statistical applications in the mathematical theory
- In fact, it has become a part of our everyday life
- In personal and management decisions, we face uncertainty and use probability theory, whether or not we admit the use of something so sophisticated
- To quote Levin, we live in a the world in which we are unable to forecast the future with complete certainty
- Our need to cope with uncertainty leads us to the study and use of probability theory
- Probability theory in fact is the foundation of statistical inference

Probability

Favourable outcome

- Probability of a given event is an expression of likelihood or chance of occurrence of an event
- It is a branch of mathematics that deals with calculating the likelihood of a given event's occurrence
- Expressed as a number between 1 and 0
  - An event with a probability of 1 can be considered a certainty
  - An event with a probability of 0 can be considered a uncertainty
- Probability has been introduced in Maths to predict how likely events are to happen
- It is explained by four different ways
  - Classical or Priori Probability * * * *
  - Relative Frequency Theory of Probability
  - Subjective Approach to Probability
  - Axiomatic Approach to Proability

something is already known

Classical or a Priori Probability

- It is the oldest and simplest approach of all
- It originated in the eighteenth century in solving the problems pertaining to the games of chances like throwing of coins, dice or deck of cards etc.
- The basic assumption underlying the classical theory is that the outcomes of a random experiment are "equally likely"
- The Experiment refers to processes which result in different possible outcomes or observations
- The "equally likely" conveys the notion that each outcome of an experiment has same chance of appearing as any other
- Thus, in a throw of a dice, occurrence of 1, 2, 3, 4, 5 and 6 are equally likely events

Probability

- The definition given by French Mathematician Laplace is adopted by the classical school which says
  - Probability is the ratio of number of "favorable" cases to the total number of equally likely cases

$$P(A) = \frac{number\ of\ favorable\ cases}{total\ number\ of\ equally\ likely\ cases}$$

total no of cases = [H,T]

- To calculate probability we need to calculate
  - Number of favorable cases
  - Total number of equally likely cases

$p(H) = \frac{1}{2} = 0.5 = 50\%$

total no of outcomes = [1,2,3,4,5,6]

$P(4) = \frac{1}{6} = 0.16 = 16\%$

- E.g.
  - If a coin is tossed, there are two equally likely results, a head or a tail. Hence the probability of head becomes 1 / 2, which is same as getting a tail
  - If a dice is thrown, the probability of obtaining an even number is 3/6 or ½ since there are six faces which means there are 6 equally likely cases

Shortcomings of Classical approach

- The definition can be applied whenever it is not possible to make a simple enumeration cases which can be considered equally likely
  - E.g. how does it apply to probability of rain ?

Mike

- It fails to answer questions like “what is the probability that a ~~make~~ will die before the age of 60 ?” as the outcomes are not equally likely

classical probability depends of total No of outcomes.

Relative Frequency Theory

- In 1800, British statisticians, were interested in theoretical foundation for calculating risk of losses in life insurance and commercial insurance
- They began defining probabilities from statistical data collected on births and deaths
- Today this approach is known as relative frequency of occurrence
- The probability of an event cam be defined as relative frequency with which it occurs in an indefinitely large number of trials
- If an event occurs a out of n, its relative frequency is a/n, the value which is approached by a/n when n becomes infinity is called the limit of the relative frequency

$$P(A) = \lim_{n \to \infty} \frac{a}{n}$$

Relative Frequency Theory

- The two approaches classical and relative though seemingly same, differ widely
- Relative Frequency Theory though useful in practice, has difficulties from a mathematical point of view, since an actual limiting number may not really exist
- Quite often people use this approach without evaluating a sufficient number of occurrences
- It may be pointed out at the very outset that probability should not be understood in the sense of certainty
  - E.g. when we say that probability of getting head or tail is 50%, it does not mean that if the coin is tossed 16 times, we must get 8 heads and 8 tails
  - This means, when a coin is tossed large number of times (n increases), we will usually get closer to 50% heads
- The probability calculated by following relative frequency definition is called as posteriori or empirical probability as distinguished from a priori probability obtained by following classical approach

Subjective Approach to Probability

- Subjective approach to assigning probabilities was introduced in year 1926 by Frank Ramsey in his book "The Foundation of Mathematics and other Logical Essays"
- It is defined as the probability assigned to an event by an individual based on whatever evidence is available at the given time
- Hence such probabilities are based on the beliefs of the person making the probability statement
- E.g. if a teacher wants to find out probability of X student topping in Y subject, he may assigned a value between zero to one according to his degree of belief for possible occurrence
- He may take into account factors like past academic performance, views of his colleagues etc.
- This approach is highly broad and very flexible

Axiomatic Approach to Probability

- It was introduced by Russian mathematician Andrey in year 1933
- He mentioned the theory of probability in his book "Foundation of Theory of Probability"
- When this approach is followed, no precise definition of probability is given
- Rather we give certain axioms or postulates on which probability calculations are based
- The whole field of probability theory is based on following axioms
  - The probability of an event ranges from zero to one
    - If event can not take place, its probability is zero
    - If it is certain, its probability is one
  - The probability of entire sample space P(S) = 1
  - If A and B are mutually exclusive or disjoint events the probability of occurrence of either A or B is denoted by A U B and given by
    - A U B = P(A) + P(B)

Definitions

- **Sample Space** (S)
  - The set of all the possible outcomes to occur in any trial
  - E.g.:
    - Tossing a coin, Sample Space (S) = {H,T}
    - Rolling a die, Sample Space (S) = {1,2,3,4,5,6}

sample space is made up of all possible sample point

- **Sample Point** →
  - It is one of the possible results
  - E.g.
    - In a deck of Cards:
      - 4 of hearts is a sample point
      - The queen of clubs is a sample point

- **Experiment or Trial** → Action to get the outcome
  - Describes an act which can be repeated under some given conditions
  - Random experiments are those whose results depend on chance
  - E.g.:
    - The tossing of a coin
    - Selecting a card from a deck of cards

Definitions

- **Event**
  - It is a single outcome of an experiment
  - Also referred as chance
  - E.g.
    - Getting a Heads while tossing a coin is an event
- **Exhaustive Events**
  - Events are said to exhaustive when they totality includes all the possible outcomes of experiments
  - E.g.
    - For throwing a dice, the possible outcomes are 1, 2, 3, 4, 5, 6
    - Hence the exhaustive number of cases is 6
- **Complimentary event** – opposite or negative
  - The non-happening events
  - The complement of an event A is the event, not A (or A')
  - E.g.
    - Standard 52-card deck, A = Draw a heart, then A' = Don't draw a heart

Definitions

- Impossible Event
  - The event cannot happen
  - E.g.
    - In tossing a coin, impossible to get both head and tail at the same time
- Mutually exclusive events
  - Two events are said to be mutually exclusive or incompatible when both can not happen simultaneously
  - E.g.
    - In tossing coin, you could get only either of head or tail not both at the same time
- Equally likely events
  - Events are said to be equally likely when one does not occur more often than the others
  - E.g.
    - In tossing coin, each side to be observed approximately the same number of times in the long run

Probability Example 1

- A coin is tossed one time. What is the probability that it will Head?

sample space (s) = [H, T]

$$p(H) = \frac{\text{favorable outcomes}}{\text{total outcomes}} = \frac{1}{2}$$

$p(H) = 0.5$ or $1/2$ or 50%

decimal (0 to 1) | ration | percentable

Probability Example 2

$S = [HH, HT, TH, TT]$

two coins are tossed

- A ~~coin is tossed two times.~~ What is the probability of getting
  - Two heads
  - Two tails
  - One head
  - No head
  - At least one head
  - At most one head

① $P(HH) = \dfrac{\text{No of times HH appeared}}{\text{total no of outcomes}} = \dfrac{1}{4} = 0.25 = 25\%$

② $P(TT) = \dfrac{1}{4} = 0.25 = 25\%$

③ $P(H) = \dfrac{2}{4} = 1/2 = 0.50 = 50\%$

④ $P(\text{No Head}) = 1/4 = 0.25 = 25\%$

① two heads = Exactly two heads

② at least one head = one head + two heads

③ at most two heads = 0 head + 1 head + 2 heads

⑤ $P(\text{at least one head}) = P(H) + P(HH) = \dfrac{3}{4} = 0.75 = 75\%$

⑥ $P(\text{at most one Head}) = P(0H) + P(H) = 3/4 = 0.75 = 75\%$

Probability Example 3

three coins are tossed

- A ~~coin is tossed three times~~. What is the probability of getting
  - Two heads
  - Two tails
  - One head
  - No head
  - At least one head
  - At most one head

$p(HH) = 3/8$

$p(TT) = 3/8$

$p(H) = 3/8$

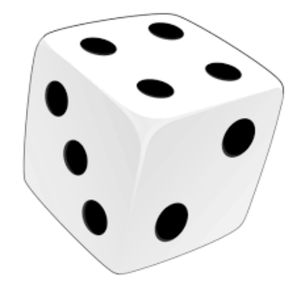
$p(0H) = 1/8$

$p(\text{at least one head}) = 7/8$

$p(\text{at most one head}) = 4/8$

$S = [HHH, HHT, HTT, THH, THT, TTT, TTH, HTH]$

Probability Example 4



- A dice is rolled one time. What is the probability of getting
  - 5
  - Number greater than 3
  - One even number
  - One prime number

$S = [1, 2, 3, 4, 5, 6]$

① $p(5) = 1/6$

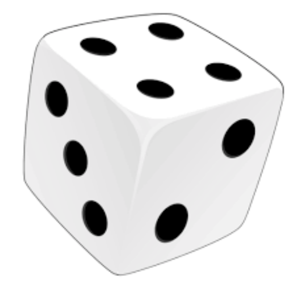
② $p(N > 3) = 3/6$

③ $p(\text{even no}) = 3/6$

④ $p(\text{prime no}) = 3/6$

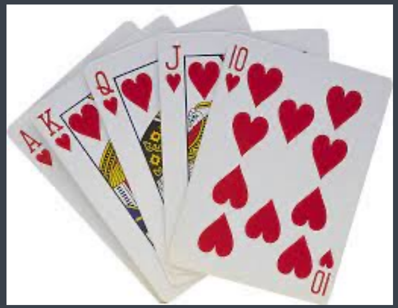
Probability Example 5

- A dice is rolled two times. What is the probability of getting
  - 6 on first time
  - Sum if 10 or more
  - Difference in numbers is equal to 3
  - Both are the even numbers
  - Both are 5



Probability Example 6

- A card is drawn from a pack of cards. What is the probability of getting
  - Red color
  - Ace of black
  - Jack of red
  - Any king
  - Dimond and Jack
  - Spade



Probability formula

- $P(A`) = 1 - P(A)$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- $P(A \cap B) = P(A) . P(B)$
- $P(A . A`) = 0$
- $P(A` . B) = P(B) - P(A.B)$
- $P(A . B`) = P(A) - P(A.B)$
- $P(A + B) = P(A` . B) + P(A . B`) + P(A . B)$

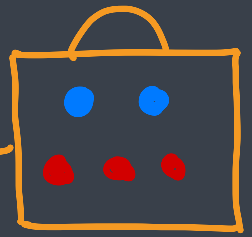
Conditional Probability

Independent Events

- Events can be independent, meaning each event is **not affected** by any other events
- Example:
  - Tossing a coin
  - Each toss of a coin is a perfect isolated thing
  - What it did in the past will not affect the current toss
  - The chance is simply 1-in-2, or 50%, just like ANY toss of the coin
  - So each toss is an **Independent Event**

Dependent Events

- Events can be ~~in~~dependent, meaning each event is **affected** by any other events
- Example
  - Consider there are 2 blue and 3 red marbles are in a bag
  - What are the chances of getting a blue marble?
  - What are the chances of getting a blue marble again?



$p(\text{Red}) = \frac{3}{5}$

with replacement

$p(\text{Red}) = 3/5$

simple probability

without replacement

$p(\text{Red}) = 2/4$

conditional probability

Conditional Probability

- The likelihood of an event occurring, assuming a different one has already happened
- The formula is

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

Example 1

- What are the chances of drawing 2 blue marbles from a bag of 2 blue and 3 red marbles?

Example 2

- Drawing 2 Kings from a Deck

Baye's Rule

- Bayes' theorem, named after 18th-century British mathematician Thomas Bayes, is a mathematical formula for determining conditional probability
- It provides a way to revise existing predictions or theories (update probabilities) given new or additional evidence
- Bayes' theorem is also called Bayes' Rule or Bayes' Law and is the foundation of the field of Bayesian statistics

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

where:

$P(A) =$ The probability of A occurring

$P(B) =$ The probability of B occurring

$P(A|B) =$ The probability of A given B

$P(B|A) =$ The probability of B given A

$P\left(A \bigcap B\right)) =$ The probability of both A and B occurring

How does it work ?

- Below is a data set of weather and corresponding target variable 'Play' (suggesting possibilities of playing). We need to classify whether players will play or not based on weather condition.

Weather	Play
Sunny	No
Overcast	Yes
Rainy	Yes
Sunny	Yes
Sunny	Yes
Overcast	Yes
Rainy	No
Rainy	No
Sunny	Yes
Rainy	Yes
Sunny	No
Overcast	Yes
Overcast	Yes
Rainy	No

How does it work ?

- Step 1: Convert the data set into a frequency table

$p(y) = 9/14$

simple

Frequency Table		
Weather	No	Yes
Overcast		4
Rainy	3	2
Sunny	2	3
Grand Total	5	9

$$\begin{array}{r} 4 \\ 5 \\ 5 \\ \hline 14 \end{array}$$

How does it work ?

- Step 2: Create Likelihood table by finding the probabilities like Overcast probability = 0.29 and probability of playing is 0.64

P(weather)

Likelihood table				
Weather	No	Yes		
Overcast	0	4	=4/14	0.29
Rainy	3	2	=5/14	0.36
Sunny	2	3	=5/14	0.36
All	5	9		
	=5/14	=9/14		
	0.36	0.64		

-P (playing game)

How does it work ?

- **Problem:** Players will play if weather is sunny. Is this statement is correct?
- We can solve it using above discussed method of posterior probability.

given

$$P(Yes \mid Sunny) = P(Sunny \mid Yes) * P(Yes) / P(Sunny)$$

$$P(A \mid B) = \frac{P(A)\, P(B \mid A)}{P(B)}$$

- Here we have

A = Yes

B = Sunny

$P(Sunny \mid Yes) = 3/9 = 0.33$
$P(Sunny) = 5/14 = 0.36$
$P(Yes) = 9/14 = 0.64$

- Which means, $P(Yes \mid Sunny) = 0.33 * 0.64 / 0.36 = 0.60$, which has higher probability.