INTRODUCTION TO PROBABILITY

Whenever we use mathematics in order to study some observational phenomena we build the mathematical model for the phenomena. There are two types of mathematical models:

1) Deterministic models
2) Non-deterministic or Probabilistic or Stochastic or Random models.

Deterministic Models:

A model which stipulates that the condition under which an experiment is performed, determine the outcome of the experiment. It is an experiment in which the outcomes can be predicted with certainty.

Ex: 1. If a particle is released with initial velocity u and acceleration a, several times, then in each release the distance traveled by the particle in time t is $ut+(1/2)at^2$

2. If we insert a battery into a simple circuit the mathematical model which gives the flow of current is $I=\frac{V}{R}$, Ohms Law. If the experiment is repeated a number of times using the same circuit i.e. keeping V & R fixed, the value of I would remain the same.

Non-deterministic Models:

Suppose we have an experiment such that the collection of all possible outcomes of the experiment is the same whenever it is conducted under identical condition, but its outcome cannot be predicted with certainty in any treat of the experiment is the experiment require a different mathematical model for their investigation.

Ex: 1) Tossing a coin
2) Drawing a card at random from a pack of cards.

Random or Stochastic Experiment:

The real-world phenomena which involve randomness and for which the non-deterministic models are appropriate is referred to as a *Random or Stochastic Experiment*. The following are the features of a random experiment.

1) Each experiment is capable of being repeated indefinitely under essentially the same condition.
2) Though it will not be possible to predict a particular outcome, it is possible to describe the 'set of all possible outcomes' of the experiment.
3) As the experiment is performed repeatedly the individual outcomes seem to occur in a haphazard manner. However, as the experiment is repeated a large number of times, a definite pattern or regularity appears.

Note: 1) Each performance in a random experiment is called a **Trial.**

2) All the trails are conducted under the same set of conditions in a random experiment.

3) The result of a trial in a random experiment is called an **Outcome.**

Ex: 1) Toss a die and observe the number that shows on top.

2) Toss a coin 4 times and observe the total number of heads obtained.

3) A tube light is tested for its life length by recording the time elapsed until it burns out.

4) From an urn containing only black balls, a ball is chosen and its colour is noted.

Sample Space:

The totality of all possible outcomes of a random experiment E is called a **sample space** S. A sample space S is said to be finite, if the no. of elements in S is finite.

Ex: 1) S={1,2,3,4,5,6}

2) S={0,1,2,3,4}

3) S={t: t>=0}

4) S={black balls}

Events:

Every subset A of S which is a disjoint union of a single element subsets of the sample space S of a random experiment E is called an **event** i.e. an event A is simply a set of possible outcomes.

Ex: 1) A={2,4,6} is an even no. occurs.

2) A={2} is 2 heads occur.

3) A={t/t<3} is tube glows less than 3 hrs.

Mutually Exclusive Events:

Two events A & B are said to be mutually exclusive if they cannot occur together i.e. $A \cap B = \phi$.

Ex: An electronic device is tested and its total time of service, t, is recorded. Let S={t/t>=0}. Consider the events A={t/t<100} B={t/50<=t<=200} C={t/t>150}. Then

$A \cup B = \{t/t \leq 200\}$

$A \cap B = \{t/50 \leq t \leq 100\}$

$B \cup C = \{t/t \geq 50\}$

$B \cap C = \{t/150 \leq t \leq 200\}$

$A \cap C = \phi$ $A \cup C = \{t/t<100 \text{ or } t>150\}$

$\overline{A} = \{t/t \geq 100\}$

$\overline{C} = \{t/t \leq 150\}$

Equally Likely Events:

Two or more events are said to be equally likely or equiprobable if they have equal chance of occurrence i.e. there is no reason to expect one in preference to the other.

Independent Events:

Two events are said to be independent if the occurrence of one event does not effect the occurrence or the non-occurrence of the other.

Classical Definition of Probability: (Due to **Laplace**)

If a trial results in n mutually exclusive and equally likely events or outcomes and m of them are favorable to the happening of an event A, then, the probability of happening of A is

$$P(A) = \frac{\text{Favourable no.of cases}}{\text{Total no.of cases}} = \frac{m}{n}$$

Ex: 1) Probability of head appearing when a coin is tossed is P(A)=1/2.

Note: 1) If m=n, then P(A)=1 i.e the event A is a certain event.
2) If m=0, then P(A)=0 i.e. A is an impossible event.

Limitations: It fails when the various outcomes of the trail are not equally likely or equally probable.

Relative Frequency: (Due to **Richard Von Mises)**

Let an experiment E be repeated n times essentially under the same conditions and let A be the event associated with E. Let m be the No. of times the event A occurs. Then f=m/n is called the relative frequency of the event A in the n repetitions of E.

It can be shown that, $f \to P(A)$ as $n \to \infty$. However, this definition has its own limitations

Geometrical Definition

The probability of an event is the status of the area favorable for an event of the total area of events. (It is the generalization of the classical definition). This definition also has its own limitations.

Axiomatic Definition of Probability: (Due to **A. N Kolmogorov, 1933**).

Let E be an experiment. Given a sample space S associated with E, **Probability** is a function which assigns a non-negative real number to every event A denoted by P(A), and is called the probability of the event A, satisfying the following axioms:

i) $0 \leq P(A) \leq 1$
ii) $P(S)=1$
iii) If A & B are mutually exclusive events, then $P(A \cup B)=P(A)+P(B)$

Note: If $A_1, A_2, \ldots, A_n \ldots$ are pairwise mutually exclusive events then

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

Theorem1:

If ϕ is the Null event, then $P(\phi)=0$

Proof: For any event A, $A=A\cup\phi$, where A and ϕ are mutually exclusive.

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

$\Rightarrow P(A)=P(A)+P(\phi)$

$\Rightarrow P(\phi)=0$

Theorem2:

If $\bar{A}$ is the complementary event of A, then $P(A) = 1 - P(\bar{A})$

Proof:

If S is the Sample Space, with A and $\bar{A}$ being disjoint events then

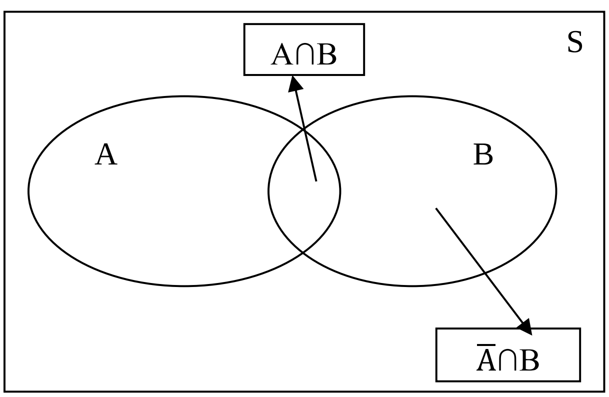
$S = A\cup\bar{A}$

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

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

Cor: If A & B are any two events, then $P(\bar{A}\cap B)=P(B)-P(A\cap B)$

Proof:



Since $\bar{A}\cap B$ and $A\cap B$ are disjoint events, we can write

$B=(A\cap B)\cup(\bar{A}\cap B)$

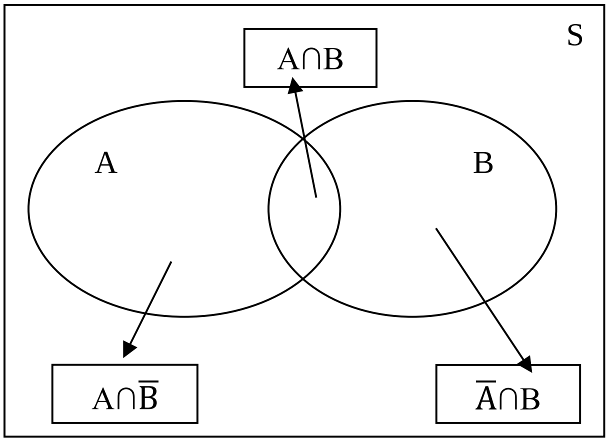
$P(B)=P(A\cap B)+P(\bar{A}\cap B)$

Or $P(\bar{A}\cap B)=P(B)-P(A\cap B)$

Theorem 3: (Addition Theorem)

If A and B are any 2 events, then $P(A \cup B)=P(A)+P(B)-P(A\cap B)$

Proof:



We can write

$A \cup B=A\cup(B\cap\bar{A})$

$P(A\cup B)=P(A)+P(B\cap\bar{A}) \longrightarrow 1$

$B=(A\cap B)\cup(B\cap\bar{A})$

$P(B)=P(A\cap B)+P(B\cap\bar{A}) \longrightarrow 2$

Subtracting 2 from 1, we get

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

Aliter:

Let $A\cup B=A\cup(B\cap\bar{A})$

$P(A\cup B)=P(A)+P(B\cap\bar{A})$

Add and Subtract $P(A\cap B)$, we get

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

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

Theorem 4:

For any three events A, B and C, prove that

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

Prove it !

Theorem 5:

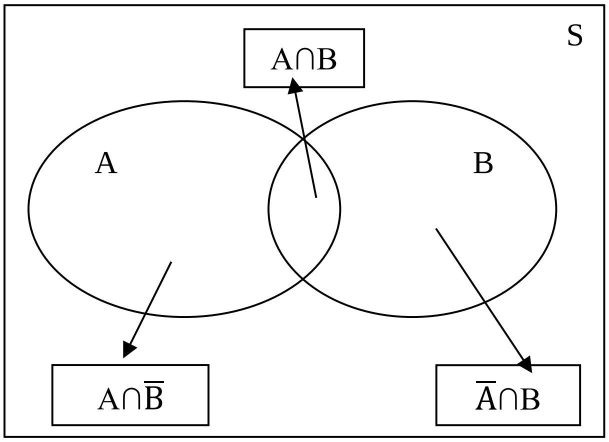
If $A\subset B$ then $P(A)\leq P(B)$

Prove it !

Results:

1) Show that the probability (that exactly one of the events A or B occurs) is given by $P(A)+P(B)-2P(A\cap B)$

Proof:



For any two events A & B,

Probability (that exactly one of the events A or B occurs) = $P((A\cap\bar{B})U(B\cap\bar{A})]$

Since $A\cap\bar{B}$ and $B\cap\bar{A}$ are mutually exclusive, we have

$P((A\cap\bar{B})U(B\cap\bar{A})]=P(A)+P(B)-2P(A\cap B)$

$\Rightarrow$ $P((A\cap\bar{B})U(B\cap\bar{A})]= P(A\cap\bar{B})+P(B\cap\bar{A})$ --------------(a)

But $A\cap B$ is disjoint with both these sets and the union of the events $A\cap\bar{B}$ and $B\cap\bar{A}$ and $A\cap B$ is nothing but AUB.

Add & Subtract $P(A\cap B)$ in (a) above, we get,

$$
\begin{aligned}
P[(A\cap\bar{B})U\ (B\cap\bar{A})] &= P(A\cap\bar{B})+P(B\cap\bar{A})+ P(A\cap B)- P(A\cap B) \\
&= [P(A\cap\bar{B})+P(B\cap\bar{A})+ P(A\cap B)]- P(A\cap B) \\
&= P(AUB)-P(A\cap B) \\
&= P(A)+P(B)-P(A\cap B)-P(A\cap B)
\end{aligned}
$$

$\therefore$ The required probability = $P(A)+P(B)-2P(A\cap B)$

2) For any two events A & B, $P(A\cap B)\leq P(A)\leq P(AUB)\leq P(A)+P(B)$

Prove it !

Problems:

1) What is the probability that a leap year selected at random contain 53 Sundays?

Solution: A leap year has 366 days i.e. 52 complete weeks and 2 days over and above.
via (S,M),(M,T),(T, W),(W, Thu),(Thu, F),(F, Sat),(Sat, S)
Let event A={A leap year contains 53 Sundays}
Total No. of cases= n=7 and favorable No. of cases= m=2
Therefore, P(A)= m/n= 2/7.

2) Suppose that A, B & C are events, Such that P(A)= P(B)=P(C)=1/4, $P(A\cap B)=P(B\cap C)=0$ and $P(A\cap C)=1/8$. Find the probability that atleast one of the events A, B or C occurs.

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

$= 1/4+1/4+1/4-0-1/8-0+0$

$= (6-1)/8$

$= 5/8.$

3) Suppose that A & B are events. Given that P(A)=x, P(B)=y, $P(A\cap B)=z$.
Find, i) $P(\bar{A}\cup\bar{B})$ ii) $P(\bar{A}\cap B)$ iii) $P(\bar{A}\cup B)$ iv) $P(\bar{A}\cap\bar{B})$

Solution:

i) $P(\bar{A}\cup\bar{B})=P(\overline{(A\cap B)})=1-P(A\cap B)=1-z$

ii) $P(\bar{A}\cap B)=P(B)-P(A\cap B)= y-z$

iii) $P(\bar{A}\cup B)=P(\bar{A})+P(B)-P(\bar{A}\cap B)$

$=1-x+y-y+z$

$=1-x+z$

iv) $P(\bar{A}\cap\bar{B})=P(\overline{(A\cup B)})=1-P(A\cup B)=1-P(A)-P(B)+P(A\cap B)$

$=1-x-y+z$

4) A class contains 10 boys and 20 girls, of which half the boys and half the girls have brown eyes. A person is chosen at random from this group. What is the probability that the chosen person is a boy or has brown eyes?

Solution: Define the events A : {The person chosen is a boy}
B : {The person chosen has brown eyes}
$\therefore$ Required probability = $P(A\cup B)$ = ? = $P(A)+P(B)-P(A\cap B)$

Now $P(A) = \frac{\binom{10}{1}}{\binom{30}{1}} = \frac{10}{30} = \frac{1}{3}$

$P(B) = \frac{\binom{15}{1}}{\binom{30}{1}} = \frac{15}{30} = \frac{1}{2}$

$P(A \cap B) = \frac{\binom{5}{1}}{\binom{30}{1}} = \frac{5}{30} = \frac{1}{6}$

$\therefore P(A \cup B) = \frac{1}{3} + \frac{1}{2} - \frac{1}{6} = \frac{2}{3}$

5) A lot consists of 10 good articles, 4 with minor defects and 2 with major defects. An article is chosen at random. Find the probability that
   a) it has no defects
   b) it has no major defects
   c) it is either good or has major defects

   Solve it!

Finite Sample Space:

A sample space consisting of a finite or countably infinite number of elements is referred to as **Finite sample space.**

Ex: $S=\{a_1,a_2\ldots.a_k\}$

In order to characterize P(A), consider an event consisting of single outcome, say, $A=\{a_i\}$ we assign a number p_i is called the probability of $\{a_i\}$ satisfying

i) $p_i \geq 0$, $i=1,2,\ldots.k$

ii) $p_1+p_2+\ldots.+p_k=1$

Suppose that an event A consists of r outcomes, $1 \leq r \leq k$, so that $A=\{a_{j1},a_{j2},\ldots a_{jr}\}$ where j1,j2…jr are any indices from 1,2,3….k. Then $P(A)=p_{j1}+p_{j2}+\ldots+p_{jr}$.

Thus, by assigning probabilities p_i to each elementary event $\{a_i\}$ subject to condition (i) and (ii) above, one can uniquely determine P(A) for each $A \subset S$.

To evaluate p_i's and hence P(A), some assumptions such as equally likely outcomes concerning the individual outcome must be made.

Note 1: In most of the experiments we are concerned with choosing at random one or more objects from a given collection of objects. Suppose we have N objects say $a_1, a_2 \ldots a_N$.

a) **To choose one object at random from N objects** means each object has the same probability of selection. i.e. P(choosing any a_i)=1/N, i=1,2….N.

b) **To choose 2 objects at random from N objects** means each pair of objects has the same probability of being chosen as any other pair. Thus if there are k such pairs, then P(choosing any pair)=1/k.

c) **To choose n objects at random from N objects** means that each n-tuple, say (a_i, $a_2 \ldots a_n$) is as likely to be chosen as any other n tuple. If there are k such groups of n objects then P(choosing any group of n objects)=1/k.

Note 2: There are several ways in which samples may be selected from a population. Here we consider, for example:

i) the samples drawn sequentially (one after another)
ii) the samples drawn simultaneously (together)

Let Z denote the set of balls in the urn. If the balls are drawn sequentially then we may describe the outcome of the game by the ordered k-tuple, ($z_1, z_2 \ldots z_k$) of elements of Z, where z_1 denotes the first ball drawn, z_2 denotes the 2^{nd} and so on and k^{th} the total no. of balls drawn. Thus we shall refer to ($z_1, z_2 \ldots z_k$) as an *ordered sample* of size k.

If the balls are drawn simultaneously, it no longer makes sense to speak of a first ball or 2^{nd} ball etc., we may describe the outcome of our sampling only by the subset $\{z_1, z_2 \ldots z_k\}$ as an *unordered sample of size k.*

1) The number of ways of choosing an unordered sample of k objects out of n objects is $\binom{n}{k}$ i.e. ${}^nC_k$
2) The no. of ways of choosing ordered sample with replacement (WR) is n^k
3) The no. of ways of choosing ordered samples without replacement (WOR) is ${}^nP_k = \dfrac{n!}{(n-k)!}$

Problems:

1) A lot consists of 10 good articles, 4 with minor defects and 2 with major defects. Two articles are chosen at random without replacement. Find the probability that

a) both are good
b) both have major defects
c) atleast one is good
d) atmost one is good

Solve it!

2) Ten persons are wearing badges marked 1 through 10. Three persons are chosen at random and asked to leave the room simultaneously with their badge no. being noted.

a) What is the probability that the smallest badge no. is 5?
b) What is the probability that the largest badge no. is 5?

Solve it!

3) A box contains tags marked 1,2,….n. Two tags are chosen at random. Find the probability that the no.s on the tag will be consecutive integers if

a) the tags are chosen without replacement
b) the tags are chosen with replacement

Solve it!

Conditional Probability:

Consider the example: Suppose we have a lot consisting of 100 items of which 80 are good and 20 are defective. Now we choose two items at random from this lot.
We define A= {the first item is defective} and B= {the second item is defective}.
(a) If the selection is With Replacement (WR), what is P(A) and P(B)?
(b) If the selection is Without Replacement (WOR), what is P(A) and P(B)?
Suppose we choose samples WR then P(A)=20/100=1/5 =P(B).
If we chose samples WOR then P(A)=1/5. But what is P(B)?
To compute we should know whether A did occur or not. Thus if A and B are two events associated with an experiment E, then P(B/A) denotes the conditional probability of the event B, given that A has occurred, i.e. P(B/A)=19/99.

Definition: If A and B are events for which P(B)>0, then we define the conditional probability of A given B denoted by P(A/B) as
$P(A/B)=P(A\cap B)/P(B);\ P(B)>0$
Similarly $P(B/A)=P(A\cap B)/P(A);\ P(A)>0$

Note: Whenever we compute P(B/A) , we are essentially computing the P(B) with respect to the reduced sample space A, rather than the originally sample space S.

Ex: Two fair dice are tossed, and the outcomes are recorded

$$S=\left\{\begin{matrix}(1,1)\ (1,2)\ldots\ldots(1,6)\\(2,1)\ (2,2)\ldots\ldots(2,6)\\ \vdots \\ \vdots \\(6,1)\ (6,2)\ldots\ldots(6,6)\end{matrix}\right\}$$

Let A={(x,y);x+y=10} and B={(x,y);x>y}

Then A={(5,5),(4,6),(6,4)} and B={(2,1),(3,1),(3,2)…….(6,5)}

P(A)=3/36 & P(B)=15/36
Now P(A/B)=? = 1/15 & P(B/A)=1/3
Also, from the above definition we have $P(A/B)= P(A\cap B)/P(B)$
Here $A\cap B$ occurs iff both A and B occur and for this only one outcome is favorable. Hence $P(A\cap B)= 1/36$.

Therefore, $P(A/B)= P(A\cap B)/P(B)= (1/36)/(15/36)=1/15$
Similarly, $P(B/A)= P(B\cap A)/P(A)= (1/36)/(3/36)=1/3$.

Note: P(A/B) satisfies the various axioms,

(i). $0 \leq P(A/B) \leq 1$

(ii). $P(S/B)=1$

(iii). $P((A_1 \cup A_2)/B)=P(A_1/B)+P(A_2/B)$ if $A_1 \cap A_2=\phi$ i.e A & A are mutually exclusive

(iv). $P(A_1 \cup A_2 \cup \ldots /B)= P(A_1/B)+P(A_2/B)+\ldots$; if $A_i \cap A_j \neq \phi$ for $i \neq j$

(v). If B=S, $P(A/S)=P(A \cap S)/P(S)=P(A)$.

Thus we have two ways of computing P(A/B).

(i). Directly, by considering, P(A) with respect to the reduced sample space B.

(ii). Using the above definition by computing $P(A), P(A \cap B)$ with respect to original sample space S.

Multiplication Theorem:

If A and B are any two events, then

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

$\quad\quad = P(A).P(B/A)$.

Using this theorem, we can compute the probability of the simultaneous occurrence of two events.

Ex: In the above example of a lot consisting of 80 good items and 20 defective items if we were to choose 2 items at random (WOR), what is the probability that both of them are defective?

Solution: Define A={first item is defective} B={second item is defective}.

$P(A)=1/5$, $P(B/A)=19/99$, $P(A \cap B)=?$

$P(A \cap B)=P(B/A).P(A)=19/495$

Note: Generalization of Multiplication Theorem.

If $A_1, A_2, A_3 \ldots\ldots, A_n$ are any n events then,

$P(A_1 \cap A_2 \cap A_3 \cap \ldots\ldots \cap A_n)=P(A_1)\, P(A_2/A_1) P(A_3/A_1 A_2) \ldots\ldots P(A_n/A_1 A_2 \ldots A_{n-1})$

Definition: The events $B_1, B_2, \ldots\ldots B_k$ are said to partition the sample space S if,

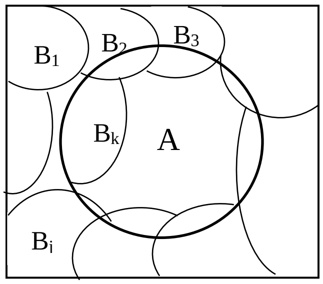
(i). $B_i \cap B_j = \phi \quad \forall\, i \neq j$

(ii). $\bigcup_{i=1}^{k} B_i = S$

(iii). $P(B_i)>0\ \forall\, i$

(i.e. when an experiment E is performed, one and only one of the B_i's occurs)

Theorem of Total Probability:



Let A be an event with respect to S and let B_1, B_2,…… B_k be a partition of S. We may decompose and write A as the union of mutually exclusive events.

i.e. $A= (A\cap B_1) \cup (A\cap B_2) \cup \ldots \cup (A\cap B_k)$ where some of $A\cap B_j$ may be $= \phi$

Therefore $P(A)=P[(A\cap B_1) \cup (A\cap B_2) \cup \ldots \cup (A\cap B_k)]$

$=P(A\cap B_1) + P(A\cap B_2)+\ldots\ldots+P(A\cap B_k)$

By Multiplication theorem,

$P(A)=P(B_1).P(A/B_1)+ P(B_2).P(A/B_2)+ \ldots\ldots+P(B_j).P(A/B_j)+\ldots+ P(B_k).P(A/B_k)$

$P(A)=\sum_{j=1}^{k} P(B_j).P(A/B_j)$ is called the theorem of total probability.

Ex: Consider the lot of 20 defective and 80 non defective items from which we choose two items WOR. What is the probability that the 2^{nd} item selected is defective?

Let A:{1^{st} selection is defective} and B:{2^{nd} selection is defective}

We have $P(B)=P(B/A)P(A)+P(B/\overline{A})P(\overline{A})$

$=(19/99)(1/5)+(20/99)(4/5)=1/5.$

Conditional Probability (Contd…)

As we have already seen the conditional probabilities may also be used as tools for computing unconditional probabilities.

For ex: $P(A\cap B)=P(A/B).P(B) \rightarrow$ (A)
i.e. one can compute $P(A\cap B)$ from the knowledge of $P(A/B)$ and $P(B)$.
Also we know that, $P(A)= P(A/B).P(B)+ P(A/\overline{B}).P(\overline{B}) \rightarrow$ (B)
(i.e. by partitioning S, using B and $\overline{B}$ s.t $B \cap \overline{B}= \phi$)
Finally if $P(A)>0$, we may use (A) and (B) to compute $P(B/A)$
i.e. $P(B/A)=[P(A/B).P(B)] / [P(A/B).P(B)+ P(A/\overline{B}).P(\overline{B})] \rightarrow$ (C)
(C) is a special case of Bayes' theorem

BAYES' THEOREM (Thomas Bayes'- 17th century)

Suppose that $B_1, B_2….B_k$ are k mutually exclusive and collectively exhaustive events and at least one and not more than one of them must have happened, but it is not known which one. Suppose also that an event A may follow any one of the events B_i, with known probability and that A is known to have happened. What is then the probability that it was preceded by a particular event B_i?
The answer to this is given by Bayes' theorem.

Statement: Let $B_1, B_2….B_k$ be a partition of the sample space S and let A be any arbitrary event associated with S such that $P(A)>0$. Then we have,

$$P(B_i/A)=P(B_i).P(A/B_i)/ \sum_{j=1}^{k} P(B_j).P(A/B_j)$$

This is called the Bayes' theorem.

Note:
1. The probability $P(B_1)…P(B_k)$ are typically subjective probabilities which represent our opinion about the nature prior to any experimentation and are termed as '<u>apriori probabilities</u>'. As they exist before we gain any information from the experiment itself
2. Probabilities $P(A/B_i)$, i=1,2,3….k are called 'Likelihood probabilities'.
3. The probabilities $P(B_i/A)$ i=1,2,3….k are called 'posteriori probabilities' as they are determined after the results of the experiments are known.(after the event A observed to occur)
4. Since B_i's are a partition of S one and only one of the events B_i occurs. Hence Bayes' theorem gives us the probability of a particular B_i given that the event has occurred. In order to apply this theorem we must know the values of $P(B_i)$s. Quite often these values are not known and this limits the applicability of the result.

Independent Events:
Quite often we want to know, for what events A and B, it is true that $P(A/B)=P(A)$?
In other words, for what events A and B, it is true that the occurrence of B provides no information about the chance that A will occur?
The answer is, $P(A/B)=P(A)$.
i.e. $P(A\cap B)/P(B)=P(A)$
i.e. $P(A\cap B)=P(A).P(B) \rightarrow$ A .
Thus we say that two events A and B are independent iff A holds.

Definition: Two events A and B are said to be independent iff $P(A\cap B)=P(A).P(B)$.

Note 1: A & B are disjoint then $P(A\cap B)=P(\phi)=0$, so that A and B cannot be independent unless either $P(A)=0$ or $P(B)=0$.

Note 2: If A and B are independent then A and $\bar{B}$ are independent, $\bar{A}$ and B are independent and $\bar{A}$ and $\bar{B}$ are independent.

Prove it!

Definition: (**Pair-wise independent events**) A set of events $A_1,A_2,\ldots.A_n$ are said to be pair- wise independent if $P(A_i\cap A_j)=P(A_i)P(A_j)\ \forall i\neq j$

Definition: (**Mutual independence of n events**) If $A_1,A_2,\ldots.A_n$ are n events, then for their mutual independence we should have
$P(A_{i1}\cap A_{i2}\cap\ldots..\cap A_{ik})=P(A_{i1})P(A_{i2})\ldots..P(A_{ik})\quad k=2,3,4\ldots n.$
i.e. we should have
i) $P(A_i\cap A_j)=P(A_i)P(A_j)$; $(i\neq j,\ i,j=1,2,3,4.....n)$
ii) $P(A_i\cap A_j\cap A_k)= P(A_i)P(A_j)P(A_k)$; $(i\neq j\neq k;\ i,j,k=1,2,3\ldots.n)$
:
:
$P(A_i\cap A_j\cap\ldots.\cap A_n)= P(A_1)P(A_2)\ldots\ldots P(A_n)$.
i.e in all there are 2^n-n-1 conditions. In particular for n=3(say A,B,C)we have $2^3-3-1=4$ conditions for their mutual independence viz,
$P(A\cap B)=P(A).P(B)$, $P(A\cap C)=P(A).P(C)$, $P(B\cap C)=P(B).P(C)$ and $P(A\cap B\cap C)=P(A).P(B).P(C)$.

Theorem: A and B are two events with nonzero probabilities. If they are mutually exclusive then they cannot be independent and conversely.

Proof: Given $P(A)>0$ & $P(B)>0 \rightarrow$ 1

a) A and B are mutually exclusive => $P(A\cap B)=0 \rightarrow$ 2
   For independence we must have $P(A\cap B)=P(A).P(B)>0$ which contradicts 2.
   Hence m.e events cannot be independent.
b) Similarly A and B to be independent => $P(A\cap B)=P(A).P(B)>0 \rightarrow$ 3
   Now for A and B to be m.e., we must have $P(A\cap B)=0$, which contradicts 3.
   Hence independent events cannot be m.e.

Problems:

1) Urn 1 contains x white balls and y red balls. Urn 2 contains z white balls and v red balls. A ball is chosen at random from Urn1 and put into Urn2. Then a ball is chosen from Urn2 at random. What is the probability that the ball is white?

Solution: Define A= {First selection is white} => $\bar{A}$ = {First selection is red}
B= {Second selection is white}
To find P(B)=?
By total probability theorem, we have $P(B)= P(B/A).P(A)+ P(B/\bar{A}).P(\bar{A})$
where $P(A)=x/(x+y)$, $P(\bar{A})=y/(x+y)$,
$P(B/A)=(z+1)/(z+v+1)$,
$P(B/\bar{A})= z/(z+v+1)$.
$\therefore P(B)= [(z+1)/(z+v+1)]. [x/(x+y)] + [z/(z+v+1)]. [y/(x+y)]$

2) 2 two defective tubes get mixed up with 2 good ones. The tubes are tested one by one until both the defectives are found. What is the probability that the last defective tube is obtained **a)** on the 2nd test **b)** on the 3rd test **c)** on the 4th test

Solution:
a) Define the events A_1:{The 1st tube tested is defective}
A_2:{The 2nd tube tested is defective}
D :{ the last defective tube is obtained on the 2nd test}
$P(D)= P(A_1 \cap A_2)= P(A_1)P(A_2/A_1)= (2/4) . (1/3) = 1/6$
b) Define A_3:{3rd tube tested is defective}
D :{last defective tube is obtained in the 3rd test}
$P(D) =P(A_1 \cap \bar{A}_2 \cap A_3)+P(\bar{A}_1 \cap A_2 \cap A_3)$
$= P(A_1)P(\bar{A}_2/A_1). P(A_3/\bar{A}_2,A_1)+ P(\bar{A}_1)P(A_2/\bar{A}_1). P(A_3/\bar{A}_1,A_2)$
$=(1/2)(2/3)(1/2)+(1/2)(2/3)(1/2)=1/6+1/6=1/3.$
c) P{the last defective tube is obtained on the 4th test}
=1-{P(getting the last defective either in the 2nd test or the 3rd test)}
=1-(1/6+1/3)
=1/2.

3) A box contains 4 bad and 6 good tubes. Two are drawn out together. One of them is tested and found to be good. What is the probability that other one is also good?

Solve it!

4) Suppose that A and B are independent events associated with an experiment E. If the probability that A or B occurs is 0.6. If the probability that A occurs is 0.4, determine the probability that B occurs.

Solution: Given : P(AUB)=0.6, P(A)=0.4
$P(AUB)=P(A)+P(B)-P(A \cap B)$
$=P(A)+P(B)-P(A).P(B)$
$0.6= 0.4+P(B)(1-0.4)= 0.4+0.6P(B)$
$P(B)=0.2/0.6=1/3$

5) Let A and B be 2 events associated with an experiment. Suppose that P(A)=0.4 while P(AUB) is 0.7. Let P(B) be p. For what choice of p are the events A and B
i) mutually exclusive? ii) independent?

Solution: P(A)=0.4 P(AUB)=0.7 P(B)=p.

i) P(AUB)=P(A)+P(B) (here $P(A\cap B)=0$)

0.7 = 0.4+P(B)

$\therefore$ p =P(B)=0.3

ii) $P(AUB)=P(A)+P(B)-P(A\cap B)$

= P(A)+P(B)-P(A).P(B)

0.7= 0.4+ P(B)0.6

$\therefore$ p =P(B)=0.3/0.6=0.5

6) An electrical assembly has 2 subsystems A and B. Given P(A fails)=0.2, P(B fails alone)=0.15, P(both A and B fail)=0.15. Evaluate i)P(A fails given B has failed) ii)P(A fails alone)

Solve it!

7) A vacuum tube may come from any of the 3 manufacturers with probabilities p_1=0.25, p_2=0.25, p_3=0.5. The probabilities that the tube will function properly during a specified period of time equal 0.1, 0.2 and 0.4 respectively for the three manufacturers. Compute the probability that a randomly chosen tube will function for the specified period of time.

Solution: Let A:{tube comes from 1st manufacturer}

B:{ tube comes from 2nd manufacturer}

C:{tube comes from 3rd manufacturer}

Let X={Tube will function for the specified period of time}

P(X)=?

$P(X)= P(X\cap A)+ P(X\cap B)+ P(X\cap C)$ {A,B,C are disjoint}

= P(X/A)P(A)+ P(X/B)P(B)+ P(X/C)P(C)

= (0.1).(0.25)+ (0.2).(0.25)+ (0.4).(0.5)= 0.275

8) Three newspapers A,B,C are published in a city and a recent survey of readers indicates the following :20%read A , 16%read B, 14% read C, 8% read A and B, 5% read A and C, 4% read B and C and 2% read A,B & C. For One adult chosen at random, compute the probability that **a)** he reads none of the papers **b)** he reads exactly one of the papers **c)** he reads at least one of A and B, if it is known that he reads least one of the papers published.

Solve it!

9) Given a) $P(\overline{A})=0.4$ P(B/A)=0.5 P(AUB)=0.95 find P(B)?

b) P(AUB)=0.7 $P(\overline{B}/\overline{A})=0.5$ find P(A)?

Solve it!

10) A,B,C,D,E are mutually independent events with P(A)=1/2, P(B)=3/4, P(C)=5/6, P(D)=1/8 and P(E)=2/3. Find P(AUBUCUDUE)?

Solution: P(AUBUCUDUE)=1-P($\overline{AUBUCUDUE}$) =1-[P($\bar{A}\cap\bar{B}\cap\bar{C}\cap\bar{D}\cap\bar{E}$)]

=1-[P($\bar{A}$)P($\bar{B}$)P($\bar{C}$)P($\bar{D}$)P($\bar{E}$)]

=1-[(1/2)(1/4)(1/6)(7/8)(1/3)]

=1145/1152.

11) It is found in manufacturing certain articles defects of Type 1 occur with probability 0.1 and defects of Type 2 occur with probability 0.05. (Assume independence with type of defects). What is the probability that :

i) an article does not have both types of defects?
ii) an article is defective?
iii) an article has one type of defect, given that it is defective?

Solve it!

12) What is the probability that in a group of n people at least 2 of them have same birthday (the same day, month, but other year) [classical birthday problem]?

Solve it!

13) Three dice are rolled independently. Let A:{sum of the digits shown is 6}, and B:{all three digits are different}. Are A and B independent?

Solve it!

14) In a bolt factory, machines A, B, and C manufacture 25%, 35%, and 40% of the total output, respectively. Of their outputs, 5%, 4%, and 2% respectively, are defective bolts. A bolt is chosen at random and found to be defective. What is the probability that the bolt came from machine A? B? C?

Solution:

Let E:{The bolt is defective}

A:{The bolt selected at random is manufactured by M/c A}

B:{The bolt selected at random is manufactured by M/c B}

C:{The bolt selected at random is manufactured by M/c C}

Given, P(A)=0.25, P(B)=0.35, P(C)=0.4

P(selected bolt is defective given it is manufactured by M/c A)=P(E/A)=0.05

Similarly, P(E/B)= 0.04, P(E/C)=0.02

Now the required probability is,

P(selected bolt is manufactured by M/c A given that bolt is defective)

=P(A/E)

=P(E/A).P(A)/ (P(E/A).P(A)+ P(E/B).P(B)+ P(E/C).P(C))

=0.05(0.25)/(0.05(0.25)+0.04(0.35)+0.02(0.4))

=0.362

Similarly we can obtain P(B/E)=0.406 and P(C/E)=0.232.

RANDOM VARIABLES

Introduction

While describing the sample space of an experiment, we did not specify that an individual outcome needs to be a number. But in many situations, we are not concerned with all aspect of the outcome of an experiment but only in a particular numerical value of the outcome, such as, the number of red balls in a sample, the height of a randomly selected man, or the income of a randomly selected family. In other words, in many experimental problems (situations), we want to assign a real number x to every element s ∈ S. A random variable is used for this purpose.

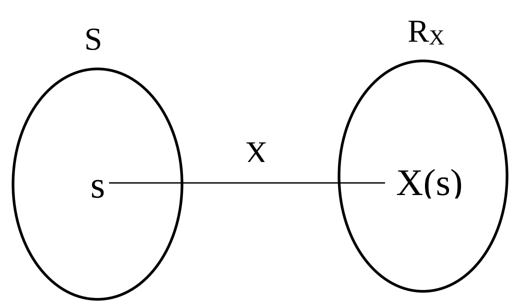
Definition: Let S be the sample space associated with a experiment E. A real valued function X defined on S and taking values in R (-∞, ∞) is called a **Random Variable** (rv) (or one-dimensional rv). i.e., a function X assigning to every element s ∈ S, a real number X(s) is called a rv and R is called the **Range space.**

Ex: Suppose a coin is tossed twice. Let X=No. of heads appearing. Then X takes values 0, 1, 2 as below:

Outcome	HH	HT	TH	TT
X: No. of Heads	2	1	1	0

Note that with every outcome s, there corresponds a real number X(s)=x. Thus, X is a rv.

One dimensional rvs will be denoted by X, Y, Z,…. The values which X, Y, Z,…… can assume are denoted by x, y, z,…………If x is a real number, the set of all s ∈ S such that X(s) = x is denoted briefly by X=x.



i.e., {X = x} = {s ∈ S : X(s) = x}.

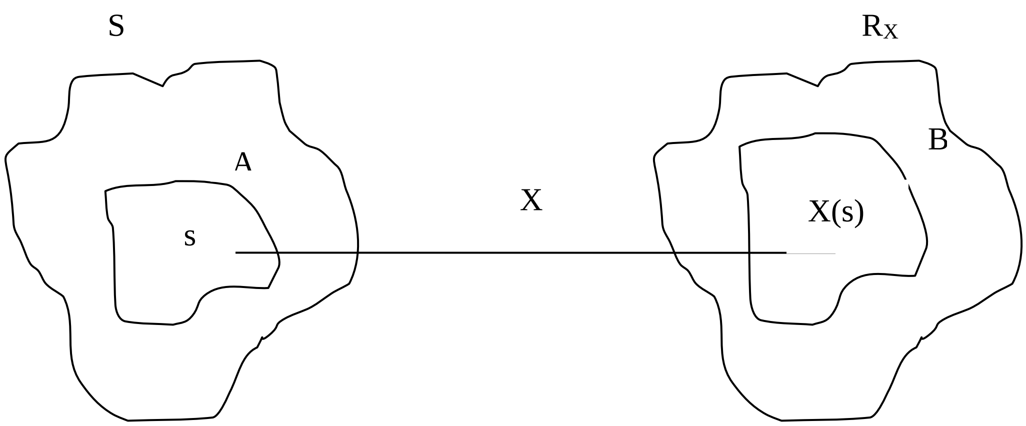
Thus, P(X = x) = P{s: X(s) = x}

$P(X \leq a) = P\{s: X(s) \in [-\infty, a]\}$

$P(a < X \leq b) = P\{s: X(s) \in [a,b]\}$

Results: If X, X_1 and X_2 are rvs, C, C_1 and C_2 are constants, then CX, X_1X_2, $X_1 + X_2$, $X_1 - X_2$, $C_1X_1 \pm C_2 X_2$, |X|, Max(X_1 , X_2), Min(X_1 , X_2), f(x) [where f(.) is a continuous function of X], etc. are all rvs.

Equivalent events: Let E be an experiment and S be the sample space. Let X be a rv defined on S and let R_X be its range space. Let B be an event w.r.t R_X; i.e. B is a subset of R_X.
Suppose that A is defined as A = {s ∈ S: X(s) ∈ B} i.e. A consists of all outcomes in S for which X(s) ∈ B. Then we say that A and B are equivalent events. i.e. A and B are equivalent events whenever they occur together.



Definition: Let B be an event in the range space R_X. We define P(B) as P(B) = P(A) where A={s ∈ S : X(s) ∈ B}. i.e. we define P(B) equal to the probability of the event A which is equivalent to B.

Note: We are assuming that probabilities may be associated with events in S. Hence the above definition makes it possible to assign probabilities to events associated with R_X in terms of probabilities defined over S.

Discrete Random Variable (drv): Let X be a rv. If X takes atmost a countable number of values, it is called a Discrete Random Variable. i.e. a real valued function defined on a discrete sample space is called drv.
In Otherwords, if the number of possible values of X is finite or countably infinite, we call X as drv. i.e. the possible values of X may be listed as $x_1, x_2, x_3, \ldots\ldots, x_n, \ldots\ldots$ In the finite case, the list terminates and in countably infinite case, the list continues indefinitely.

Ex: (i) X = No. of α – particles emitted by a radioactive source in a given period.

(ii) X = No. of telephone calls received at a telephone exchange in a specified time interval.

Definition: Probability Mass Function (pmf) and Probability Distribution of a drv X

Let X be a drv taking atmost a countably infinite number of values $x_1, x_2, \ldots, x_n, \ldots$ with each possible outcome x_i we associate a number $p(x_i) = P(X = x_i)$ called the probability of x_i. The numbers $p(x_i)$, i=1,2,3,………must satisfy the following conditions.

(i) $p(x_i) \geq 0 \ \forall\, i$

(ii) $\sum_{i=1}^{\infty} p(x_i) = 1$

This function $p(x_i)$ is called the pmf of the drv X and the set $\{p(x_i)\}$ or $\{x_i, p(x_i)\}$, i=1,2,3,… is called the probability distribution of rv X.

Depending on the nature or the functional form of the pmf, we have several Discrete Probability Distributions.

Binomial Distribution:

Consider an experiment E. Let A be an event associated with E. Let P(A) = p so that $P(\bar{A}) = 1-p$. Consider 'n' independent repetitions of E. The sample space consists of all possible sequences $\{a_1, a_2, \ldots, a_n\}$, where a_i is either A or $\bar{A}$ depending on whether A or $\bar{A}$ occurred on the i^{th} repetition of E.

Let the rv X denote the number of times the event A occurred. We call X a Binomial rv with parameters n and p and its possible values are 0,1,2,3,…..,n (i.e. we say X has a Binomial distribution with parameters n and p). The individual repetitions are called **Bernoulli trials.**

Theorem: Let X be a Binomial rv based on n repetition.

Then $P(X = k) = \binom{n}{k} . p^k . (1-p)^{n-k}$, k = 0,1,2,3,……,n.

Proof: Consider a particular order in an outcome satisfying order in an outcome satisfying X = k as AAAAA…..A (k times) and $\bar{A}\,\bar{A}\,\bar{A}\,\bar{A}$ ……… $\bar{A}$ (n-k times).

i.e., 1^{st} k repetition resulted in the occurrence of A while the last n-k repetition resulted in the occurrence of $\bar{A}$. Since all the repetition are independent, we have,

P(AAAAA…..A (k times) and $\bar{A}\,\bar{A}\,\bar{A}\,\bar{A}$ ……… $\bar{A}$ (n-k times)) = $p^k . (1-p)^k$

But exactly the same probability would be associated with any other outcome for which X = k. The total number of such outcomes in $\binom{n}{k}$ (i.e., we must choose exactly k (out of n) positions for A's). Now these $\binom{n}{k}$ outcomes are all mutually exclusive.

Hence using addition theorem of probability, we get the expression

$P(X = k) = \binom{n}{k} . p^k . (1-p)^{n-k}$, k = 0,1,2,3,……,n.

Continuous Random Variable (crv): A rv X is said to be continuous if it can take all possible values between certain limits. In otherwords, a rv is said to be continuous when its different values cannot be put in one to one correspondence with a set of positive integers.

Definition: Probability Density Function (pdf) of a crv X

Let X be a crv. Then the function f is said to be the probability density function of X if it satisfies the following condition:

(i) $f(x) \geq 0 \;\; \forall x$

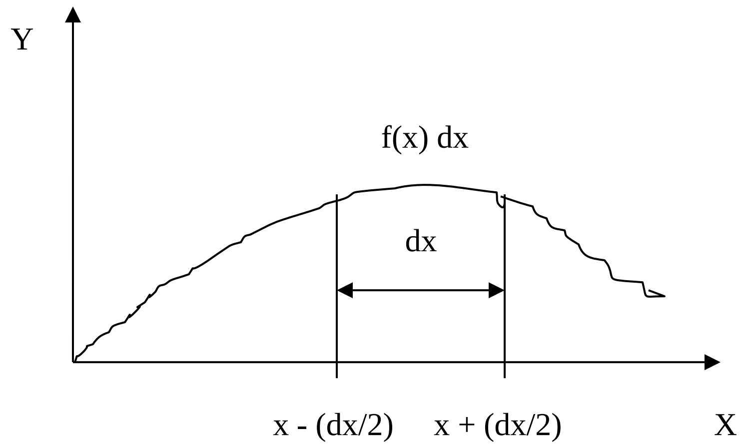
(ii) $\int_{-\infty}^{\infty} f(x)dx = 1$

(iii) $P(a \leq X \leq b) = \int_a^b f(x)\,dx \;; -\infty < \epsilon\, a < b < \infty$.

Definition: Let X be a crv. Let f(x) be any continuous function of x so that f(x)dx represents the probability that X falls in a small interval (x - (dx/2), x + (dx/2)).

i.e. $f(x)dx = P(x - (dx/2) \leq X \leq x + (dx/2))$.

In the figure, f(x)dx represents the area bounded by the curve. Thus f(x) is called the pdf of X.



It is a consequence of (iii) above that for any specified value of X, say x_0, we have

$P(X = x_0) = 0$, as $P(X = x_0) = P(x_0 \leq X \leq x_0) = \int_{x_0}^{x_0} f(x)dx = 0$

Accordingly, for a continuous rv X, we have

$P(a \leq X \leq b) = P(a \leq X < b) = P(a < X \leq b) = P(a < X < b)$

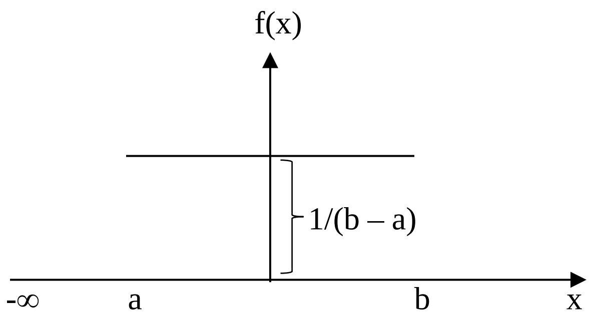
Again, depending on the nature or the functional form of the pdf, we have several Continuous Probability Distributions.

Uniform distribution:

If X is a crv assuming all values in [a,b] $(a < \infty$ and $b < \infty)$ and its pdf is given by

$$f(x) = \begin{cases} \dfrac{1}{b-a}; & a \leq x \leq b \\ 0 \quad ; & \text{Otherwise} \end{cases}$$

then X is said to be Uniformly distributed over the interval [a,b].



Note 1: Uniformly distributed random variables have a pdf constant over the given interval.

Note 2: For any subinterval [c, d] such that $a \leq c < d \leq b$, $P(c \leq X \leq d)$ is the same for all subintervals having the same length.

i.e. $P(c \leq X \leq d) = \int_c^d f(x)\,dx = (d - c)/(b - a)$

So, it depends only on the length of the interval and not on the location of that interval.

Note 3: By choosing a point at random in [a, b], we mean that the x – coordinate of the chosen point, say x, is uniformly distributed over [a, b].

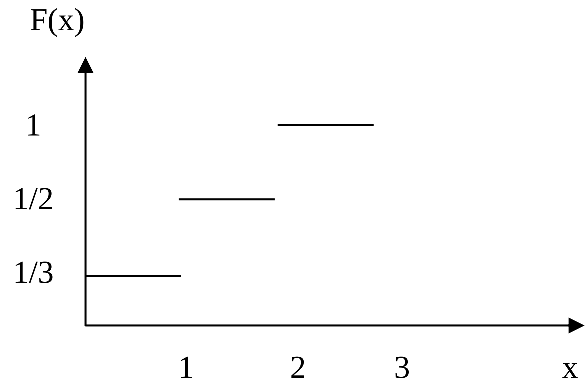
Cumulative Distribution Function (cdf): Let X be a rv. Then the cdf of X denoted by F(x) is given by $F(x) = P(X \leq x)$.

i.e $$F(x) = \begin{cases} \sum_{(j;\, x_j \leq x)} p(x_j) & \text{if X is a drv} \\ \int_{-\infty}^{x} f(s)\,ds & \text{if X is a crv} \end{cases}$$

Ex 1: Let X be a rv taking 3 values 0,1,2 with probabilities 1/3, 1/6, and 1/2. Then, we have,

$F(x) = P(X \leq x)$

$= \sum_{(j;\, x_j \leq x)} p(x_j)$ if X is a drv



$$F(x)=\begin{cases}0; & x<0\\ \frac{1}{3}; & 0\le x<1\\ \frac{1}{2}; & 1\le x<2\\ 1; & x\ge 2\end{cases}$$

Note 1: If X is a drv taking values $x_1, x_2, \ldots\ldots, x_n$, then we have

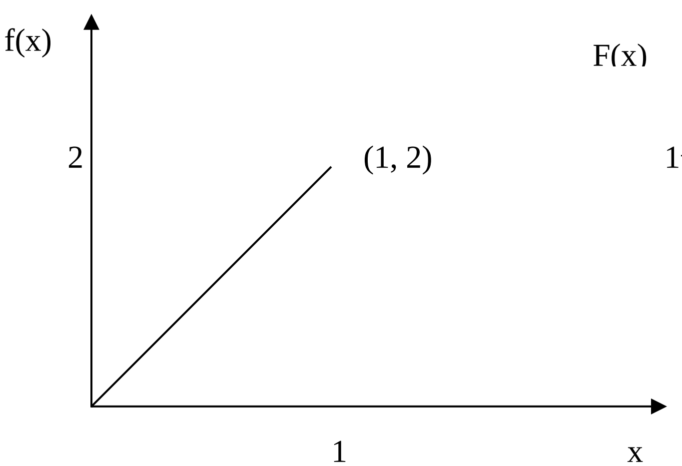
$$F(x)=\begin{cases}0 & ; \quad x<x_1\\ p(x_1) & ; x_1\le x<x_2\\ p(x_1)+p(x_2) & ; x_2\le x<x_3\\ \quad\vdots & \\ p(x_1)+p(x_2)+\cdots+p(x_{n-1}) & ; x_{n-1}\le x<x_n\\ p(x_1)+p(x_2)+\cdots+p(x_n)=1; & \quad x\ge x_n\end{cases}$$

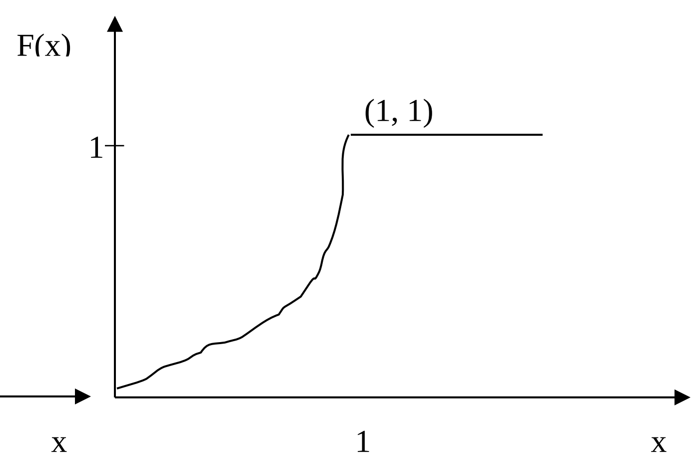
Note 2: For a drv X, F(x) makes jumps at specific values of x i.e. F(x) increases in steps and is called a step function.

Ex 2: Let X be a crv having pdf $f(x)=\begin{cases}2x; & 0<x<1\\ 0; & \text{Otherwise}\end{cases}$

Then, $F(x) = P(X \le x) = \int_{-\infty}^{x} f(s)ds$

$$F(x)=\begin{cases}0; & x\le 0\\ \int_0^x 2s\,ds = x^2; & 0<x<1\\ 1; & x\ge 1\end{cases}$$





Remarks:

1. $0 \le F(x) \le 1$, $-\infty < x < \infty$.

2. F is a non-decreasing function. i.e., if $x_1 < x_2$, then $F(x_1) \le F(x_2)$.

3. $F(-\infty) = \lim_{x\to-\infty} F(x) = 0$; $F(+\infty) = \lim_{x\to\infty} F(x) = 1$.

4. If X is a drv with pmf $p(x_j)$ and cdf $F(x)$, then

$$p(x_j) = P(X = x_j) = F(x_j) - F(x_j - 1).$$

If X is a crv with pdf $f(x)$ and cdf $F(x)$ then,

$$f(x) = \frac{d}{dx}[F(x)] \ \forall \ x \text{ at which F is differentiable,}$$

$$= F'(x)$$

5. Let X be a crv with pdf $f(x)$ and cdf $F(x)$. Then,

$$P(a \le X \le b)=F(b)-F(a)$$

Proof:

$$P(a \le X \le b) = \int_a^b f(x)dx = \int_{-\infty}^b f(x)dx - \int_{-\infty}^a f(x)dx$$

$$= P(X \le b) - P(X \le a)$$

$$= F(b) - F(a).$$

Since, for a crv X, we have $P(X=x_0)=0$, we get

$P(a \le X \le b)=P(a < X < b) = P(a < X \le b) = P(a \le X < b) = F(b) - F(a)$.

However, if X is a drv with pmf $p(x_i)$ and cdf $F(x)$, we have

$P(a < X \le b) = F(b) - F(a)$

$P(a \le X \le b) = P(X = a) + F(b) - F(a)$

$P(a < X < b) = F(b) - F(a) - P(X = b)$

$P(a \le X < b) = F(b) - F(a) + P(X = a) - P(X = b)$.

Prove it!

Problems:

1. A coin is known to come up heads three times as often as tails. This coin is tossed three times. Let X be the number of heads that appear. Writeout the probability distribution of the rv X. Also obtain the cdf of X.

Solution: Given that P(H) = ¾ =p $\Rightarrow$ P(T) = ¼ = q and n=3,

Let X = No. of heads; (taking values {0,1,2,3}).

Clearly, X~B(n,p). Thus we have,

$P(X = x) = \binom{n}{x}.p^x.\ q^{n-x}$; x=0,1,2,3…n; p+q=1

$P(X = 0) = \binom{n}{0}.p^0.\ q^n = 1(1/4)^3 = 1/64$

$P(X = 1) = \binom{n}{1}.p^1.\ q^{n-1} = 3(3/4)(1/4)^2 = 9/64$

$P(X = 2) = \binom{n}{2}.p^2.\ q^{n-2} = 3(3/4)^2(1/4) = 27/64$

$P(X = 3) = \binom{n}{3}.p^3.\ q^0 = 1(3/4)^3(1) = 27/64$

Now, $F(x) = P(X \leq x)$

$= \sum_{(j;\ x_j \leq x)} p(x_j)$

$$F(x) = \begin{cases} 0\ ; & x < 0 \\ \frac{1}{64}; & 0 \leq x < 1 \\ \left(\frac{1}{64}\right) + \left(\frac{9}{64}\right) = \frac{10}{64}; & 1 \leq x < 2 \\ \left(\frac{10}{64}\right) + \left(\frac{27}{64}\right) = \frac{37}{64}; & 2 \leq x < 3 \\ \left(\frac{37}{64}\right) + \left(\frac{27}{64}\right) = 1; & x \geq 3 \end{cases}$$

2. Given $f(x) = \begin{cases} kx^3; & 0 < x < 1 \\ 0; & \text{elsewhere} \end{cases}$

   Find k so that the above is a pdf and hence find

   (i) P(1/4 < X < 3/4)

   (ii) P(X < 1/2)

   (iii) P(X > 0.8)

   (iv) CDF of X.

Solution: We have $\int_{-\infty}^{\infty} f(x)dx = 1$.

Here $\int_0^1 f(x)dx = 1 \Rightarrow \int_0^1 k\,x^3 dx = 1 \Rightarrow [k\,x^4/4]_0^1 = 1 \Rightarrow k/4 = 1$ or k=4.

(i). $P(1/4 < X < 3/4) = \int_{1/4}^{3/4} 4x^3\,dx = [4\,x^4/4]_{1/4}^{3/4} = (3/4)^4 - (1/4)^4 = 80$.

(ii). $P(X < 1/2) = \int_0^{1/2} 4x^3\,dx = [4\,x^4/4]_0^{1/2} = (1/2)^4 = 1/16$.

(iii). $P(X > 0.8) = \int_{0.8}^{1} 4x^3 dx = [4\,x^4/4]_{0.8}^{1} = 1-(0.8)^4 = 0.5904$.

(iv). $F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt$

For $x \leq 0$, $f(x) = 0 \Rightarrow F(x) = 0$.

For $0 < x < 1$, $F(x) = \int_{-\infty}^{0} f(t)dt + \int_0^x f(t)dt = 0 + \int_0^x 4t^3\,dt = \frac{4t^4}{4}]_0^x = x^4$

For $x \geq 1$, $F(x) = \int_{-\infty}^{0} f(t)dt + \int_0^1 f(t)dt + \int_1^x f(t)dt = 0 + \int_0^1 4t^3 + 0 = [t^4]_0^1 = 1$.

Hence, $F(x) = \begin{cases} 0; & x \leq 0 \\ x^4; & 0 < x < 1 \\ 1; & x \geq 1 \end{cases}$

3. A rv X assumes 4 values with probabilities (1+3x)/4, (1-x)/4, (1+2x)/4 and (1-4x)/4. For what range of values of x is this a probability distribution?

   Solve it!

4. Suppose that the rv X has possible values 1,2,3,4,……..and

   $P(X = j) = 1/2^j$, j=1,2,3,…………

   Compute: a) P(X is even) b) $P(X \geq 5)$ c) P(X is divisible by 3)

   Solve it!

5. Let X be a crv with pdf given by

$$f(x) = \begin{cases} ax\,; & 0 \le x \le 1 \\ a; & 1 \le x \le 2 \\ -ax + 3a; & 2 \le x \le 3 \\ 0; & \text{elsewhere} \end{cases}$$

(a) Determine the constant ‘a’

(b) Obtain the cdf F(x).

Solve it!

6. The diameter on an electric cable, say X, is assumed to be a crv with pdf,

$f(x) = 6x(1-x)$; $0 < x < 1$.

(a) Check whether the above f(x) is a pdf

(b) Obtain the cdf of X

(c) Determine a number ‘b’ such that $P(X > b) = 2\,P(X > b)$

(d) Compute $P[(X \le 1/2) / (1/3 < X < 2/3)]$.

Solve it!

7. Suppose that X is a uniformly distributed rv, over the interval (-a, +a) where a > 0. Determine ‘a’, wherever possible, so that the following are satisfied:

(a) $P(X > 1) = 1/3$

(b) $P(X < 1) = 1/2$

(c) $P(X < 1/2) = 0.7$

Solve it!

8. Let the rv K be uniformly distributed over the interval [0, 5]. What is the probability that the roots of the $4x^2 + 4xk + k + 2 = 0$ are real?

Solution: Since K is uniformly distributed,

We have $f(k) = \begin{cases} \frac{1}{5}; & 0 \leq k < 5 \\ 0; & \text{elsewhere} \end{cases}$

To find: P(the roots of the equation $4x^2 + 4xk + k + 2 = 0$ are real)=?

Now, the roots of the equation $4x^2 + 4xk + k + 2 = 0$ are given by $\frac{-4 \pm \sqrt{16k^2 - 16(k+2)}}{8}$

For the roots to be real, we must have, $16k^2 - 16(k+2) \geq 0 \Rightarrow k^2 - k - 2 \geq 0$

$\Rightarrow (k - 2)(k + 1) \geq 0$

∴ Required probability = $P((k - 2)(k + 1) \geq 0)$=?

Now, we have 2 possibilities:

(i) $k-2 \geq 0$ and $(k + 1) \geq 0$ implies $k \geq 2$ and $k \geq -1 \Rightarrow k \geq 2 \Rightarrow k \in (2, 5)$.
Hence, probability of $k \in (2, 5) = P(2 \leq k \leq 5) = \int_2^5 1/5 \, dk = 3/5$.

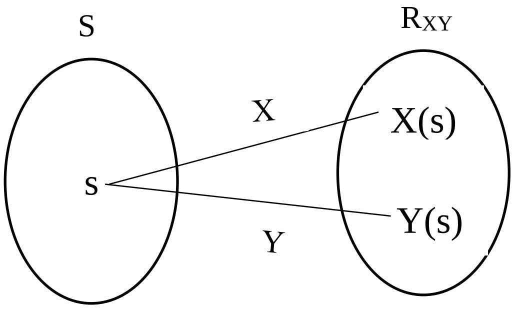
OR

(ii) When $k-2 \leq 0$ and $k+1 \leq 0$ implies $k \leq 2$ and $k \leq -1. \Rightarrow k \leq -1 \Rightarrow k \in (-\infty, -1)$.
Hence, $P(k \in (-\infty, -1)) = P(-\infty \leq k \leq -1) = 0$

∴ Required probability = 3/5.

TWO AND HIGHER DIMENSIONAL RANDOM VARIABLES

Definition:
Let S be the sample space associated with a given random experiment E. Let X=X(s) and Y=Y(s) be two functions each assigning a real number to each outcome $s \in S$. Then we call the pair (X,Y) a 2-dimensional random variable.



Ex: Height and weight of a randomly chosen person.

Definition:
If $X_1=X_1(s)$, $X_2=X_2(s)$……$X_n=X_n(s)$ are n functions each assigning a real number to every outcome $s \in S$, then we call $(X_1,X_2….X_n)$ a n-dimensional random variable.

Definition:

Discrete case: A 2-dimensional rv (X,Y) is said to be discrete if the possible values of (X,Y) are finite or countably infinite.

Continuous case: A 2-dimensional rv (X,Y) is said to be continuous if it can take all values in some non-countable set R of the Euclidean plane.

Probability distribution:

Discrete case: Let (X,Y) be a 2-dimensional drv. With each possible outcome (x_i,y_j) we associate a No. $p(x_i,y_j) = P(X=X_i, Y= Y_j)$, satisfying

1. $p(x_i, y_j) \geq 0 \ \forall \ i, j$
2. $\sum_{j=1}^{\infty} \sum_{i=1}^{\infty} p(x_i, y_j) = 1$

The function $p(x_i,y_j)$ defined above is called the joint probability mass function (pmf) of the 2-dimensional drv (X,Y) and the set of triplets$(x_i, y_j, p(x_i,y_j))$, where i,j=1…n is called the joint probability distribution of the 2-dimensional drv (X,Y).

Continuous case: Let (X,Y) be a 2-dimensional crv. Then there exists a function f(x,y) called the joint pdf of the 2-dimensional crv (X,Y) satisfying

1. $f(x, y) \geq 0 \ \forall \ x, y \ \in R.$
2. $\iint_R f(x, y)dxdy = 1$ or equivalently $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)dxdy = 1.$

Also, 3. $P(a \leq X \leq b, c \leq Y \leq d) = \int_c^d \int_a^b f(x, y)dxdy$

Joint Cumulative distribution function:

Let (X,Y) be a 2-dimensional random variable then the joint cdf F of (X,Y) is defined as
$F(x,y) = P(X \le x, Y \le y) = P(-\infty < X \le x, -\infty < Y \le y)$

$$= \sum_{\{(i,j) : (x_i,y_j) \le (x,y)\}} \sum p(x_i,y_j) \quad \text{if (X,Y) is a 2-dimensional drv}$$

$$= \int_{-\infty}^{x} \int_{-\infty}^{y} f(s,t)\, ds\, dt \quad \text{if (X,Y) is a 2-dimensional crv}$$

Marginal probability distribution:

Discrete case: Let us associate with each 2-dimensional drv (X,Y), two one dimensional random variables, say, X and Y individually. Our interest may be in finding the probability distribution of X and the probability distribution of Y respectively. We have the joint probability distribution of X and Y given by $p_{ij} = p(x_i,y_j) = P(X=x_i, Y=y_j)\ \forall\, i,j$.

This is usually represented in the tabular form as shown below:

X\Y	y_1	y_2	…………	…………	y_m	Total
x_1	p_{11}	p_{12}	…………..	…………	p_{1m}	$P_{1.}$
x_2	p_{21}	p_{22}	…………	………..	p_{2m}	$P_{2.}$
:	:	:	………..	…………	:	:
:	:	:	………..	…………	;	:
x_n	p_{n1}	p_{n2}	………..	…………	p_{nm}	Pn.
Total	$P_{.1}$	$P_{.2}$	………..	………..	P.m	1

Note: $\sum_{i=1}^{n} \sum_{j=1}^{m} p(x_i, y_j) = 1$.

The probability distribution of X called the Marginal probability distribution of X is given by
$p(x_i) = P(X=x_i) = P(X=x_i, Y=y_1) + P(X=x_i, Y=y_2) + \ldots\ldots\ldots + P(X=x_i, Y=y_j) + \ldots.$

$$= \sum_{j=1}^{n} p_{ij}$$

$$= \sum_{j=1}^{\infty} p(x_i, y_j)\ \forall\, i.$$

(since $X=x_i$ must occur with $Y=y_j$ for some j and can occur with $Y=y_j$ for only one j)
Similarly, $q(y_j) = P(Y=y_j) = \sum_{i=1}^{\infty} p(x_i, y_j)\ \forall\, j$ is the Marginal probability distribution of Y.

Continuous case: Let f (x,y) be the joint pdf of a 2-dimensional crv (X,Y). Then g(x) and h(y), the marginal probability density functions of X and Y, are respectively given by

$$g(x) = \int_{-\infty}^{\infty} f(x,y)\,dy$$

$$h(y) = \int_{-\infty}^{\infty} f(x,y)\,dx$$

Note: These pdfs g(x) and h(y) correspond to the basic pdfs of one-dimensional random variables X and Y respectively.

Also, $P(a \le X \le b) = P(a \le X \le b; -\infty < Y < \infty)$

$$= \int_{a}^{b} \int_{-\infty}^{\infty} f(x,y)\,dy\,dx = \int_{a}^{b} g(x)\,dx$$

Similarly, we obtain, $P(a \le Y \le b) = \int_{a}^{b} h(y)\,dy$.

Conditional probability distribution:

Discrete case: Let (X,Y) be a 2-dimensional drv with joint probability distribution $p(x_i,y_j)$. Let $p(x_i)$ and $q(y_j)$ be the marginal probability distributions of X and Y respectively. Then the conditional probability distribution of $X=x_i$ given $Y=y_j$ is defined as

$$p(x_i/y_j) = P(X=x_i/Y=y_j)$$
$$= P(X=x_i,Y=y_j) / P(Y=y_j)$$
$$= p(x_i,y_j)/q(y_j);\ q(y_j)>0$$

Similarly, the conditional probability distribution of $Y=y_j$ given $X=x_i$ is defined as

$$q(y_j /x_i)= P(Y=y_j / X=x_i)$$
$$= p(x_i ,y_j) / p(x_i);\ p(x_i)> 0$$

Continuous case: Let (X,Y) be a 2-dimensional crv with joint pdf f(x,y). Let g(x) and h(y) be the marginal pdfs of X and Y respectively. Then the conditional pdf of X given Y is defined as
$g(x/y) = f(x,y) / h(y)$; $h(y)>0$

Similarly, the conditional pdf of Y given X is defined as
$h(y/x) = f(x,y) / g(x)$; $g(x)>0$.

Independent random variables:

Discrete case: Let (X,Y) be a 2-dimensional drv with joint probability distribution $p(x_i,y_j)$. Let $p(x_i)$ and $q(y_j)$ be the marginal probability distributions of X and Y respectively. Then X and Y are said to be independent random variables iff

$p(x_i,y_j) = p(x_i).\ q(y_j)\ \forall$ i,j.

i.e. $P(X=x_i,Y=y_j) = P(X=x_i)\ P(Y=y_j)\ \forall$ i,j.

Continuous case: Let (X,Y) be a 2-dimensional crv with joint pdf f(x,y). Let g(x) and h(y) be the marginal pdfs of X and Y respectively. Then X and Y are said to be independent random variables iff

$f(x,y) = g(x).h(y)\ \forall$ x,y.

In other words

(a) Let (X,Y) be a 2-dimensional drv. Then X and Y are said to be independent iff
$p(x_i/y_j) = p(x_i)\ \forall$ i,j. [or≡ iff $q(y_j/x_i) = q(y_j), \forall i, j$]

(b) Let (X,Y) be a 2-dimensional crv. Then X and Y are said to be independent iff
$g(x/y) = g(x)\ \forall$ x,y. [or≡ iff $h(y/x) = h(y), \forall x, y$]

PROBLEMS

1. Suppose that the joint pdf of the 2-dimensional crv (X,Y) is given by

$$f(x,y) = \begin{cases} x^2 + \frac{xy}{3}; & 0 \le x \le 1\ , 0 \le y \le 2 \\ 0 & ; \quad \text{elsewhere} \end{cases}$$

Find (i). g(x)
(ii). h(y)
(iii). g(x/y)
(iv). h(y/x)
(v). P(X>$\frac{1}{2}$)
(vi). P(Y<X)
(vii).P[(Y<$\frac{1}{2}$)/(X<$\frac{1}{2}$)]

Solution:

(i). g(x) =$\int_{-\infty}^{\infty} f(x,y)dy$=$\int_0^2 (x^2 + \frac{xy}{3})dy = 2x^2 + \frac{2x}{3}$, 0≤x≤1

(ii). h(y) =$\int_{-\infty}^{\infty} f(x,y)dx = \int_0^1 (x^2 + \frac{xy}{3})dx$=$\frac{y}{6}$+$\frac{1}{3}$, 0≤y≤2

(iii). g(x/y) = f(x,y) / h(y) ; h(y)>0

$$= \frac{x^2 + \frac{xy}{3}}{\frac{1}{3} + \frac{y}{6}} = \frac{6x^2 + 2xy}{2 + y}\ , 0\le x\le 1 \text{ and } 0\le y\le 2$$

(iv). h(y/x) = f(x,y)/g(x) ; g(x)>0.

$$= \frac{x^2 + \frac{xy}{3}}{2x^2 + \frac{2x}{3}} = \frac{3x + y}{6x + 2}\ , 0 \le x \le 1 \text{ and } 0 \le y \le 2$$

(v). P(X>$\frac{1}{2}$)= P(X>$\frac{1}{2}$,0≤ Y ≤)=$\int_0^2 \int_{\frac{1}{2}}^1 f(x,y)dxdy$

$$P\left(X > \frac{1}{2}\right) = \int_0^2 \int_{\frac{1}{2}}^1 \left(x^2 + \frac{xy}{3}\right) dxdy = \int_0^2 [(\frac{x^3}{3} + \frac{x^2y}{6})]_{\frac{1}{2}}^1 dy$$

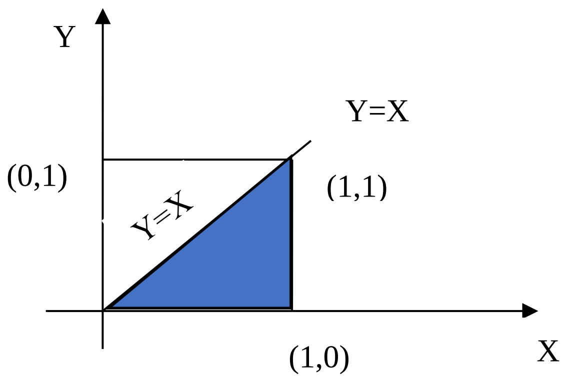
$$= \int_0^2 (\frac{1}{3} + \frac{y}{6} - \frac{1}{24} - \frac{y}{24})dy = \int_0^2 \left(\frac{7}{24} + \frac{3y}{24}\right) dy$$

$$= \left[\frac{7}{24}y\right]_0^2 + \left[\frac{y^2}{16}\right]_0^2 = \frac{5}{6}$$

(vi). $P(Y < X) = \int_0^1 \int_0^x \left[x^2 + \frac{xy}{3}\right] dydx$

$$= \int_0^1 \left[x^3 + \frac{x^3}{6}\right] dx = \left[\frac{x^4}{4} + \frac{x^4}{24}\right]_0^1$$

$$= \frac{7}{24}$$



(vii). $P(Y < \frac{1}{2}/X < \frac{1}{2}) = \frac{P(Y < \frac{1}{2} \cap X < \frac{1}{2})}{P(X < \frac{1}{2})}$

$$P\left(Y < \frac{1}{2} \cap X < \frac{1}{2}\right) = \int_0^{\frac{1}{2}} \int_0^{\frac{1}{2}} \left(x^2 + \frac{xy}{3}\right) dxdy = \int_0^{\frac{1}{2}} \left(x^2 y + \frac{xy^2}{6}\right)_0^{\frac{1}{2}} dx$$

$$= \left[\frac{x^3}{6} + \frac{x^2}{48}\right]_0^{\frac{1}{2}} = \frac{5}{19}$$

$$P(X < \frac{1}{2}) = \int_0^2 \int_0^{\frac{1}{2}} \left[x^2 + \frac{xy}{3}\right] dxdy = \int_0^2 \left[\frac{x^3}{3} + \frac{x^2 y}{6}\right]_0^{\frac{1}{2}} dy$$

$$= \left[\frac{y}{24} + \frac{y^2}{48}\right]_0^2 = \frac{1}{6}$$

$$\therefore\ P(Y < \frac{1}{2}/X < \frac{1}{2}) = \frac{\frac{5}{192}}{\frac{1}{6}} = \frac{5}{32}$$

2. Suppose that the 2-dimensional crv (X,Y) has the pdf

$$f(x,y) = \begin{cases} Cx(x-y)\ ; & 0 < x < 2, -x < y < x \\ 0\ ; & \text{elsewhere} \end{cases}$$

a) Evaluate C
b) Find the marginal pdf of X
c) Find the marginal pdf of Y

Solve it!

3. Let $f(x,y) = \begin{cases} \frac{2}{a^2} \; ; & 0 \leq x < y \leq a \\ 0 \; ; & \text{elsewhere} \end{cases}$ be the joint pdf of the 2-dimensional crv (X,Y).
   Find g(x/y) and h(y/x)?

Solution:

We have $g(x/y) = \frac{f(x,y)}{h(y)}$; $h(y) > 0$ and $h(y/x) = \frac{f(x,y)}{g(x)}$; $g(x) > 0$

$$g(x) = \int_x^a f(x,y)dy = \int_x^a \frac{2}{a^2} dy = \left[\frac{2y^2}{a}\right]_x^a = \frac{2}{a^2}(a-x), 0 \leq x \leq a$$
$$= 0, \text{elsewhere}$$

$$h(y) = \int_0^y \frac{2}{a^2} dx = \left[\frac{2}{a^2} x\right]_0^y = \frac{2y}{a^2}, 0 \leq y \leq a$$
$$= 0, elsewhere$$

$$g(x/y) = \frac{f(x,y)}{h(y)} = \frac{\frac{2}{a^2}}{\frac{2y}{a^2}} = \frac{1}{y}, o \leq y \leq a$$
$$= 0, \text{elsewhere}$$

$$h(y/x) = \frac{f(x,y)}{g(x)} = \frac{\frac{2}{a^2}}{\frac{2(a-x)}{a^2}} = \frac{1}{a-x}; 0 \leq x \leq a$$
$$= 0, \text{elsewhere}$$

4. For what values of ‘k’, is $f(x,y)=ke^{-(x+y)}$ a joint pdf of (X,Y) over the region $0<x<1;0<y<1$?

   Solve it!

5. Two rvs X and Y have their joint pdf given by
   $f(x,y) = \begin{cases} 6(e^{-2x-3y}); & x,y \geq 0 \\ 0 \; ; & \text{elsewhere} \end{cases}$
   Find
   a) $P(1<X<2, 2<Y<3)$
   b) $P(0<X<2, Y>2)$
   c) Marginal pdfs of X and Y. Also the Conditional pdfs of X given Y and Y given X

   Solve it!

6. Test for the independence of the rvs X and Y, given the joint pdf
   $f(x,y) = \begin{cases} 2(e^{-x-y}); & 0 < x < y < \infty \\ 0 \; ; & \text{elsewhere} \end{cases}$

Solution:

X and Y are independent iff $f(x,y)=g(x).h(y) \; \forall \; x,y$
Consider,

$$g(x) = 2\int_x^\infty (e^{-x-y})dy = 2e^{-2x}$$
$$h(y) = 2\int_0^y (e^{-x-y})dx = -2e^{-2y} + 2e^{-2y}$$
$$g(x).h(y) = 2e^{-2x}(-2e^{-2y} + 2e^{-2y}) \neq f(x,y)$$

Hence X and Y are not independent.

7. Test for the independence of the rvs X and Y, given the joint pdf
$$f(x, y) = \begin{cases} 8xy \ ; & 0 < x < y < 1 \\ 0 \ ; & \text{elsewhere} \end{cases}$$

Solve it!

8. Test for independence of the rvs X and Y whose joint probability function is given below:

X/Y	1	2	3
1	1/8	1/8	2/8
2	3/8	0	0
3	0	1/8	0

Solution:

X/Y	1	2	3	$P(X=x_i)$
1	1/8	1/8	2/8	4/8
2	3/8	0	0	3/8
3	0	1/8	0	1/8
$P(Y=y_j)$	4/8	2/8	4/8	1

$$P(Y = 1) = \frac{1}{8} + \frac{3}{8} = \frac{4}{8} = \frac{1}{2}, \qquad P(X = 1) = \frac{4}{8} = \frac{1}{2}$$
$$P(Y = 2) = \frac{2}{8}, \qquad P(X = 2) = \frac{3}{8}$$
$$P(Y = 3) = \frac{3}{8}, \qquad P(X = 3) = \frac{1}{8}$$

Now for X and Y to be independent we must have $P(X=x_i, Y=y_j) = P(X=x_i)\, P(Y=y_j) \quad \forall\ i,j$.
Consider,

$$P(X = 1, Y = 1) = \frac{1}{8} \text{ and } P(X = 1).P(Y = 1) = \frac{1}{2}.\frac{1}{2} = \frac{1}{4}$$
$$P(X = 1, Y = 1) \neq P(X = 1).P(Y = 1)$$

Hence X and Y are not independent.

9. Test for the independence of the rvs X and Y whose joint probability function is given below:

Y/X	-1	0	1
-1	1/12	1/12	2/12
0	0	0	0
1	2/12	2/12	4/12

Solve it!

10. Suppose X and Y are independent rvs and X takes values 2, 5 and 7 with probabilities $\frac{1}{2}, \frac{1}{4}$ and $\frac{1}{4}$ respectively while Y takes values 3 and 5 with probabilities $\frac{1}{3}$ and $\frac{2}{3}$ respectively. Determine the joint pmf of X and Y. Also determine the probability distribution of Z=X+Y?

Solution:

Since X and Y are independent, we have $P(X=x_i, Y=y_j) = P(X=x_i)\, P(Y=y_j)\ \ \forall\, i,j$.

Y/X	2	5	7	$P(Y=y_j)$
3	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{3}$
5	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{2}{3}$
$P(X=x_i)$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	1

To find the probability distribution of Z=X+Y

Note that Z takes values 5,8,10,7,12

$P(Z=5) = P(X=2,Y=3) = \frac{1}{6}$

$P(Z=8) = P(X=5,Y=3) = \frac{1}{12}$

$P(Z=10) = P(X=5,Y=5) + P(x=7,Y=3) = \frac{1}{6} + \frac{1}{12} = \frac{1}{4}$

$P(Z=7) = P(X=2,Y=5) = \frac{1}{3}$

$P(Z=12) = P(X=7,Y=5) = \frac{1}{6}$

Joint pmf: $P(X=x_i,Y=y_j) = P(X=x_i).P(Y=y_j)$; since X and Y are independent

$$P_{11} = P(X=2, Y=3) = P(X=2).P(Y=3) = \frac{1}{2}.\frac{1}{3} = \frac{1}{6}$$

$$P_{12} = P(X=2, Y=5) = P(X=2).P(Y=5) = \frac{1}{2}.\frac{2}{3} = \frac{1}{3}$$

$$P_{21} = P(X=5, Y=3) = P(X=5).P(Y=3) = \frac{1}{4}.\frac{1}{3} = \frac{1}{12}$$

$$P_{22} = P(X=5, Y=5) = P(X=5).P(Y=5) = \frac{1}{4}.\frac{2}{3} = \frac{1}{6}$$

$$P_{31} = P(X=7, Y=3) = P(X=7).P(Y=3) = \frac{1}{4}.\frac{1}{3} = \frac{1}{12}$$

$$P_{32} = P(X=7, Y=5) = P(X=7).P(Y=5) = \frac{1}{4}.\frac{2}{3} = \frac{1}{6}$$

Mathematical Expectation or Expected Value of a random variable

Let X be a rv. Then the mathematical expectation of X is defined as

$$E(X) = \begin{cases} \sum_{i=1}^{\infty} x_i p(x_i) \ ; \ \text{if x is a drv.} \\ \int_{-\infty}^{\infty} x f(x) dx \ ; \text{if x is a crv} \end{cases}$$

Properties:

(i). $E(C) = C$
(ii). $E(CX) = CE(X)$
(iii). $E(X+Y) = E(X)+E(Y)$
Generalization:
$E(X_1 + X_2 \ldots .. + X_n) = E(X_1) + E(X_2) + \ldots + E(X_n)$
i.e. $E(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} E(X_i)$
(iv). $E(XY) = E(X).E(Y)$ iff X and Y are independent

Variance:
Let X be a rv. Then the variance of X, denoted by σ^2, is given by
$V(X) = E(X - E(X))^2$,
$= E(X^2) - (E(X))^2$

Properties:
(i). $V(C)=0$
(ii). $V(CX) = C^2 V(X)$
(iii). $V(X+Y) = V(X) + V(Y)$ iff X and Y independent

$$= E(X+Y)^2 - (E(X+Y))^2$$
$$= E(X^2) + E(Y^2) + 2E(XY) - \left(E(X)\right)^2 - \left(E(Y)\right)^2 - 2E(X)E(Y)$$
$$= E(X^2) - \left(E(X)\right)^2 + E(Y^2) - \left(E(Y)\right)^2 + 2[E(XY) - E(X)E(Y)]$$
$$= V(X)+V(Y)+0$$

$\therefore V(X+Y) = V(X) + V(Y)$
Also $V(aX+bY) = a^2V(X)+b^2V(Y)$ iff X and Y are independent

Covariance of (XY)
$Cov(XY) = E[(X-E(X))(Y-E(Y))]$
$= E(XY)-E(X).E(Y)$
Note 1: If X and Y are independent then $Cov(XY)=0$. However, $Cov(XY)=0$ does not necessarily mean that the rvs are independent.
Note 2: If X and Y are any two rvs then we have
$V(X+Y)=V(X)+V(Y)+2[E(XY)-E(X)E(Y)]=V(X)+V(Y)+2Cov(XY)$

Also, $V(aX + bY) = a^2V(X) + b^2V(Y) + 2ab \ Cov(XY)$

Mean and Variance of Binomial distribution:

Let X~B(n,p)

$P(X = x) = \binom{n}{x} p^x q^{n-x}$; x=0,1,2,…,n

$E(X) = \sum_{x=0}^{n} xp(x) = \cdots = np$

Thus mean of B.D is np

Now, V(X)= $E(X^2) - (E(X))^2$

where, $E(X^2) = \sum_{x=0}^{n} x^2 \binom{n}{x} p^x q^{n-x} = \cdots = n(n-1)p^2 + np$

$\therefore V(X) = n(n-1)p^2 + np - (np)^2$

$= npq$

Thus for the B.D. mean > variance; for 0<p<1

Mean and variance of Uniform distribution:

Let X~U[a,b]

We have $f(x) = \begin{cases} \frac{1}{b-a}; & a \leq x \leq b \\ 0 \quad ; & \text{elsewhere} \end{cases}$

The mean of uniform distribution is

E(X) = $\int_{-\infty}^{\infty} x. f(x)dx = \int_a^b x \frac{1}{b-a} dx = \frac{a+b}{2}$

Now, V(X) = $E(X^2) - (E(X))^2$

where, $E(X^2) = \int_a^b x^2 \frac{1}{b-a} dx = \frac{1}{3}(b^2 + ab + a^2)$

Thus V(X) = $\frac{1}{3}(b^2 + ab + a^2) - \frac{(a+b)^2}{4}$

V(X) = $\frac{1}{12}(b-a)^2$

Chebyshev's inequality:

It will give us a means of understanding precisely how the variance measures the variability about the expected value of a rv. If we know the probability distribution of a rv we may then compute E(X) and V(X), however the converse is NOT TRUE. Nevertheless, we can give a very useful upper or lower bound to such a probability.

Let X be a rv with $E(X) = \mu$ and let 'c' be any real number. Then, if E(X-c)<∞ and $\varepsilon > 0$, we have

$$P\{|X - c| \geq \varepsilon\} \leq \frac{1}{\varepsilon^2} E(X - c)^2$$

This is known as Chebyshev's inequality

Prove it!

Alternate forms:

(i). $P(|X - c| < \varepsilon) \geq 1 - \frac{1}{\varepsilon^2} E(X - c)^2$ (Complimentary event)

(ii). If $c = \mu$ ($=$ mean) then $P(|X - \mu| \geq \varepsilon) \leq \frac{V(X)}{\varepsilon^2}$

(iii). If $c = \mu, \varepsilon = k\sigma;\ \sigma^2 = V(X)$ then $P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$

If variance V(X) is small, most of the probability distribution of X is concentrated near the mean, $E(X) = \mu$ and when V(X)=0 we have X=E(X), i.e. all the values assumed by the rv X coincide with its mean.

Correlation Coefficient:

Correlation coefficient between two rvs X and Y denoted by ρ_{XY} or ρ is a measure of the degree of linearity or linear relationship between two rvs and is given by,

$$\rho = \frac{Cov(XY)}{\sqrt{V(X)V(Y)}} = \frac{E(XY) - E(X)E(Y)}{\sigma_X \sigma_Y}$$

If $\rho = 0$ then the rvs are said to be uncorrelated. But the converse need not be true i.e. if the rvs are uncorrelated then they need not be independent.

If $\rho = \pm 1$ then the rvs are said to be perfectly correlated.

$-$ve values of $\rho \Rightarrow X \uparrow Y \downarrow$

$+$ve values of $\rho \Rightarrow X \uparrow Y \uparrow$

Theorem 1:

The correlation coefficient between two rvs lies between -1 and +1.

i.e. $-1 \leq \rho \leq +1$,

Prove it!

Theorem 2:
The Correlation coefficient is independent of change of origin and scale.
i.e., $\rho_{UV} = \pm\rho_{XY}$; where U=a+bX; V=c+dY

Prove it!

Theorem 3:
If X and Y are linearly related, then $\rho = \pm 1 (\equiv \rho^2 = 1)$ and conversely.

Prove it!

Problems

1. To Show that Cov(XY)=0 does not necessarily imply that the rv are independent

Solution:
Let X be a rv with pdf

$$X \sim f(x) = \begin{cases} \frac{1}{2}; & -1 \leq x \leq 1 \\ 0; & \text{elsewhere} \end{cases}$$ (Can you identify the distribution?)

Let $Y=X^2$ (dependence is quadratic)

Now $E(X) = \int_{-\infty}^{\infty} xf(x)dx = \int_{-1}^{1} x\frac{1}{2}dx = 0$
$E(XY)=E(X^3) = \int_{-1}^{1} x^3 f(x)dx = \int_{-1}^{1} \frac{x^3}{2} dx = 0$
$\therefore \text{Cov}(XY) = E(XY) - E(X)E(Y) = 0$

2. Find the mean and the variance of probability distribution where $g(0) = \frac{16}{31}$, $g(1) = \frac{8}{31}$, $g(2) = \frac{4}{31}$, $g(3) = \frac{2}{31}$, $g(4) = \frac{1}{31}$

Solution:

Given function is a probability distribution (discrete)

Mean = E(X) = $\sum_{x=0}^{4} xg(x)$ = $0.\frac{16}{31} + 1.\frac{8}{31} + 2.\frac{4}{31} + 3.\frac{2}{31} + 4.\frac{1}{31} = \frac{26}{31}$

Variance = V(X) = $E(X^2) - (E(X))^2$

where, $E(X^2) = \sum_{x=0}^{4} x^2 g(x) = 0.\frac{16}{31} + 1.\frac{8}{31} + 4.\frac{4}{31} + 9.\frac{2}{31} + 16.\frac{1}{31} = \frac{58}{31}$

$\therefore V(X) = \frac{58}{31} - \left(\frac{26}{31}\right)^2 = \frac{1122}{961}$

3. A fair die is tossed 72 times. Given that X: No. of times 6 appears. Evaluate $E(X^2)$?

   Solve it!

4. Suppose a rv X has mean 10 and variance 25, for what +ve values of 'a' and 'b' does the rv Y= aX-b have expectation (mean) 0 and variance 1.

   Solve it!

5. The rv (X,Y) has a joint pdf given by $f(x,y) = \begin{cases} x+y; & 0 \le (x,y) \le 1 \\ 0; & \text{elsewhere} \end{cases}$.

   Find correlation coefficient ρ_{XY}.

Solution: We have

$$\rho_{XY} = \frac{Cov(XY)}{\sqrt{V(X)V(Y)}} = \frac{E(XY) - E(X)E(Y)}{\sqrt{V(X)V(Y)}}$$

$$E(XY) = \int_0^1 \int_0^1 xy(x+y)dxdy = \int_0^1 y\left[\frac{x^3}{3} + \frac{x^2y}{2}\right]_0^1 dy = \int_0^1 y\left[\frac{1}{3} + \frac{y}{x}\right]$$

$$= \int_0^1 \left[\frac{4}{3} + \frac{y^2}{x}\right]dy = \frac{y^2}{6} + \frac{y^3}{6}\Big]_0^1 = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

$$E(X) = \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} f(x,y)dydx = \int_{-\infty}^{\infty} xg(x)dx$$

where, $g(x) = \int_0^1 (x+y)dy = x + \frac{1}{2}$

$$\therefore E(X) = \int_0^1 x\left(x + \frac{1}{2}\right)dx = \left[\frac{x^3}{3} + \frac{x^2}{4}\right]_0^1 = \frac{1}{3} + \frac{1}{4} = \frac{7}{12}$$

$$E(Y) = \int_0^1 yh(y)dy$$

$$h(y) = \int_0^1 (x+y)dx = y + \frac{1}{2}$$

$$\therefore E(Y) = \int_0^1 y\left(y + \frac{1}{2}\right) = \frac{7}{12}$$

$$V(X) = E(X^2) - (E(X))^2$$

$$E(X^2) = \int_0^1 x^2\left(x + \frac{1}{2}\right)dx = \frac{x^4}{6} + \frac{x^2}{4} = \frac{10}{24} = \frac{5}{12}$$

$$V(X) = \frac{5}{12} - \frac{49}{144} = \frac{11}{144} = V(Y)$$

$$\rho_{XY} = \frac{\left(\frac{1}{3}\right) - \left(\frac{7}{12}\right)\left(\frac{7}{12}\right)}{\sqrt{\left(\frac{11}{144}\right)\left(\frac{11}{144}\right)}} = \frac{\left(\frac{1}{3}\right) - \left(\frac{49}{144}\right)}{\frac{11}{144}} = \frac{\frac{-1}{144}}{\frac{11}{144}} = \frac{-1}{11}$$

6. Find ρ_{XY}, given the joint pdf of (X,Y) as $f(xy) = \begin{cases} 2 - x - y; & 0 \le (xy) \le 1 \\ 0 & ; \text{otherwise} \end{cases}$

Solve it!

7. Two independent variates X_1,X_2 have means 5, 10 and variances 4, 9 respectively. Find covariance between $U=3X_1+4X_2$; $V=3X_1-X_2$

Solve it!

8. Let X_1,X_2,X_3 be uncorrelated rvs having same standard deviation. Find the correlation coefficient between U and V where $U=X_1+X_2$ and $V=X_2+X_3$

Solve it!

9. If X,Y,Z are uncorrelated rvs with V(X)=25, V(Y)=144 and V(Z)=81. Find ρ_{UV} where U=X+Y and V=Y+Z.

Solve it!

PROBABILITY DISTRIBUTION

Discrete Probability Distributions

Bernoulli Distribution:

A drv X is said to follow the Bernoulli distribution with parameter p, if it assumes only non-negative values 0 and 1 and its probability function is given by
$P(X=x) = p^{x} q^{1-x}$; x=0,1 and p+q= 1

Now, to obtain the mean and variance of the Bernoulli distribution, consider
$E(X) = \sum_x xp(x) = \sum_0^1 xp^x q^{1-x} = p$ and
$V(X)=E(X^2)-(E(X))^2$
where $E(X^2)= \sum_x x^2 p(x) = \sum_0^1 x^2 p^x q^{1-x} = p$
$V(X)=p-p^2=p(1-p)=pq$

Binomial Distribution:

A drv X is said to follow the Binomial distribution with parameters n and p, if it assumes only non-negative values 0,1,2,3,……..n and its probability function is given by
$P(X=x)=\binom{n}{x} p^x q^{n-x}$; x=0,1,2,3,4,……,n and p+q=1
Note: $\sum_{x=0}^{n} P(X = x) = \sum_{x=0}^{n}\binom{n}{x} p^x(1-p)^{n-x} = [p + (1-p)]^n = 1$
Hence P(X=x) is the probability distribution.

Now, to obtain mean and variance of Binomial distribution:
$E(X)= \sum_x xp(x) = np$
and $V(X)=E(X^2)-(E(X))^2=npq$

Poisson Distribution:

A discrete random variable X is said to follow the Poisson distribution with parameter λ, ($\lambda > 0$) if it assumes only non-negative values 0,1,2,3,4,… and its probability function is given by
$P(X=x)= e^{-\lambda}\lambda^x / x!$, x=0,1,2,3,4,…
Note: $\sum_{x=0}^{\infty} P(X = x) = \sum_{x=0}^{\infty} e^{-\lambda}\lambda^x / x! = 1$
Hence P(X=x) is the probability distribution.

Now, to obtain mean and variance of Poisson distribution:
$E(X)= \sum_x xp(x) = \sum_{x=0}^{\infty} xe^{-\lambda}\lambda^x / x! = \ldots = \lambda$ (Prove it!)
and $V(X)=E(X^2)-(E(X))^2 = \ldots = \lambda$ (Prove it!)

Geometric Distribution:

A drv X is said to follow the geometric distribution with parameter p if it takes values 1,2,3,4,… and its probability function is given by
$P(X=x) = p\, q^{x-1}$; x=1,2,3,4,… and p+q=1

Now, to obtain mean and variance of Geometric distribution:

$E(X) = \sum_x xp(x) = \sum_{x=1}^{\infty} xp\, q^{x-1}$ =……= 1/p (Prove it!)

and $V(X) = E(X^2)-(E(X))^2$ =…= q/p^2 (Prove it!)

Uniform Distribution:

A drv X is said to follow the Uniform distribution if it takes values 1,2,3,…,n and its probability function is given by

P(X=x) = 1/n; x=1,2,3,…,n

We have, E(X) = (n+1)/2 and V(X) = $(n^2-1)/12$.

Poisson Distribution as an approximation to the Binomial Distribution / Poisson Distribution as a Limiting case of the Binomial Distribution

Now what happens to the binomial probability $\binom{n}{x} p^x q^{n-x}$, x=0,1,2,3,4,…n, when n becomes large i.e. $n \to \infty$? For this consider a rv X~B(n,p) we have,

$P(X=x) = \binom{n}{x} p^x q^{n-x}$, x=0,1,2,3,4,…n

Now we put the following conditions:

1) n, the number of trials is indefinitely large i.e. $n \to \infty$

2) p, the constant probability of success for each trial is indefinitely small i.e. $p \to 0$.

3) $np = \lambda$ say, is finite, so that $p = \lambda/n$ so that $q = 1-(\lambda/n)$

Under these conditions it can be shown that, (Prove it!)

$\lim_{n\to\infty} P(X=x) = \frac{e^{-\lambda}\lambda^x}{x!}$; $x = 0,1,2,\ldots n$; $n \to \infty$

Or $P(X=x) = \frac{e^{-\lambda}\lambda^x}{x!}$; $x = 0,1,2,\ldots$, which is nothing but the Poisson distribution.

Problems:

1) What is the probability of getting a 6 atleast once in two throws of fair die?

Solution: Let X=No. of times 6 appears.

Given n=2 and p=P(getting 6 in a single throw of a die)=1/6

Clearly X~B(n,p) with n=2; p=1/6

$P(X=x) = \binom{n}{x} p^x q^{n-x}$, x=0,1,2…,n

P(getting a 6 atleast once in 2 throws) = P(X=1) + P(X=2)

$= \binom{2}{1}(1/6)(5/6) + \binom{2}{2}(1/6)^2$

= 11 / 36.

2) 6 coins are tossed. Find the probability of getting **a)** exactly 3 heads **b)** atmost 3 heads **c)** atleast 3 heads **d)** atleast 1 head

Solution: Let X = No. of heads appearing; Given n=6, p=P(getting a H)= 1/2.
Clearly X~B(n,p) with n=6; p=1/2.
$P(X=x) = \binom{n}{x} p^x q^{n-x}$; x=0,1,2,…n.

a) P(exactly 3 heads) = P(X =3) = $\binom{6}{3}$(1/64) = 20/64 = 5/16.

b) P(atmost 3 heads) = P(X=0) + P(X=1) + P(X=2) + P(X=3)
$= \binom{6}{0}(1/64) + \binom{6}{1}(1/64) + \binom{6}{2}(1/64) + \binom{6}{3}(1/64)$
= 21 / 32.

c) P(atleast 3 heads) = P(X=3) + P(X=4) + P(X=5) + P(X=6)
$= \binom{6}{3}(1/64)+ \binom{6}{4}(1/64) + \binom{6}{5}(1/64) + \binom{6}{6}(1/64)$
= 21 / 32.

d) P(atleast one head) = P(X=1) + P(X=2) + ------ +P(X=6)
= (1/64) (6+15+20+15+6+1)
= 63 / 64

3) If X has a Poisson probability distribution with parameter λ, show that
i) P(X is even) = $\frac{1}{2}(1 + e^{-2\lambda})$
ii) P(X is odd) = $\frac{1}{2}(1- e^{-2\lambda})$

Solution: Since X~ P(λ) , we have
$P(X=k) = e^{-\lambda} \lambda^k / k!$; k=0,1,2,…
Now P(X is even) = P(X=0)+P(X=2)+P(X=4)+…

$$= \sum_{k=0}^{\infty} P(X = 2k)$$
$$= \sum_{k=0}^{\infty} e^{-\lambda} \lambda^{2k}/(2k)!$$
$$= e^{-\lambda} \sum_{k=0}^{\infty} \lambda^{2k}/(2k)! \longrightarrow (1)$$

Now Consider, $(e^{\lambda} + e^{-\lambda})/2$.

$$=1/2\left(\sum_{k=0}^{\infty} \lambda^k/k! + \sum_{k=0}^{\infty}(-\lambda)^k/k!\right)$$
$$=1/2\left(\sum_{k=0}^{\infty} 2\lambda^{2k}/(2k)!\right)$$
$$=\sum_{k=0}^{\infty} \lambda^{2k}/(2k)! \longrightarrow (2)$$

Using (2) in (1), we get
P(X is even) = $e^{-\lambda} [(e^{\lambda} + e^{-\lambda})/2]$

$=1/2[1 + e^{-2\lambda}]$

Similarly show that P(X is odd) = $\frac{1}{2}(1 - e^{-2\lambda})$

4) If X has a Poisson distribution with parameter λ and if P(X=0) = 0.2 , evaluate P(X>2).

Solve it!

5) Suppose that X has a Poisson distribution with parameter λ. If P(X=2) = $\frac{2}{3}$ P(X=1), evaluate P(X=0) and P(X=1)

Solve it!

6) 6 coins are tossed 6400 times. Using Poisson distribution obtain the approximate probability of getting 6 heads 100 times.

Solution: Let X= number of times 6 heads appearing
Given n=6400 and p= P(getting 6 heads in a throw of 6 coins)= $1/2^6$ = 1/64
Clearly X~B(n,p) with n=6400 p=1/64
$P(X=x)=\binom{n}{x}p^x q^{n-x}; x = 0,1,2,\ldots,n$
But since n is large and p is small we may use Poisson approximation, with λ=np=100
Thus we have, with $X \sim P(\lambda)$,
$P(X=x)=e^{-\lambda}\lambda^x/x!$; x=0,1,2,…
$\therefore$ P(getting 6 heads 100 times)=P(X=100) = $e^{-100}(100)^{100}/100!$

7) Suppose that the probability that an item produced by a particular machine is defective equals 0.2. If 10 items produced from this machine are selected at random. What is the probability that not more than one defective is found? Use the binomial and Poisson distributions and compare the answers.

Solve it!

8) Suppose that a container contains 10,000 particles. The probability that such a particle escapes from the container equals 0.0004. What is the probability that more than 5 such escapes occur? (assume that the various escapes are independent of one another).

Solve it!

Continuous Probability Distributions

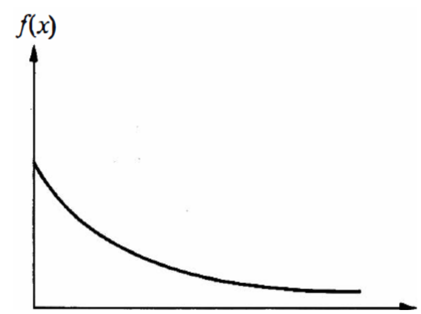
Uniform Distribution:
A crv X taking all possible values in the interval [a,b]; (a<∞, b<∞) is said to follow the Uniform distribution over [a,b], if its pdf is given by

$$f(x) = \begin{cases} \frac{1}{b-a}; & a \le x \le b \\ 0 \ ; & \text{otherwise} \end{cases}$$

Now, to obtain mean and variance of Uniform distribution
E(X)= $\int_a^b xf(x)dx$ = (b+a)/2 and
V(X) =E(X^2)-(E(X))^2=(b-a)^2/12.

Exponential Distribution:
A continuous rv X taking all possible non-negative values is said to follow the Exponential distribution with parameter λ, if its pdf is given by

$$f(x) = \begin{cases} \lambda e^{-\lambda x}; & \lambda > 0, x > 0 \\ 0 & ; \text{elsewhere} \end{cases}$$



Note: f(x) ≥ 0 ∀x and $\int_0^\infty f(x)dx = \int_0^\infty \lambda e^{-\lambda x} dx = 1$
Hence f(x) is a pdf

Now, to obtain mean and variance of exponential distribution
E(X) = $\int_0^\infty x\lambda e^{-\lambda x} dx = 1/\lambda$
and V(X) = E(X^2)-(E(X))^2 = $1/\lambda^2$

Normal Distribution:
A crv X taking all possible values in $(-\infty, \infty)$ is said to follow the Normal distribution with parameters μ and σ^2 $(\text{i.e.}\, X\sim N(\mu, \sigma^2))$, if its pdf is given by

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\ e^{-(1/2)[(x-\mu)/\sigma)^2]};\ -\infty < x < \infty, -\infty < \mu < \infty \text{ and } \sigma > 0$$

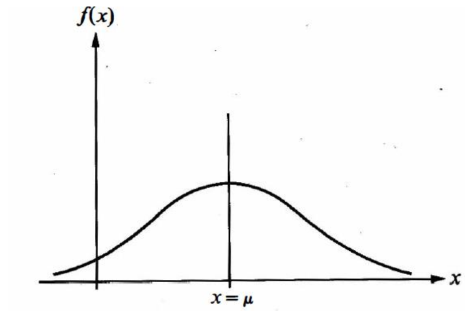
Properties of Normal distribution
1) $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\ e^{-(1/2)[(x-\mu)/\sigma)^2]}$ is a pdf

Prove it!

2) The mean and variance of normal distribution are given by
E(X) = $\int_{-\infty}^{\infty} xf(x)dx = \mu$ and V(X)=E(X^2)-(E(X))^2=σ^2

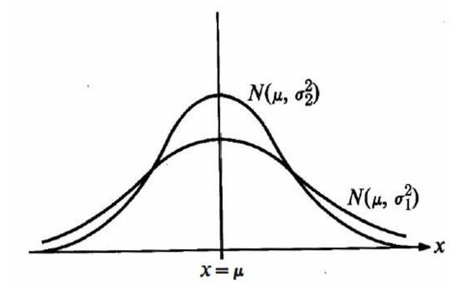
Prove it!

3) f(x) the pdf has the well-known bell shape.



4) Since f depends on x only through $(x-\mu)^2$, f(x) is symmetric about μ.

5) If the variance is large the spread will be more while if the variance is small the spread will be less, and the different values assumed by the r.v fall close to the mean.



6) $P(\mu-3\sigma \le X \le \mu+3\sigma) = 0.9973.$

Standard Normal Distribution:

If $X \sim N(\mu, \sigma^2)$, then the rv $Z = \frac{X-\mu}{\sigma}$ has a Standardized Normal Distribution (~N(0,1)) with mean=0 an variance=1 and Z is called Standardized Normal Variable (SNV). The pdf of Z is given by,

$\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}; -\infty < z < \infty$

The cdf of the SNV Z denoted by $\phi(z)$, is defined as

$\phi(z) = P(Z \le z) = \int_{-\infty}^{z} \varphi(t)dt$

Now suppose $X \sim N(\mu, \sigma^2)$ then what is $P(a \le X \le b)$? We have

$P(a \le X \le b) = \int_a^b f(x)dx$

where f(x) is the normal pdf. This cannot be integrated by ordinary methods.

Thus, we consider $P(a \le X \le b) = P\left(\frac{a-\mu}{\sigma} \le \frac{x-\mu}{\sigma} \le \frac{b-\mu}{\sigma}\right)$

$= P(c_1 \le Z \le c_2)$

$= \phi(c_2) - \phi(c_1)$ $\quad (\because P(a \le X \le b) = F(b) - F(a))$

Note 1: For different values of c_1, c_2, …, the values of $\phi(z)$ are tabulated in the standard normal tables.

Note 2: Also we have, $\phi(-z) = 1 - \phi(z)$.

Gamma Distribution:

A crv X taking all possible non-negative values is said to follow the Gamma distribution with parameters λ, α i.e. $X \sim \Gamma(\lambda, \alpha)$ where $\lambda > 0, \alpha > 0$ if its pdf is given by

$$f(x) = \begin{cases} \dfrac{\lambda^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}; & x > 0, \lambda > 0, \alpha > 0 \\ 0; & \text{elsewhere} \end{cases}$$

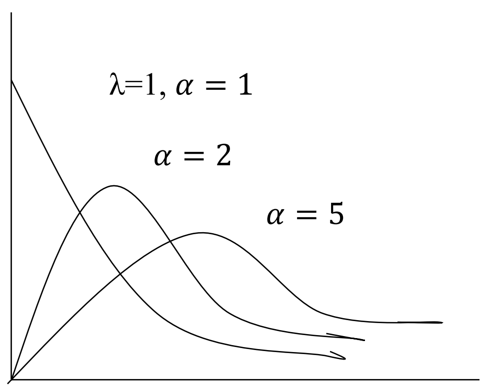
where $\Gamma(p) = \int_0^{\infty} x^{p-1} e^{-x} dx = (p-1)! \quad p > 0$

and $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

Now, to obtain mean and variance of Gamma distribution:

E(X) = $\int_0^{\infty} xf(x)dx = \frac{\alpha}{\lambda}$

and V(X) = E(X^2)-(E(X))^2 = $\frac{\alpha}{\lambda^2}$



When $\alpha = 1$, Gamma pdf reduces to exponential pdf with E(X) = $\frac{1}{\lambda}$ and $V(X) = \frac{1}{\lambda^2}$

Chi-square (χ^2) distribution:

A continuous rv X taking all possible non-negative values is said to follow the Chi-square (χ^2) distribution with n-degree of freedom (i.e. $X \sim \chi^2_{(n)}$), if its pdf is given by

$$f(x) = \begin{cases} \dfrac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}; & x > 0 \\ 0; & \text{Otherwise} \end{cases}$$

Now, to obtain mean and variance of Chi-square distribution:

E(X) = $\int_0^{\infty} xf(x)dx$ = n

and V(X) = E(X^2)-(E(X))^2 = 2n

When λ=1/2, $\alpha = n/2$, Gamma pdf reduces to Chi-square pdf with E(X) = n and $V(X) = 2n$

Problems

1. The diameter of an electric cable is normally distributed with mean 0.8 and variance 0.0004. What is the probability the diameter will exceed 0.81 units?

Solution:

If $X \sim N(\mu, \sigma^2), f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$

Now $X \sim N(0.8, 0.0004)$

To find P(X > 0.81) = ?

$\therefore P(X > 0.81) = P\left(\frac{X-\mu}{\sigma} > \frac{0.81-0.8}{0.02}\right)$

= P(Z > 0.5)

$= 1 - P(Z \leq 0.5)$

$= 1 - \phi(0.5)$

= 1-0.6915 (from SN tables)

= 0.3085

2. Let $X \sim N(2, 0.16)$. Find (i) $P(X \geq 2.3)$ (ii) $P(1.8 \leq X \leq 2.1)$

Solution:

(i) $P(X \geq 2.3) = P\left(\frac{X-\mu}{\sigma} \geq \frac{2.3-2.0}{04}\right)$

= $P(Z \geq 0.75)$

= 1- $P(Z \leq 0.75)$

= 1- $\phi(0.75)$

= 1- 0.7734 (from SN tables)

= 0.2266

(ii) $P(1.8 \leq X \leq 2.1)$ = ?

Solve it!

3. Suppose $X \sim N(\mu, \sigma^2)$. Determine 'c' as a function of μ and σ such that $P(X \leq c) = 2P(X \geq c)$.

Solution:

Consider $P(X \leq c) = 2P(X \geq c)$

$$P\left(\frac{X-\mu}{\sigma} \leq \frac{c-\mu}{\sigma}\right) = 2P\left(\frac{X-\mu}{\sigma} \geq \frac{c-\mu}{\sigma}\right)$$

$$P\left(Z \leq \frac{c-\mu}{\sigma}\right) = 2P\left(Z \geq \frac{c-\mu}{\sigma}\right) = 2\left[1 - P\left(Z \leq \frac{c-\mu}{\sigma}\right)\right]$$

$$\phi\left(\frac{c-\mu}{\sigma}\right) = 2\left[1 - \phi\left(\frac{c-\mu}{\sigma}\right)\right]$$

$$\Rightarrow \phi\left(\frac{c-\mu}{\sigma}\right) = 2/3 = 0.667$$

$$\Rightarrow \left(\frac{c-\mu}{\sigma}\right) = 0.43 \text{ (from SN tables)}$$

$$\Rightarrow c = \mu + 0.43\,\sigma$$

4. Let $T \sim N(50,4)$. Find $P(48 < T < 53)$?

   Solve it!

5. The lifetimes of two electronic devices are normally distributed with $D_1 \sim N(40,36)$ and $D_2 \sim N(45,9)$.
   a. If the device is to be used for a 45 hour period, which device to be preferred?
   b. If the device is to be used for a 48 hour period, which device to be preferred?

   Solve it!

6. In a normal distribution 31% of the items are under 45 and 48% are over 64. Find the mean and variance of the distribution.

   Solve it!

Table 3 THE STANDARD NORMAL DISTRIBUTION

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = P(X \le x)$$

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5369
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9131	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767

The Standard Normal Distribution (Continued)

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9989	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Moments and Moment Generating Function (mgf)

Moments:

Let X be a rv. Then its k^{th} moment about the mean is defined as

$$\mu_k = E(X - E(X))^k$$

$$= \begin{cases} \sum_x \left(x - E(X)\right)^k p(x) & ; \text{ if x is a drv} \\ \int_{-\infty}^{\infty} \left(x - E(X)\right)^k f(x)dx & ; \text{ if x is a crv} \end{cases}$$

The k^{th} moment about the origin is defined as

$$\mu'_k = E(X)^k$$

for k=2; $\mu_2 = V(X)$

Moment Generating Function (mgf)

Let X be a rv then the Moment Generating Function (mgf) is defined as

$M_X(t) = E(e^{tX})$ where t is a real variable

$$= \begin{cases} \sum_{x=0}^{\infty} e^{tx} p(x) ; & \text{if x is a drv} \\ \int_{-\infty}^{\infty} e^{tx} f(x)dx & ; \text{ if x is a crv} \end{cases}$$

$$M_X(t) = E(e^{tX})$$

$$= E\left[1 + \frac{tX}{1!} + \frac{t^2X^2}{2!} + \frac{t^3X^3}{3!} + \cdots + \frac{t^nX^n}{n!} + \cdots\right]$$

$$= E(1) + tE(X) + \frac{t^2}{2!}E(X^2) + \cdots + \frac{t^n}{n!}E(X^n) + \cdots$$

$$\therefore M_X(t) = 1 + t\mu'_1 + \frac{t^2}{2!}\mu'_2 + \cdots + \frac{t^n}{n!}\mu'_n + \cdots$$

i.e. the n^{th} order moment about the origin of any distribution is obtained as coefficient of $\left(\frac{t^n}{n!}\right)$ is to expansion of its mgf.

Differentiating w.r.t ‘t’ we get,

$$M'_X(t) = \mu'_1 + t\mu'_2 + \frac{t^2}{2!}\mu'_3 + \cdots + \frac{t^{n-1}}{(n-1)!}\mu'_n + \cdots$$

put t=0; $M'_X(0) = \mu'_1 = E(X)$

Differentiating again

$$M''_X(t) = \mu'_2 + t\mu'_3 + \cdots + \frac{t^{n-2}}{(n-2)!}\mu'_n + \cdots$$

put t=0; $M''_X(0) = \mu'_2$

$$\vdots$$

$$M^n_X(0) = \mu'_n$$

i.e. the n^{th} moment about the origin can be obtained by differentiating its mgf n times and putting $t = 0$.

Theorem:

If the rv X has the mgf M_X and Y is the rv $Y = \alpha X + \beta$, then mgf of Y is given by

$M_Y(t) = e^{\beta t} M_X(\alpha t)$, where α, β are constant.

Prove it!

Result:

If X and Y are two rvs with mgfs $M_X(t)$ and $M_Y(t)$ and if $M_X(t) = M_Y(t)\ \forall\ t$. Then their probability distribution must be the same.

Theorem:

Let X and Y be the two independent rvs and let Z=X+Y, (sum of two independent rvs) then mgf of Z is given by

$M_Z(t) = M_X(t)M_Y(t)$

Prove it!

Generalization: Let $X_1, X_2 \ldots, X_n$ be n independent rvs. Let $Z = \sum_{i=1}^{n} X_i$, then

$M_Z(t) = \prod_{i=1}^{n} M_{X_i}(t)$

Now, obtain the MGFs of the following distributions:

1. Binomial distribution
2. Poisson distribution
3. Normal distribution
4. Chi Square distribution
5. Gamma distribution

Also obtain the mean and the variance of all the above distributions using their mgfs.

Problems

1. Let X be the outcome when a fair die is rolled. Find the mgf of X and hence obtain its mean and variance.

   Solve it!

2. Suppose that the mgf of a rv X is of the form $M_X(t) = (0.4e^t + 0.6)^8$. Find the mgf of the rv Y=3X+2 and also obtain E(X) using the mgf.

   Solve it!

3. Let the rv X has the pdf

$$f(x) = \begin{cases} \lambda e^{-\lambda(x-\alpha)}; & x \geq a \\ 0; & \text{Otherwise} \end{cases}$$

   find mgf of X and hence mean and variance.

   Solve it!

4. A rv X has a pdf $f(x) = \frac{1}{2} e^{-|x|}; -\infty < x < \infty$. Obtain the mgf of X and hence the mean and the variance.

   Solve it!

5. Find the mgf of a rv X which is uniformly distributed over [-a,a] and hence evaluate $E(X^{2n})$?

   Solve it!

6. A rv X follows $X \sim N(\mu, \sigma^2)$. Show that $E(X-\mu)^{2n} = 1.3.5. \ldots (2n-1)\sigma^{2n}\ ;\ n = 1,2, \ldots.$

   Solve it!

7. A rv X takes values 0,1,2,3 etc.. with $p(x)=ab^x$ where a and b are positive and a+b=1. Find the mgf of X. If $E(X)=m_1$ and $E(X^2)=m_2$, show that $m_2=m_1(2m_1+1)$

   Solve it!

Reproductive property (Additive property)

If two or more independent rvs having certain distribution are added, then the resulting rv also will have a distribution of the same type as its summands. This is known as the reproductive property.

Some of the probability distribution exhibit this property and some do not.

Check whether the following distributions exhibit the reproductive property

1. Binomial distribution
2. Poisson distribution
3. Normal distribution
4. Chi Square distribution
5. Exponential distribution

STOCHASTIC PROCESSES

Introduction

Most of the naturally occurring phenomena show randomness in their behavior. They depend heavily on chance and do not follow the deterministic laws formulated by man. A natural mathematical tool to describe such a phenomena is the Random Function defined on a suitable parameter space. The prediction of the future, though possible only in a statistical sense, is achieved by the study of the evolution of random function with reference to the characteristics of the parameter.

A Stochastic Process is the family of random variables $\{X(t); t \in T\}$ indexed by parameter t varying in an index set T.

In order to describe a stochastic process, one has to specify the probability measure associated with the process that characterizes it.

Usually, the probability measure is specified by joint pdf, $f(x_1, x_2, \ldots, x_n)$.

In a stochastic process $\{X(t); t \in T\}$ generally 't' denotes time, but other parameters like space, length, width, area etc. are also used. The values assumed by the process are called the states and the set of all possible values is called the state space denoted by 'S'. Set of possible values of the parameter is called the parameter space denoted by 'T'

Both the Parameter Space (PS) and State Space (SS) may be discrete or continuous. Accordingly, stochastic processes are classified as:

(i) Discrete PS, discrete SS.
Eg: Consumer preference observed on a monthly basis

(ii) Continuous PS, discrete SS
Eg: No. of telephone calls arriving at an exchange over a period of time

(iii) Discrete PS, Continuous SS
Eg: Inventory on hand is observed only at discrete time points

(iv) Continous PS, Continous SS
Eg: Water level/ content of a dam being observed over a period of time

Example:
What are the SS and PS for a stochastic process which is the score during a football match?
Answer: T : {[0,90]}

S : {(x,y) : x,y = 0,1,2,……}

Types of Stochastic processes:

Quite often we find that X_n, the member of the family are not mutually independent. The nature of their dependence varies.

Stochastic processes are classified according the nature of dependence that exists amongst the members of the family.

I. Processes with independent increments

A Stochastic process $\{X(t); t \in T\}$ such that for any $(t_1, t_2, t_3, \ldots t_n) \in T$ where $t_1 < t_2 < t_3 < \ldots < t_n$, the rvs $[X(t_2) - X(t_1)]$; $[X(t_3) - X(t_2)]; \ldots ; [X(t_n) - X(t_{n-1})]$ are all independent is called a process with independent increments.

II. Stationary Processes

A Stochastic process $\{X(t) ; t \in T\}$ is said to be stationary if for any $(t_1, t_2, t_3, \ldots t_n) \in T$ the joint pdf of $\{X(t_1 + h), X(t_2 + h), \ldots X(t_n + h)\}$ and $\{X(t_1), X(t_2), \ldots, X(t_n)\}$ are the same $\forall\, h \in (-\infty, \infty)$

A stationary process is said to be strictly stationary if it is stationary of order n for any integer 'n'. Otherwise the process is said to be weakly or covariance stationary.

A process which is not stationary in any sense is called evolutionary.

III. Markov Processes

A Stochastic process $\{X(t); t \in T\}$ is called a Markov process if for any $t_1 < t_2 < t_3 < \ldots < t_n$, the conditional probability distribution of $X(t_n)$ for given values $X(t_1), X(t_2), \ldots, X(t_{n-1})$ depends only on $X(t_{n-1})$ i.e most recent known value of the process.

$$P[X(t_n) = x_n / X(t_{n-1}) = x_{n-1}, X(t_{n-2}) = x_{n-2}, \ldots, X(t_1) = x_1]$$
$$= P[X(t_n) = x_n / X(t_{n-1}) = x_{n-1}]$$

This property is called the Markov Property.

Markov Process is a Stochastic Process for which the future event or state depends only on

the immediately preceding event or state.

Every Stochastic process with independent increments is a Markov Process.

Markov Chain: A Markov Chain is a Markov process $\{X_n; n \in T\}$ with a Discrete PS and Discrete SS. i.e. the stochastic process $\{X_n; n = 0,1,2,...\}$ is a Markov chain if for $i, j, i_1, i_2, ..., i_{n-1} \in$ N(or I)

$$P[X_n = j/X_{n-1} = i, X_{n-2} = i_1, \ldots, X_0 = i_{n-1}] = P[X_n = j/X_{n-1} = i] = p_{ij}$$

where p_{ij} denotes probability that state a process at the n^{th} time point is j given that it was in state i in the (n-1) th time point. These p_{ij}'s are called One Step Transition Probabilities of the process, as p_{ij} refers to states i and j in 2 successive trials.

These p_{ij}'s may or may not be independent of 'n' (the time). p_{ij}'s may depend only on difference in time epochs (m-n) instead of n and m. If the p_{ij}'s are independent of 'n' then the process is said to be time homogeneous or simply homogeneous.

In more general case we are concerned with the states i and j at 2 non-successive trials, say state i at n^{th} trial and state j at $(n+m)^{th}$ trial. The corresponding transition probability is called the m-step transition probability, given by

$$p_{ij}^{(m)} = P[X_{n+m} = j/X_n = i]$$

Transition Probability Matrix (TPM):

The transition probabilities p_{ij} satisfies:

$p_{ij} \geq 0 \; \forall \; i, j$

$\sum_{j \in s} p_{ij} = 1 \; \forall \; i$

These probabilities may be conveniently written in the matrix form as,

$$P = \begin{array}{c} \\ i=0 \\ i=1 \\ \vdots \\ i=n \end{array} \begin{array}{c} \begin{array}{cccc} j=0 & j=1 & \dots & j=n \end{array} \\ \begin{bmatrix} p_{00} & p_{01} & \cdots & p_{0n} \\ p_{10} & p_{11} & \cdots & p_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ p_{n0} & p_{n1} & \cdots & p_{nn} \end{bmatrix} \end{array}$$

This 'P' is called the homogenous Transition Probability Matrix or simply TPM. This P is a Stochastic Matrix or Markov Matrix. [i.e. a square matrix with non-negative elements and unit row sums]. In addition, if the column sum is also unity, it is then called a Doubly Stochastic Matrix.

Similarly, for m- step transition probability we have.

$$P^{(m)} = \begin{array}{c} \\ i=0 \\ i=1 \\ . \\ . \\ i=n \end{array} \begin{array}{c} \begin{array}{cccc} j=0 & j=1 & \dots & j=n \end{array} \\ \begin{bmatrix} p_{00}^{(m)} & p_{01}^{(m)} & \dots & p_{0n}^{(m)} \\ p_{10}^{(m)} & p_{11}^{(m)} & \dots & p_{1n}^{(m)} \\ \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \dots & \vdots \\ p_{n0}^{(m)} & p_{n1}^{(m)} & \dots & p_{nn}^{(m)} \end{bmatrix} \end{array}$$

Order of a Markov Chain

A Markov Chain $\{ X_n ; n \in T \}$ is said to be of order s; s=1,2,3…..n,

if $\forall$ n, $P[X_n=j / X_{n-1}=i,\ldots, X_{n-s}=i_{s-1},\ldots] = P[X_n=j / X_{n-1}=i,\ldots, X_{n-s}=i_{s-1}]$

And a Markov Chain is said to be order 1 if above is $P[X_n=j / X_{n-1}=i]$ i.e s=1

Finite State Markov Chain

A Stochastic process $\{Xn ; n \in T \}$ is said to be a finite state Markov Chain if it has the following properties.

1. A finite number of states
2. Markovian property
3. Stationary or time homogeneous transition probabilities
4. Set of initial probabilities $P(X_0 = i) \; \forall \; i$

Thus, a tpm together with the initial probabilities associated with the states completely describes a Markov chain.

Note: The joint probability distribution of the stochastic process is given by,

$P\{X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n\}$

$= P(X_0 = i_0)\, P(X_1 = i_1 / X_0 = i_0)\, P(X_2 = i_2 / X_0 = i_0, X_1 = i_1) \ldots P(X_n = i_n / X_0 = i_0 \ldots, X_{n-1} = i_{n-1})$

$= P(X_0 = i_0)\, P(X_1 = i_1 / X_0 = i_0)\, P(X_2 = i_2 / X_1 = i_1) \ldots P(X_n = i_n / X_{n-1} = i_{n-1})$ ($\because$ Markovian Property)

The first factor on RHS i.e. $P(X_0 = i_0)$ is called the initial probability distribution of the chain and the remaining factors are the conditional probabilities which are the one step transition probabilities of the Markov Chain.

Hence, to be able to write down probability law, we need to know initial probability distribution of the chain as well as all one step transition probabilities.

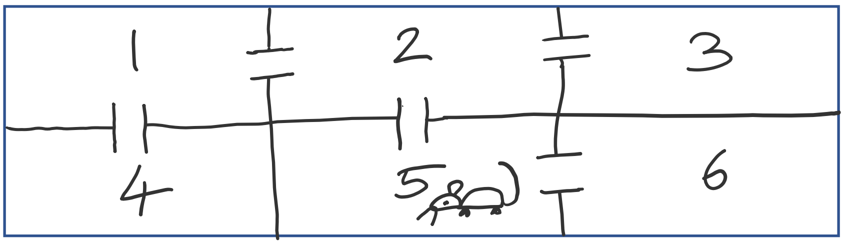
Problems

1) What is the TPM of a Markov Chain $\{Z_n\}$, where each Z_n is independently distributed as the rv Z, which has the distribution $P(Z=k) = p_k$?

Solution: Let $p_{ij} = P[Z_1=j / Z_0=i] = P(Z_1=j) = p_j$

$$\therefore\ T = \begin{array}{c} \\ i=0 \\ i=1 \\ . \\ . \\ i=n \end{array} \begin{array}{c} \begin{array}{cccc} j=0 & j=1 & \ldots & j=n \end{array} \\ \begin{bmatrix} p_0 & p_1 & \cdots & p_n \\ p_0 & p_1 & \cdots & p_n \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ p_0 & p_1 & \cdots & p_n \end{bmatrix} \end{array}$$

2) A rat is putting a maze as shown in the figure. At each time instant it changes its room choosing its exit at random. What is the TPM of the Markov Chain $\{Z_n\}$, where Z_n is the room the rat is occupying during the interval (n,n+1) ?



Solution:

$$\therefore\ T = \begin{array}{c} \\ i=1 \\ i=2 \\ i=3 \\ i=4 \\ i=5 \\ i=6 \end{array} \begin{array}{c} \begin{array}{cccccc} j=1 & j=2 & j=3 & j=4 & j=5 & j=6 \end{array} \\ \begin{bmatrix} 0 & 1/2 & 0 & 1/2 & 0 & 0 \\ 1/3 & 0 & 1/3 & 0 & 1/3 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \end{array}$$

3) A factory has two machines, but on any given day not more than one is in use. This machine has a constant probability 'p' of failure and if it fails the breakdown occurs at the end of the day's work. A single repairman is employed. It takes him two days to repair a machine and he works on only one machine at a time. Construct a stochastic process which will describe the working of the factory (i.e determine the state space and parameter space). List the possible transitions between the states of the factory and obtain the tpm of the Markov Chain $\{X_n\}$ where X_n is the number of days that would be needed to get both the machines back in working order and X_n is recorded at the end of day n. Also write the transition diagram.

Solve it!

Higher Transition Probabilities: Chapman – Kolmogorov Equations (C-K Equations)

The m- step transition probability is given by $p_{ij}^{(m)}$ = P [X_{n+m} = j/ X_n = i] where $p_{ij}^{(m)}$ gives the probability that from state i at the n^{th} trial the state j is reached at $(n+m)^{th}$ trial in m-steps. i.e it is the probability of transition from state i to state j in exactly m- steps. Since p_{ij} is time homogeneous, consider,

$$p_{ij}^{(2)} = P[X_{n+2} = j/ X_n = i]$$

Let k be some intermediate state. Consider a fixed value of k, then we have

$$p_{ij}^{(2)} = P[X_{n+2} = j/ X_n = i] = P[X_{n+2} = j, X_{n+1} = k/ X_n = i]$$

$$= P[X_{n+2} = j/ X_{n+1} = k, X_n = i].\ P[X_{n+1} = k/ X_n = i]$$

$$= P[X_{n+2} = j/ X_{n+1} = k].\ P[X_{n+1} = k/ X_n = i]$$

$$= p_{kj} \cdot p_{ik}$$

$$p_{ij}^{(2)} = p_{ik} \cdot p_{kj}$$

Since, k= 1,2,3,…… are all mutually exclusive we have

$$p_{ij}^{(2)} = \sum_{k \in S} p_{ik} p_{kj}$$

Thus, by mathematical induction we have;

$$p_{ij}^{(m+1)} = P[X_{n+m+1} = j/ X_n = i]$$

$$= \sum_k P[X_{n+m+1} = j/ X_{n+m} = k]\ P[X_{n+m} = k/ X_n = i]$$

$$= \sum_k p_{kj}\, p_{ik}^{(m)} = \sum_k p_{ik}^{(m)}\, p_{kj}$$

Thus in general we have,

$$p_{ij}^{(m+n)} = \sum_k p_{ik}^{(m)}\, p_{kj}^{(n)}$$

$$\equiv\ p_{ij}^{(n)} = \sum_{k \in S} p_{ik}^{(n-1)}\, p_{kj}$$

This is called the Chapman-Kolmogorov (CK) equation which is satisfied by the transition probabilities of a Markov Chain.

We can write these results in terms of transition probability matrices.

Let $P = (p_{ij})$ denote the one step tpm

$P^{(m)} = \left(p_{ij}^{(m)}\right)$ denote the m-step tpm

$P^{(2)} = \left(p_{ij}^{(2)}\right)$, where $p_{ij}^{(2)} = \sum_k p_{ik}p_{kj}$; i.e the elements of $P^{(2)}$ are the elements of the matrix obtained by multiplying P by itself.

$P^{(2)} = P.\ P$

Similarly, $P^{(m)} = P.\ P^{(m-1)} = P^{(m-1)}.\ P = P^m$

Thus, the m-step tpm is obtained by multiplying the one step tpm itself m-times. In other words, the probability of finding m-step tpm is one of finding the powers of the given one step tpm.

The CK equation are the basic equations in the study of Markov Process as they provide an efficient means of studying the m- step transition probabilities. i.e they enable us to build convenient relationship for transition probabilities between two points in the parameter space at which the process exhibits the Markov dependence property.

Problems

1) Let $\{ X_n ; n \geq 0 \}$ be a Markov Chain with states 0,1,2 and tpm

$$P = \begin{array}{c} \\ i=0 \\ i=1 \\ i=2 \end{array} \begin{array}{c} \begin{array}{ccc} j=0 & j=1 & j=2 \end{array} \\ \begin{bmatrix} 3/4 & 1/4 & 0 \\ 1/4 & 1/2 & 1/4 \\ 0 & 3/4 & 1/4 \end{bmatrix} \end{array}$$

and with initial probability distribution $P(X_0 = i) = 1/3, \forall i$

Find (i) $P[X_1 = 1/ X_0 = 2]$

(ii) $P[X_2 = 1/ X_1 = 1]$

(iii) $P[X_2 = 2/ X_1 = 1]$

(iv) $P[X_2 = 2, X_1=1/ X_0 = 2]$

(v) $P[X_2 = 2, X_1=1, X_0 = 2]$

(vi) $P[X_3 =1, X_2 = 2, X_1=1, X_0 = 2]$

Solution:

(i) $P[X_1 = 1/ X_0 = 2] = p_{21} = 3/4$

(ii) $P[X_2 = 1/ X_1 = 1] = p_{11} = 1/2$

(iii) $P[X_2 = 2/ X_1 = 1] = p_{12} = 1/4$

(iv) $P[X_2 = 2, X_1=1/ X_0 = 2] = P[X_2 = 2/ X_1=1, X_0 = 2]. P[X_1=1/ X_0 = 2]$

$= P[X_2 = 2/ X_1=1]. P[X_1=1/ X_0 = 2]$

$= p_{12}. p_{21} = (3/4)(1/4) = 3/16$

(v) $P[X_2 = 2, X_1=1, X_0 = 2] = P[X_2 = 2/ X_1=1, X_0 = 2]. P[X_1=1/ X_0 = 2]. P[X_0 = 2]$

$= p_{12}. p_{21}.(1/3) = (3/4)(1/4)(1/3) = 1/16$

(vi) $P[X_3 =1, X_2 = 2, X_1=1, X_0 = 2]$

$= P[X_3 =1, X_2 = 2, X_1=1/ X_0 = 2]. P[X_0 = 2]$

$= P[X_3 =1, X_2 = 2/ X_1=1].P[X_1 = 1/ X_0 = 2]. P[X_0 = 2]$

$= P[X_3 =1/ X_2 = 2].P[X_2 = 2/ X_1 = 1]. P[X_1 = 1/ X_0 = 2]. P[X_0 = 2]$

$= p_{21}.p_{12}.p_{21}.(1/3) = (3/4)(1/4)(3/4)(1/3) = (3/64)$

2) A Markov Chain $\{X_n ; n \geq 0\}$ has the tpm $P = \begin{bmatrix} x & y \\ 2x & y/3 \end{bmatrix}$ with initial probability distribution $p^0 = \{y,x\}$ Find p^0 ?

Solution:

Since P is a Markov matrix, we have $x+y = 1$ and $2x+ y/3=1$. Solving we get $y= 3/5$ and $x=2/5$

State occupation probabilities

The unconditional probabilities with which a Markov Chain occupies its various states are called state occupation probabilities. Let S= {1,2,3,….,m}. Then we define, $p_j^{(n)} = P[X_n = j]$ (i.e absolute probabilities at time 'n') and $p^{(n)} = \{p_1^{(n)}, p_2^{(n)}, \dots \dots, p_m^{(n)}\}$ is the absolute probability vector at time n. In particular $p^{(0)}$ is called the initial probability vector or initial probability distribution.

Result: It can be shown that $p^{(n)} = p^{(0)}. P^{(n)}$ or $\equiv p_j^{(n)} = \sum_{i=1}^{m} p_i^{(0)} p_{ij}^{(n)}$

Problems

1) Let $P = \begin{array}{c|ccc} & j=1 & j=2 & j=3 \\ \hline i=1 & 0 & 1 & 0 \\ i=2 & 1/4 & 1/2 & 1/4 \\ i=3 & 5/8 & 1/8 & 1/4 \end{array}$ be the tpm of a Markov Chain $\{X_n\}$

Find: (i) p_{13} (ii) $P[X_1=2/ X_0=2]$ (iii) $p_{12}^{(3)}$

Solution:

(i) $p_{13} = 0$

(ii) $P[X_1=2/ X_0=2] = p_{22} = ½$

(iii) $p_{12}^{(3)} \Longrightarrow$ It is the (1,2)th element of $P^{(3)}$ and we have $P^{(3)} = P^3$. Thus obtain P^3 as $P^3 = P^2.P$ and obtain $p_{12}^{(3)} = 17/32$

2) A particle performs a random walk with absorbing barriers at 0 & 4. Whenever it is at any position 'r' ($0 < r < 4$) it moves to r+1 with probability 'p' or to r-1 with probability 'q', so that p+q =1. As soon as it reaches 0 or 4 it remains there itself. Let X_n be the position of the particle after 'n' moves. Write the tpm of {Xn}.

Solve it!

3) Let M be a Markov Chain with tpm $P = \begin{bmatrix} 1/4 & 3/4 \\ 0 & 1 \end{bmatrix}$ and $p^{(0)} = \{1/3,2/3\}$, where $p^{(0)}$ is the initial probability distribution. Find (1) $P^{(2)}$ and (2) $p^{(2)}$

Solution:

(1) $P^{(2)} = P^2 = P.P = \begin{bmatrix} 1/16 & 15/16 \\ 0 & 1 \end{bmatrix}$

(2) $p^{(2)} = \{p_1^{(2)}, p_2^{(2)}\}$

Where $p_1^{(2)} = p_1^{(0)} p_{11}^{(2)} + p_2^{(0)} p_{21}^{(2)} = (1/3)(1/16) + (2/3)(0) = (1/48)$

$p_2^{(2)} = p_1^{(0)} p_{12}^{(2)} + p_2^{(0)} p_{22}^{(2)} = (1/3)(15/16) + (2/3)(1) = (47/48)$

$\therefore p^{(2)} = \{p_1^{(2)}, p_2^{(2)}\} = \{1/48, 47/48\}$

4) Consider a set of coin tossing experiments where the outcomes of the n^{th} trial are denoted by 1 for a Head and 0 for a tail. Let X_n be the rv denoting the outcome of the n^{th} trial and $S_n = X_1 + X_2 + X_3 + \ldots\ldots + X_n$ be the partial sum. The possible values of S_n are 0,1,2,3,…..,n. (i.e the states of S_n). S_n is a Markov Chain find its tpm.

Solve it!

5) Suppose that it has rained for the past two days, then it will rain tomorrow with a probability 0.7. If it rained today but not yesterday then it will rain tomorrow with a probability 0.5. If it rained yesterday but not today then it will rain tomorrow with a probability 0.4. If it has not rained for the past two days then it will rain tomorrow with a probability 0.2.
   (i). Write the tpm for the weather conditions.
   (ii). Given that it rained on Monday and Tuesday, what is the probability that it will rain on Thursday?
   (iii). Suppose it did not rain yesterday, nor the day before yesterday, what is the probability that it will rain tomorrow?

Solve it!

6) Customers in a city are continuously switching the brand of soap they use. If a customer is using brand A, the probability that he continues to use it the next week is 0.5 and the probabilities that he switches to brands B and C are 0.3 and 0.2 respectively. On the other hand, if he now uses brand B, the probability that he continues to use it is 0.6 while the probability that he switches to brand C is 0.4. Moreover, if he now uses brand C, the probability that he stays with it is 0.4 and the probabilities of switching to brands A and B are 0.2 and 0.4 respectively.
   (i). Write the tpm of the process
   (ii). If a customer is using brand A now, what is the probability that he will still be using it for two weeks?
   (iii). What brand of soap is most likely to be in use in two weeks?

Solve it!

7) A monkey is being trained to read the word 'BANANA'. A series of words are flashed on a screen. If it pushes the button when 'BANANA', appears he receives a banana, if he doesn't, he receives a shock (for an incorrect response). Suppose that the probability that a correct response is followed by an incorrect one is 0.1 while the probability that an incorrect response followed by a correct one is 0.3. If the probability that the first response is correct is 0.1, find the probability that
   (i). the second response is correct
   (ii). the third response is correct
   (iii). the second and the third responses are correct

Solve it!

Classification of states of Markov Chain:

If $p_{ij}^{(n)} > 0$ and $n \geq 0$ then we say that state j can be reached from state i [state j is accessible from state i ($i \overset{@}{\rightarrow} j$)]. If $i \overset{@}{\rightarrow} j$ and in addition $j \overset{@}{\rightarrow} i$, then states i and j are said to communicate

($j \overset{©}{\leftrightarrow} i$).

Reflexivity: Any state communicates with itself. i.e. $p_{ij}^{(n)} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$

Symmetry: If $i \overset{©}{\leftrightarrow} j$ then $j \overset{©}{\leftrightarrow} i$

Transitivity: If $i \overset{©}{\leftrightarrow} j$ and $j \overset{©}{\leftrightarrow} k$ then $i \overset{©}{\leftrightarrow} k$

If all the states communicate then the Markov chain is said to be Irreducible. i.e. every state of a Markov chain can be reached from every other state, after a finite number of transitions

The corresponding TPM is said to be an Irreducible TPM.

Let C be a set of states such that no state outside can be reached from any state in C. Then the set C is said to be closed. i.e. once the system is in one of the states of C it will continue to remain in C indefinitely.

If C is a closed set and if $i \in C$ and $j \notin C$, then $p_{ij}^{(n)} = 0$.

A special case of a closed set is single state j with transition probability $p_{jj} = 1$. Then, we call state j as an Absorbing State. If a state is an absorbing state, the process will never leave it once it enters it. All states of an irreducible Markov Chain must form a closed set.

Chains which are not irreducible are said to be non-irreducible or reducible or decomposable Markov Chains. Every Markov Chain must contain at least one closed set. If the number of closed sets is 2 or more, then the chain is said to be reducible.

The set of all states of a Markov Chain that communicate with each other are grouped into a class called the Equivalence class. A Markov Chain may have one or more such equivalence classes. If there are more than one equivalence class, then it is not possible to have communicating states in different equivalence classes. However, it is possible to have states in one class that are accessible from another class.

If a Markov Chain has all its states belonging to one equivalence class it is then said to be irreducible.

Problems:

1) Let P = $\begin{array}{c|cc} & j=1 & j=2 \\ \hline i=1 & 0.8 & 0.2 \\ i=2 & 0.2 & 0.8 \end{array}$

Solution:

Irreducible - since both states communicate

2) Are the following tpms irreducible or reducible? Justify your answer.

a) $P = \begin{array}{c} \\ i=1 \\ i=2 \\ i=3 \\ i=4 \\ i=5 \end{array} \begin{array}{c} \begin{array}{ccccc} j=1 & j=2 & j=3 & j=4 & j=5 \end{array} \\ \begin{bmatrix} 0.2 & 0.7 & 0.07 & 0.02 & 0.01 \\ 0.2 & 0.7 & 0.07 & 0.02 & 0.01 \\ 0 & 0.2 & 0.7 & 0.07 & 0.03 \\ 0 & 0 & 0.2 & 0.7 & 0.1 \\ 0 & 0 & 0 & 0.2 & 0.8 \end{bmatrix} \end{array}$

Solve it!

b) $P = \begin{array}{c} \\ i=1 \\ i=2 \\ i=3 \\ i=4 \\ i=5 \end{array} \begin{array}{c} \begin{array}{ccccc} j=1 & j=2 & j=3 & j=4 & j=5 \end{array} \\ \begin{bmatrix} 0.6 & 0.1 & 0 & 0.3 & 0 \\ 0.2 & 0.5 & 0.1 & 0.2 & 0 \\ 0.2 & 0.2 & 0.4 & 0.1 & 0.1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$

Solve it!

First Passage Time:

It is often desirable to make probabilistic statements about number of transitions made by the process in going from state i to state j for the first time. This length of time is called First Passage Time in going from state i to state j.

When j=i, this first passage time is just the number of transitions until the process returns to the initial state i. In this case the first passage time is called the First Return Time or Recurrence Time for state i.

The first passage times are in general are rvs and hence have probability distributions associated with them. These probability distributions depend upon the transition probabilities of the process.

Recurrent State (Persistent State):

A state i is said to be a Recurrent or Persistent if a return to state i is certain. i.e. a process from state i returns to state i with probability 1.

Transient State:

A state i is said to be Transient if a return to state i is uncertain. i.e. once the process is in state i it ever returns to state i.

Let μ_{ij} denote the Expected First Passage Time from state i to state j.

If j=i, then μ_{ii} is the Expected Recurrence Time.

A recurrent state is called Null-recurrent, if $\mu_{ii} = \infty$ and Non-null or Positive recurrent if μ_{ii} is finite ($\mu_{ii} < \infty$).

In a finite state Markov Chain there are no Null-recurrent states.

Periodic State:

A state is to be periodic with period t (t >1) if a return is possible only at t,2t,3t,… etc. steps, else Aperiodic or Non-periodic (say period 1 i.e. t=1).

A recurrent state is said to be Ergodic if it is Non-null and Aperiodic.

Note: A Markov chain is said to be Ergodic if it has Ergodic states.

Regular Matrix:

A stochastic matrix is said to be regular if some positive power of the matrix contains only positive entries. If there are zeros in the matrix they are eliminated during exponentiation.

Whenever a TPM is regular, we say that the chain is regular.

Result: If the TPM of a Markov chain is regular then the Markov chain is ergodic.

Rule to find whether a TPM is regular or not:

Suppose the TPM of a Markov chain is of order 'm' then the highest power to which the TPM must be raised to check whether it is regular or not is $m^2 - 2m + 2$.

Problems

1) Is the following matrix regular?

$P = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$

Solution:

Here P is not a stochastic matrix. Hence it is not a tpm. ∴ P is not regular

2) Is the following matrix regular?

$P = \begin{bmatrix} 1 & 0 \\ 1/2 & 1/2 \end{bmatrix}$

Solution:

Here $m^2 - 2m + 2 = 2^2 - 4 + 2 = 2$. So we raise P upto P^2.

Consider $P^2 = P.P = \begin{bmatrix} 1 & 0 \\ 3/4 & 1/4 \end{bmatrix}$ ∴ P is not regular

3) Are the following matrices regular?

a) $P = \begin{bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{bmatrix}$

Solve it!

b) $P = \begin{bmatrix} 0 & 0 & 1 \\ 1/2 & 1/4 & 1/4 \\ 0 & 1 & 0 \end{bmatrix}$

Solve it!

4) Show that the matrix $P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ a & 0 & 1 - a \end{bmatrix}$ is regular for $0 < a < 1$

Solve it!

Long-run behavior / Limiting behavior of Markov Chains:

Suppose a Markov system is in operation for a sufficiently long period of time and suppose that a large number of transitions take place during this period of time.

As the number of transitions increase the absolute probability (i.e. the probability that the system will be in state j) becomes independent of the initial state i or initial distribution.

In other words the system reaches the condition when the initial disturbances or start-up effects die out and the system may be regarded as having reached some kind of STEADY STATE OR EQUILIBRIUM STATE / DISTRIBUTION.

This called the Long-run behavior / limiting behavior of Markov Chains.

For an irreducible ergodic Markov Chain it can be shown than $\lim_{n\to\infty} p_{ij}^{(n)}$ exists and is independent of initial state i.

i.e. $\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j \ \forall\, j;\, j = 0,1,2, \ldots.$, where π_j exists uniquely satisfying following steady state equations:

$$\pi_j = \sum_{i=0}^{\infty} \pi_i\, p_{ij} \quad \forall\, j;\, j = 0,1,2, \ldots$$

$$\sum_{j=0}^{\infty} \pi_j = 1$$

$$\pi_j \geq 0 \ \forall\, j$$

The $\pi_j{}'s$ are called the steady state probabilities of the Markov Chain and are equal to reciprocal of expected recurrence time.

$$\pi_j = \frac{1}{\mu_{jj}} \ \forall\, j$$

Here, the term Steady State Probability means that the probability of finding the process in a certain state j after a large number of transitions tends to the value π_j which is independent of initial probability distribution defined over the states.

Note: Steady state probability does not mean that the process settles down into one of the states. In fact, the process continues to make transitions from state to state and at any step n, the transition probability from state i to state j is still p_{ij}.

Suppose j=0,1,2,…,m ; then we have the steady state equations with (m+2) equations in (m+1) unknowns. Since, it has a unique solution one of the equations must be redundant and hence can be deleted. This cannot be $\sum_{j=0}^{m} \pi_j = 1$, since $\pi_j = 0 \ \forall\, j$ will satisfy the other m+1 equations.

Thus, one of the equations $\pi_j = \sum_{i=0}^{m} \pi_i\, p_{ij}$ is redundant.

Thus, we have a set of linear equations which may be solved by any of the methods to obtain $\pi_j{}'s$, the steady state probabilities.

We may also write these steady state equations in the matrix form as $\Pi = \Pi. P$, where Π is a matrix with 'm' identical rows, each represented by a vector. $\Pi = (\pi_0, \pi_1, \pi_2, \dots \dots, \pi_m)$ and P is the tpm of the irreducible ergodic Markov Chain with (m+1) states.

Problems

1) Consider the tpm $P = \begin{bmatrix} 0.7 & 0.2 & 0.1 \\ 0.2 & 0.6 & 0.2 \\ 0.1 & 0.4 & 0.5 \end{bmatrix}$ Obtain the steady state probabilities.

Solution:

Calculating the powers of P we get

$P^2 = \begin{bmatrix} 0.54 & 0.30 & 0.16 \\ 0.28 & 0.48 & 0.24 \\ 0.20 & 0.46 & 0.34 \end{bmatrix}$; $P^4 = \begin{bmatrix} 0.4076 & 0.3796 & 0.2128 \\ 0.3336 & 0.4248 & 0.2416 \\ 0.3048 & 0.4372 & 0.2580 \end{bmatrix}$; P^8 =……..; P^{16} =…….; P^{32}=…….;

Note that although the rows of P are quite different, the rows of P^3, P^4, P^6, etc. are more and more similar. In fact if we were to calculate the higher powers of P we would see that P^n approaches L $= \begin{bmatrix} 0.353 & 0.412 & 0.235 \\ 0.353 & 0.412 & 0.235 \\ 0.353 & 0.412 & 0.235 \end{bmatrix}$, called the limiting matrix of P. Note that the rows of L are identical. This implies that, if n is large the probability of the process being in state 1 at time n does not depend upon weather the chain was initially in state 1,2 or 3.

For an ergodic Markov Chain such a limit exists. i.e $\lim_{n\to\infty} P^n = L = \Pi$ This Π is called the limiting distribution or equilibrium distribution or stationary distribution of the Markov Chain or the steady states of the chain or the vector of stable probabilities.

Note:

1) For a Markov Chain with μ_{jj}, the mean recurrence time, it can be shown that $\mu_{jj} = \frac{1}{\pi_j}$. This $\mu_{jj} < \infty$ if $\pi_j \neq 0$ and if $\pi_j = 0$ then $\mu_{jj} = \infty$, then state j is a Null-recurrent state and hence the Markov Chain is not Ergodic.

2) If the TPM of a Markov chain is doubly stochastic and has 'r' states, then the steady state distribution is $\Pi = (1/r, 1/r, 1/r, \dots \dots, 1/r)$

1) Consider the tpm P = $\begin{bmatrix} 0.2 & 0.8 \\ 0.6 & 0.4 \end{bmatrix}$ Obtain the steady state probabilities.

Solution:

We have $P^2 = \begin{bmatrix} 0.52 & 0.48 \\ 0.36 & 0.64 \end{bmatrix}$; $P^4 = \begin{bmatrix} 0.443 & 0.557 \\ 0.417 & 0.583 \end{bmatrix}$; $P^8 = \begin{bmatrix} 0.4281 & 0.5719 \\ 0.4274 & 0.5726 \end{bmatrix}$; P^{16} =…….; P^{32}=…….;

We note that it reaches equilibrium condition. Now consider the steady state equations with i = 0 and j = 1, we have

$\pi_0 = \pi_0 p_{00} + \pi_1 p_{10}$
$\pi_1 = \pi_0 p_{01} + \pi_1 p_{11}$
$\pi_0 + \pi_1 = 1$
Solving we get $\pi_0 = 0.4286$ and $\pi_1 = 0.5714$

Also, the mean recurrence times are obtained as $\mu_{00} = 2.333$ steps and $\mu_{11} = 1.75$ steps

2) Obtain the steady state probabilities given the following tpms.

(a). P = $\begin{bmatrix} 0.080 & 0.184 & 0.368 & 0.368 \\ 0.632 & 0.368 & 0 & 0 \\ 0.264 & 0.368 & 0.368 & 0 \\ 0.080 & 0.184 & 0.368 & 0.368 \end{bmatrix}$

Solve it!

(b). P = $\begin{bmatrix} 0 & 2/3 & 1/3 \\ 3/8 & 1/8 & 1/2 \\ 1/2 & 1/2 & 0 \end{bmatrix}$

Solve it!

(c). P = $\begin{bmatrix} 0 & 2/3 & 1/3 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{bmatrix}$

Solve it!

(d). P = $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ a & 0 & 1-a \end{bmatrix}$; 0 < a < 1

Solve it!

3) Is the Markov Chain with tpm $P = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/4 & 1/2 & 1/4 \\ 0 & 1/2 & 1/2 \end{bmatrix}$ ergodic? If so find its limiting distribution.

Solve it!

4) A Markov Chain has a tpm $P = \begin{bmatrix} 1/3 & 0 & 2/3 & 0 \\ 1/4 & 0 & 3/4 & 0 \\ 1/3 & 1/3 & 0 & 1/3 \\ 0 & 1/2 & 0 & 1/2 \end{bmatrix}$

In which state the chain is most likely to be found in the long run?

Solve it!

5) Consider $P = \begin{bmatrix} 1/2 & 1/4 & 1/4 \\ 1/2 & 1/4 & 1/4 \\ 0 & 1/2 & 1/2 \end{bmatrix}$ Obtain the steady state probabilities.

Solution:

The above matrix is a doubly stochastic matrix. In such a case, the steady state probabilities are given by, $\pi_j = \frac{1}{s}$ ∀ j=1,2,3,…..,s. where s is the number of states.

Thus, we have, $\pi_0 = \frac{1}{3}$, $\pi_1 = \frac{1}{3}$ and $\pi_2 = \frac{1}{3}$

6) Given $P = \begin{bmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{bmatrix}$ Obtain the steady state probabilities.

Solve it!

Continuous Parameter Markov Process

Quite often we come across situations where a continuous parameter is required.

Let $\{ X(t): t \geq 0 \}$ be a Markov Process with discrete states 0,1,2……..m (i.e. m+1 states) and stationary transition probability function

$p_{ij} = P\{X(t+s)=j/X(s)=i\} \ \forall \ i, j=0,1,2\ldots\ldots$

This function is assumed to be continuous at t=0 with

$$\lim_{t\to 0} p_{ij}(t) = \begin{cases} 1 & \text{if } \ i = j \\ 0 & \text{if } \ i \neq j \end{cases} \quad \forall \quad i, j = 0,1, 2\ldots.$$

This P_{ij} again satisfies the C-K equations i.e. for any state i & j and positive numbers

t & s ($0\leq s\leq t$), we have $p_{ij} = \sum_{k=o}^{m} p_{ik}(s)p_{kj}(t-s)$

Also, classifications of states are made as earlier. i.e. state i communicates with state j if $\exists \ t_1$ & $t_2 \ni p_{ij}(t_1)>0$ & $p_{ji}(t_2)>0$. All states that communicate are said to form a class. If all states in a chain form a single class (irreducible chain) then $p_{ij}(t)>0 \ \forall \ t>0$ & $\forall$ i & j. Further, $\lim_{t\to 0} p_{ij}(t)=\pi_j$ always exist and independent of the initial state i $\forall$ i=0, 1, 2…..m and these $\pi_j{}'$ s satisfy $\pi_j = \sum \pi_j p_{ij}(t) \ \forall \ j=0, 1, 2\ldots\ldots m$ and $t \geq 0$.

Poisson Process

Here we shall study some stochastic processes in continuous time with discrete state space. One such process is the Poisson Process. Let N(t) denote the number of events occurring in the interval of length t, say (0, t]. Then, the process {N(t)} is a stochastic process with state space S = {0,1, 2,….} and parameter space T = {t; t ≥ 0}. Then the process {N(t)} is called a counting process. Let $p_n(t)$ = P[N(t)=n]. We proceed to show that under certain conditions N(t) follows the Poisson distribution with parameter λt, where λ is a constant.

i.e. N(t) ~ P(λt) ⇒ $p_n(t)$ = P(N(t)=n) = $\frac{e^{-\lambda t}(\lambda t)^n}{n!}$; n = 0,1,2,3………..; $\lambda > 0$

The postulates for Poisson Process are:

1) Independence: N(t) is independent of the number of occurrences of the event in an interval prior to (0,t] i.e. {N(t)} has independent increments.

2) Homogeneity in time: $p_n(t)$ depends only on the length t of the time interval and is independent of where this time interval is located.

3) Regularity: In an interval of infinitesimal (very small) length 'h',
   P (exactly one occurrence) = λh + O(h)
   P (more than one occurrence) = O(h)

where O(h) → 0 more rapidly than h i.e. as h→0, $\frac{O(h)}{h} \to 0 \equiv \lim_{h\to 0}\frac{O(h)}{h} = 0$

Accordingly, we have $p_1(h)$ = P[N(h) = 1] = λh + O(h)

$p_0(h) = 1 - \lambda h$ + O(h)

$p_n(h)$ = O(h); n > 1

Under the above postulates it can be shown that N(t) follows a Poisson distribution with parameter λt

i.e. N(t) ~ P(λt) ⇒ $p_n(t)$ = P(N(t)=n) = $\frac{e^{-\lambda t}(\lambda t)^n}{n!}$; n = 0,1,2,3………..; $\lambda > 0$

i) Consider, when n ≥ 1

$p_n(t+h)$ = P[N(t+h)=n] = P[n occurrences by epoch t+h starting from t=0]

This is the probability that n events occurring by epoch t and no events occurring in (t,t+h) OR (n-1) events occurring by epoch t and one event occurring in (t,t+h) etc. etc. OR no events occurring by epoch t and n events occurring in (t,t+h)

i.e $p_n(t+h)$ = P[n occurrences by epoch t and no occurrence during h]

+P[n-1 occurrences by epoch t and one occurrence during h]

+ …………

$= p_n(t)\, p_0(h) + p_{n-1}(t)\, p_1(h) + \ldots.. + p_0(t)\, p_n(h)$

$= p_n(t)[1-\lambda h+O(h)]+p_{n-1}(t)[\lambda h+O(h)]+O(h)$

$p_n(t+h) = p_n(t)[1-\lambda h]+p_{n-1}(t)[\lambda h]+O(h)$

$$\frac{(p_n(t+h) - p_n(t))}{h} = \frac{\lambda h(p_{n-1}(t) - p_n(t))}{h} + \frac{O(h)}{h} = \lambda(p_{n-1}(t) - p_n(t)) + \frac{O(h)}{h}$$

As $h \to 0$, $p_n'(t) = \lambda(p_{n-1}(t) - p_n(t))$; $n \geq 1$ ----------- **(1)**

(ii) When n = 0 we have,

$p_0(t+h) = P[N(t+h)=0] = p_0(t)\, p_0(h)$

$= p_0(t)[1-\lambda h+O(h)]$

$$\frac{(p_0(t+h) - p_0(t))}{h} = -\lambda\, p_0(t) + \frac{O(h)}{h}$$

As $h \to 0$, $p_0'(t) = -\lambda p_0(t)$; $n = 0$ ----------- **(2)**

Equations **(1)** and **(2)** are called differential difference equations which along with the given initial conditions completely specify the process.

Initial Conditions:

Assuming that the process starts at t=0 i.e. N(0) = 0, we have

$p_0(0) = P[N(0)=0] = 1$

$p_n(0) = 0$; $n \neq 0$

These differential difference equations may be solved by Laplace transformation technique or generating functions technique etc. and the solution is given by

$$p_n(t) = P(N(t)=n) = \frac{e^{-\lambda t}(\lambda t)^n}{n!} \; ; \quad n = 0,1,2,3\ldots\ldots\ldots.; \; \lambda > 0$$

which is the Poisson distribution with mean and variance λt i.e. if the mean number of occurrences in an interval of length t is λt then the mean number of occurrences per unit time is λ. λ is called the rate of occurrence.

Problems

1. Suppose that the customers arrive at a bank according to a Poisson process at a rate 3 per minute. What is the probability that in an interval of two minutes, the number of customers arriving is (i) exactly 4 (ii) greater than 4 (iii) less than 4?

    Solve it!

2. Let vehicles arrive at a junction according to a Poisson process at a rate (λ) 5 per hour. What is the probability that

    (i) Exactly 3 vehicles arrive at the traffic junction in an hour?

    (ii) more than 10 vehicles arrive in 2 hours?

    Solve it!

A Note on Differential-Difference Equations:

Difference Equations

Let f(n) be a function defined only for non-negative integral values of the argument n.
The first difference of f(n) is defined by the increment of f(n) and is denoted by $\Delta f(n)$, i.e. $\Delta f(n) = f(n+1) - f(n)$
The second and higher differences are defined by,
$\Delta^{k+1} f(n) = \Delta^k f(n+1) - \Delta^k f(n); k > 0$
By a Difference Equation we mean an equation involving a function evaluated at the arguments which differ by any of a fixed number of values.
Ex: 1) f(n+2)-f(n+1)-f(n) = 0
2) $a_0u_x+a_1u_{x+1}+\ldots+ a_ku_{x+k} = g(x)$

Differential-Difference Equations

Suppose that $u_n(t)$, n=1,2,3,…, is a function of t having a derivative $\frac{du_n(t)}{dt} = u_n'(t)$
An equation involving $u_n'(t)$, $u_n(t)$, $u_{n+1}(t)$, etc. is called a Differential-Difference Equation.
Ex: 1) $u_n'(t)= u_{n-1}(t)$; $t\geq 0$; n=1,2,3…
2) $p_n'(t) = -\lambda[p_n(t) - p_{n-1}(t)]$

To arrive at a solution for a set of differential difference equation several techniques exist, such as 1) Generating Functions Technique 2) Laplace Transforms Technique etc.

Further results associated with Poisson process

Result 1: The interval between two successive occurrences of a Poisson process {N(t)} having a parameter λt has a negative exponential distribution with mean $\frac{1}{\lambda}$

Result 2: Further it can be shown that X_1, X_2, X_3,….., the intervals between successive occurrences of events E_i and E_{i+1}, i= 1,2,3,….., for a Poisson process {N(t)} are all independent and have identical exponential distribution with mean $\frac{1}{\lambda}$.

Result 3: If the intervals between successive occurrences of events E_i are independently distributed with common exponential distribution with mean $\frac{1}{\lambda}$, then the events E_i form a Poisson process with mean λt.

Memoryless property or Forgetfulness property of exponential distribution:

The exponential distribution has the property that the time until the next occurrence of event is independent of the time that elapsed since the occurrence of the last event i.e. the future is independent of past i.e. the process forgets its past history.

Mathematically, if T is the random variable (Inter occurrence time which is exponentially distributed) then $P(T>t+s/T>s) = P(T > t)$, where s is the occurrence time of the last event.

Proof : Given $T \sim \text{Exp}(\alpha) \Longrightarrow f(t) = \alpha e^{-\alpha t}$; $t > 0$

and $F(t) = P(T \leq t) = 1 - e^{-\alpha t} \Longrightarrow P(T > t) = e^{-\alpha t}$

Consider $P(T>t+s/T>s) = \frac{P(T>t+s, T>s)}{P(T>s)} = \frac{P(T>t+s)}{P(T>s)} = \frac{e^{-\alpha(t+s)}}{e^{-\alpha s}} = e^{-\alpha t} = P(T > t)$

$\therefore P(T>t+s/T>s) = P(T > t)$

This property demonstrates that the process is completely random as it shows that the time that has elapsed since the occurrence of the last event has no effect on the time of occurrence of the next event.

Pure Birth Process

Consider situations where only arrivals take place i.e. the **customers join the queuing system but never leave**. In this case the arrivals may be thought of as the occurrence of events. Such a process is called a **Pure birth process or an Arrival process** (Here birth refers to the arrival of a new customer).

Example: State health department records the birth of new babies effective from a given date. i.e. Birth information for each baby is recorded.
Our objective is to obtain an expression for $P_n(t)$, the probability of n arrivals during an interval of length t (say (0, t]). It can be shown that,
$P_n(t) = \frac{(\lambda t)^n e^{-\lambda t}}{n!}$; n = 0,1,2…. ; $\lambda > 0$.
(where λ = arrival rate or rate of occurrence of events)

i.e. The distribution of $P_n(t)$ is Poisson with parameter λt.
It can be shown that the inter arrival times are exponentially distributed with parameter λ.

The above expression for $P_n(t)$ may be obtained under the assumptions or postulates made earlier. We may arrive at differential - difference equations, which may be solved using Laplace-transform techniques, to obtain $P_n(t)$.

Pure Death Process

Here we assume that the system starts with a given No. of customers, say N, who leave the facility at the rate μ after being serviced. **But no new customer is allowed to join the system.** Such a process is called a **pure death process** or **a departure process** (Here death refers to departure of a customer).

Example: Inventory situations may be modeled as pure death process. Say, the inventory consists of N items to start with and the items are withdrawn from the stock at a rate μ (units per unit time).
Let $q_n(t)$ = Probability of n departures during an interval of length t.
Analogous to pure birth process (with λ=μ here), solving the differential- difference equations we may show that

$q_n(t) = \frac{(\mu t)^n e^{-\mu t}}{n!}$; $\mu > 0$; n= 0,1,2,3,…..N−1

$q_N(t) = 1-\sum_{n=0}^{N-1} q_n(t)$; n = N

= P[N(all) customers departed during t]

i.e. The distribution of $q_n(t)$ is Poisson with parameter μt.
Also the inter departure times or service times are exponentially distributed with parameter μ.

Birth and Death Process

For a Pure Birth process we have

$p_k(h)$ = P[No. of births between t and t+h is k, given that the No. of births by epoch t is n]

= P[N(h)=k/N(t)=n]

$$= \begin{cases} \lambda h + O(h) & ;\ k = 1 \\ O(h) & ;\ k \geq 2 \\ 1 - \lambda h + O(h) & ;\ k = 0 \end{cases} \qquad (\lambda \text{ is the arrival rate}) \qquad \text{--------- (A)}$$

Similarly, for a Pure Death process we have

$q_k(h)$ = P[No. of deaths between t and t+h is k, given that the No. of deaths by epoch t is n]

$$= \begin{cases} \mu h + O(h) & ;\ k = 1 \\ O(h) & ;\ k \geq 2 \\ 1 - \mu h + O(h) & ;\ k = 0 \end{cases} \qquad (\mu \text{ is the departure rate}) \qquad \text{----------- (B)}$$

With (A) and (B) above together we have a **birth and death process.**

Let N(t) = No. of occurrences by epoch t starting from t = 0

Let $p_n(t)$ = P[N(t)=n]

Consider $p_n(t+h)$ = P[N(t+h)=n] = P[n occurrences by epoch t+h starting from t=0]

This can happen in the following mutually exclusive ways.

$p_n(t+h)$ = P[n occurrences by t and no births and no deaths during h]

+P[n-1 occurrences by t and one birth and no death during h]

+ P[n+1 occurrences by t and no birth and one death during h]

$= p_n(t)(1-\lambda h)(1-\mu h)+p_{n-1}(t)\lambda h+p_{n+1}(t)\mu h+O(h)$

for n > 0 we have $p_n(t+h)= p_n(t)(1-\lambda h-\mu h)+p_{n-1}(t)\lambda h+p_{n+1}(t)\mu h+O(h)$ --------------(1)

and for n=0, we have {Note: P[zero deaths during h] = 1},

$p_0(t+h)= p_0(t)(1-\lambda h).1+p_1(t)\ \mu h+O(h)$ ----------------(2)

From (1) and (2) we have,

$$\frac{p_n(t+h)- p_n(t)}{h} = -(\lambda+ \mu)p_n(t) + \mu\, p_{n+1}(t) + \lambda\, p_{n-1}(t)+ \frac{O(h)}{h}$$

$$\frac{p_0(t+h)- p_0(t)}{h} = -\lambda p_0(t) + \mu\, p_1(t) + \frac{O(h)}{h}$$

Taking limits as $h \to 0$ we get

$p'_n(t) = -(\lambda + \mu)p_n(t) + \mu\, p_{n+1}(t) + \lambda p_{n-1}(t)$; $n > 0$ -------------(3)

$p'_0(t) = -\lambda p_0(t) + \mu\, p_1(t)$; $n = 0$ ------------(4)

Equations (3) and (4) are the differential-difference equations of the Birth and Death process which play a vital role in the study of queuing theory.

QUEUING THEORY

Consider

I. <u>Commercial Service Systems:</u>

Where customers receive service from commercial organizations.

Example: 1. Barber Shop

2. Hotel / Cafeteria

3. Customers / Shoppers in a supermarket.

II. <u>Transportation Service Systems:</u>

Where vehicles receive service from such organizations.

Example: 1. Cars waiting at a traffic light

2. Planes waiting to land or take off

3. Trucks waiting to be loaded or unloaded.

III. <u>Business – Industrial Service Systems:</u>

Example: Machines waiting for repair.

IV. <u>Social Service Systems:</u>

Example: 1. Patients waiting for treatment in a hospital

2. Cases waiting to be tried in courts

3. Letters waiting to be typed by a secretary.

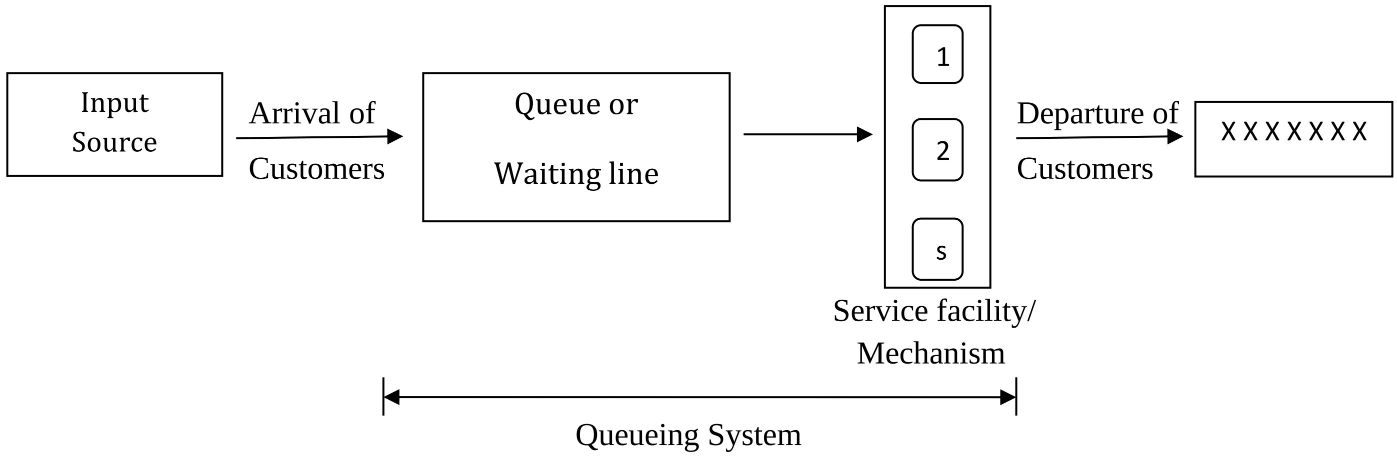
All the above examples have in common the phenomenon of **waiting**. In many of the real life situations we observe this phenomenon of **waiting**. The waiting phenomenon is the direct result of **randomness** in the operation of the service facilities.

In general, the customer's arrival & his service times are random. Our objective is to obtain some characteristics that measure the performance of the system understudy. Queuing theory deals with the mathematical study of 'queues' or 'waiting lines'.

<u>Some measures of performance are:</u>

(i) Expected waiting time of a customer.
(ii) Expected idle time of the service facility or the degree of utilization of the service facility.

Basic Structure of Queueing models:



Following assumptions are made:

1. Customers requiring service are generated over time by an **Input Source/Calling Source**.
2. As the customer arrives at the Queuing System/facility he joins the queue.
3. The server chooses a customer for service by some rule (**service discipline**).
4. The required service is performed for the customer by the **`service mechanism'.**
5. Upon the completion of a service, the customer leaves the system and the process of choosing a new (waiting) customer is repeated.

There are many alternative assumptions that can be made.

1. **Input source (Calling Population)**: finite or infinite.

   a) customers may arrive & get served individually or in groups – bulk queues.

2. **Queue size** - finite or infinite.

3. **Service discipline**: FCFS, LCFS or SIRO, priority queues.

4. **Service facility / mechanism**: Single server, or parallel servers or queues in series (tandem queues), network queues.

5. **Human behavior**: (i) Jockeying – jumping from one queue to another

   (ii) Balking – do not join the line anticipating long delay.

   (iii) Reneging – may walkout after being in the queue for a while because the wait has been too long.

A queuing system is completely specified by the following **6 main characteristics**:

1. Input or Arrival time distribution / **a**

2. Output or service time distribution / **b**

3. No. of service channels / Service facility /**c**

4. Service discipline /**d**

5. Capacity of Queuing System (Queue + in Service) /**e**

6. Calling source /**f**

a/**b**/**c** are due to $\longrightarrow$ Kendal (1953)

d/**e**/**f** are due to $\longrightarrow$ A M Lee (1966)

The standard notations used are:

M : Poisson (Markovian) arrival or departure distribution, equivalently ,

Inter arrival or service time follow exponential distribution.

D : Constant or deterministic inter arrival or service time.

$\boldsymbol{E_k}$: Erlangian or Gamma distribution of inter arrival or service time with parameter k

$$\left(T \sim f(t) = \frac{(k\mu)^k}{(k-1)!}\, e^{-k\mu t}\, t^{k-1} \,,\ t \geq 0 \,,\ k \ \&\ \mu \ \text{are the parameters}\right)$$

G : General distribution of departures (service time)

G I : General independent distribution of arrivals.

Ex : M | M | 1 : GD | N |∞

Poisson arrival | Departure | Servers: Service discipline | queue capacity | calling source

Thus we have

M | M | 1 : FIFO | ∞ | ∞ , M | M | S : FIFO | N | N , M | M | 2 : GD | ∞ | N ,

M | G | S : PRI | ∞ | ∞ , etc..

We use the following **terminologies and notations**:

State of the system = No. of customers in the Queueing system.

Queue length = No. of customers waiting for service.

= (State of the system) – (No. of customers being serviced)

N(t) = No. of customers in the system at time t.

$p_n(t)$ = P (there are `n' customers in the system at time t)

= P (N(t) =n) = Transient state probability ; assuming N(0) = 0

W(t) = Waiting time of a customer in the system at time t.

λ = Arrival rate (No. of customers arriving per unit time)

μ = Service rate (No. of customers served per unit time)

s = No. of servers.

$\rho = \frac{\lambda}{s\mu}$ = **Utilization factor** for the service facility or the **traffic intensity**.

= Expected fraction of time the individual server in busy.

In general, if the behavior of the system depends on time then such a system is said to be in a TRANSIENT STATE. This usually occurs at the early stage of operation of the system where its behavior will depend on the initial conditions. However, after sufficient time has elapsed, the behavior of the system becomes independent of time. The system is then said to be in a STEADY STATE OR EQUILIBRIUM STATE.

Due to the complexity involved in the analysis of transient state behavior and that our interest lies in obtaining an expression for the steady state probabilities, we consider only the steady state analysis. Under the steady state conditions, we use the following notations:

π_n : P (there are n customers in the system)

L_s : Expected No. of customers in the system

L_q : Expected queue length (excludes customers being served)

W_s : Expected waiting time of a customer in the system (includes service time)

W_q : Expected waiting time of a customer in the Queue. (excludes service time)

Relationship between L_s, L_q, W_s and W_q (Little's formula) :

For most of the Queuing systems the following relationships hold good

(due to John D.C. Little)

(i) $L_s = \lambda W_s$

(ii) $L_q = \lambda W_q$

(iii) $W_s = W_q + \frac{1}{\mu}$ (where $\frac{1}{\mu}$ = mean service time)

$$\lambda W_s = \lambda W_q + \frac{\lambda}{\mu}$$

$$L_s = L_q + \rho$$

Thus, if one of the four quantities in known, then the remaining can be found.

It may be easier to calculate L_s, where $L_s = \sum_{n=0}^{\infty} n\ \pi_n$

Or $L_q = \sum_{n=s}^{\infty} (n - s)\,\pi_n$

M/M/1: GD/∞/ ∞

This is the basic model widely used in queuing system. Here there is one server with no limit on the capacity on queuing system or calling source. Arrivals & departures occur according to a Poisson process with rates λ and μ respectively. Here $\lambda < \mu$ so that, there is no unending queue and all the customers are served. Service discipline being general discipline.(π_n is independent of service discipline). Thus, we use GD in the above notation. Our objective is to obtain an expression for π_n.

We first derive the differential- difference equations for $p_n(t)$, i.e. Probability of having `n' customers in the system during time 't'. Then under appropriate conditions, we take limits as t→∞ to obtain steady state probability π_n .

To derive the differential- difference equations, we make the following assumptions:

1) For an infinitesimal interval of length 'h'
P[one arrival in `h'] = λh + O(h).

P[no arrival in `h'] = 1- λh + O(h).

P[one departure in `h'] = μh + O(h).

P[no departure in `h'] = 1- μh + O(h).
Where $O(h) \to 0$ as $h \to 0$ and $\lim_{h\to 0} \frac{O(h)}{h} = 0$

2) Further, it is assumed that atmost only one event (either an arrival or a departure) can occur in h. Consider,
(i) When n>0 ; $p_n(t+h) = P[N(t\ +\ h) = n]$

The above probability is the sum of the following probabilities:

$$
\begin{aligned}
p_n(t+h) = {} & P[N(t) = n]\ P(\text{no arrival in h})\ P(\text{no departure in h}) \\
& + P[N(t) = n-1]\ P(\text{one arrival in h})\ P(\text{no departure in h }) \\
& + P[N(t) = n+1]\ P(\text{no arrival in h})\ P(\text{one departure in h}) \\
= {} & p_n(t)\ (1 - \lambda h + O(h))(1 - \mu h + O(h)) \\
& + p_{n-1}(t)\ (\lambda h + O(h))(1 - \mu h + O(h)) \\
& + p_{n+1}(t)\ (1 - \lambda h + O(h))(\mu h + O(h))
\end{aligned}
$$

$$p_n(t+h) = p_n(t)\ (1 - \lambda h - \mu h\ ) + p_{n-1}(t)\, \lambda h + p_{n+1}(t)\, \mu h + O(h)$$

$$p_n(t+h) - p_n(t) = \lambda h\ (p_{n-1}(t) - p_n(t)) + \mu h\ (p_{n+1}(t) - p_n(t)) + O(h)$$

$$\frac{p_n(t+h) - p_n(t)}{h} = \lambda\,(p_{n-1}(t) - p_n(t)) + \mu\,(p_{n+1}(t) - p_n(t)) + \frac{O(h)}{h}$$

As $h \to 0$: $p_n'(t) = \lambda(p_{n-1}(t) - p_n(t)) + \mu(p_{n+1}(t) - p_n(t))$

Or $p_n'(t) = \lambda\, p_{n-1}(t) + \mu\, p_{n+1}(t) + (\lambda + \mu)p_n(t)$; $n > 0$ --------- (1)

(ii) When n = 0 ; $p_0(t+h) = P[N(t+h) = 0]$

Noting that, for n=0, probability of occurrence of 0 departures during h is one,

we have,

$p_0(t+h) = p_0(t)\,(1 - \lambda h + O(h)) + p_1(t)\,(1 - \lambda h + O(h))(\mu h + O(h))$

$p_0(t+h) = p_0(t)\,(1 - \lambda h + O(h)) + p_1(t)\,(\mu h + O(h))$

$p_0(t+h) - p_0(t) = p_1(t)\mu h - p_0(t)\,\lambda h + O(h)$

$$\frac{p_0(t+h) - p_0(t)}{h} = \mu\, p_1(t) - \lambda\, p_0(t) + \frac{O(h)}{h}$$

As $h \to 0$: $p_0'(t) = \mu\, p_1(t) - \lambda\, p_0(t)$; n=0 ------------ (2)

Solving (1) and (2) by Laplace transformation technique a solution to $p_n(t)$, the transient state probabilities can be obtained. But the procedure is quite complex.
Due to the complexity involved in solving the above equations and that our interest lies in obtaining an expression for the steady state probabilities, we consider only the steady state analysis.

We obtain the steady state equations by noting that
as $t \to \infty$, $p_n'(t) \to 0$ and $p_n(t) \to \pi_n$ for all n.

Thus as $t \to \infty$, (1) and (2) reduce to

$\lambda\,(\pi_{n-1} - \pi_n) + \mu\,(\pi_{n+1} - \pi_n) = 0$; $n > 0$ ------------------- (3)

$\mu\pi_1 - \lambda\,\pi_0 = 0$; $n = 0$ --------------------------- (4)

Solution for π_n : from (4) we get, $\pi_1 = \frac{\lambda}{\mu}\pi_0 = \rho\pi_0$

In (3) put n=1 : $\lambda(\pi_0 - \pi_1) + \mu(\pi_2 - \pi_1) = 0$

$\mu\pi_2 = (\lambda + \mu)\pi_1 - \lambda\pi_0$

or $\pi_2 = \frac{\lambda^2}{\mu^2}\pi_0 = \rho^2\pi_0$

Similarly, put n=2 : $\lambda(\pi_1 - \pi_2) + \mu(\pi_3 - \pi_2) = 0$

$\mu\pi_3 = (\lambda + \mu)\pi_2 - \lambda\pi_1$

or $\pi_3 = \frac{\lambda^3}{\mu^3}\pi_0 = \rho^3\pi_0$

Thus we may arrive at: $\boldsymbol{\pi_n = \frac{\lambda^n}{\mu^n}\pi_0 = \rho^n\pi_0}$

To obtain π_0 :

Since π_n's are Probabilities, we have

$$\sum \pi_n = 1 \Rightarrow \sum_{n=0}^{\infty}\left(\frac{\lambda}{\mu}\right)^n \pi_0 = 1$$

$$\pi_0 = \frac{1}{\sum_{n=0}^{\infty}\left(\frac{\lambda}{\mu}\right)^n} = \frac{1}{\frac{1}{1-\frac{\lambda}{\mu}}} \quad ; \text{(since } \frac{\lambda}{\mu} < 1)$$

$$\therefore\ \pi_0 = \frac{\mu - \lambda}{\mu} = 1 - \frac{\lambda}{\mu} = 1 - \rho$$

$$\Rightarrow \boldsymbol{\pi_n = \rho^n(1 - \rho)}$$

Now $\pi_0 =$ Probability that server is idle or no queue = $\left(1 - \frac{\lambda}{\mu}\right) = 1 - \rho$

$\therefore$ Probability that server is busy = $1 - \pi_0 = \rho$

Consider,

$$L_s = \sum_{n=0}^{\infty} n\,\pi_n = \sum_{n=0}^{\infty} n\left(\left(\frac{\lambda}{\mu}\right)^n - \left(\frac{\lambda}{\mu}\right)^{n+1}\right)$$

$$= \left(1-\frac{\lambda}{\mu}\right)\sum_{n=0}^{\infty} n\left(\frac{\lambda}{\mu}\right)^n = \left(1-\frac{\lambda}{\mu}\right)\left(0+1.\left(\frac{\lambda}{\mu}\right)+2.\left(\frac{\lambda}{\mu}\right)^2+\cdots\right)$$

$$L_s = \left(1-\frac{\lambda}{\mu}\right)\frac{\frac{\lambda}{\mu}}{\left(1-\frac{\lambda}{\mu}\right)^2} \quad ; \text{ Since } \frac{\lambda}{\mu} < 1 \quad \left[\, s = \frac{a}{(1-a)^2}\right]$$

$$L_s = \frac{\rho}{1-\rho} = \frac{\lambda}{\mu-\lambda}$$

$$L_q = L_s - \rho = \frac{\rho}{1-\rho} - \rho = \frac{\rho^2}{1-\rho} = \frac{\lambda^2}{\mu(\mu-\lambda)}$$

Or $L_q = \sum_{n=1}^{\infty} (n-1)\,\pi_n$

$$W_s = \frac{L_s}{\lambda} = \frac{\lambda}{\lambda(\mu-\lambda)} = \frac{1}{(\mu-\lambda)} = \frac{1}{\mu(1-\rho)} = \frac{\rho}{\lambda(1-\rho)}$$

$$W_q = \frac{L_q}{\lambda} = \frac{\lambda}{\mu(\mu-\lambda)} = \frac{\rho}{\mu(1-\rho)} = \frac{\lambda}{\mu^2(1-\rho)} = \frac{\rho^2}{\lambda(1-\rho)}$$

Waiting time distribution of a customer in the system:

Let $W(t)$ denote the waiting time of a customer in the queuing system for the customer joining the system at time t. In steady state, $\lim_{t\to\infty} W(t) = W$.

W is the waiting time for the customer joining the system at any time.

Note: Although derivation of π_n is completely independent of service discipline and so also the expected waiting times. The probability distribution of waiting times depends on service discipline adopted. This means that while the expected value of waiting times remain the same their variance differ.

Let the service discipline be FCFS $(M/M/1 \text{ Model})$.

Let there be n customers in the system of which one be undergoing service. Let 'W' be the waiting of (n+1)th customer joining the system. Let $t_2, t_3,, t_4 \dots . t_n$ be the service times of $(n-1)$ customers in the queue and let t_1' be the remaining service time of the customer being served and t_{n+1} be the service time of arriving customer.

Clearly, W = $t_1' + t_2 + t_3 + \dots + t_n + t_{n+1}$

Let $f(w/n)$ be the conditional pdf of W given that there are n customers in the system ahead of the arriving customer. Since the variables t_i; i =2,3,4….,n+1 each has exponential distribution with parameter μ and further due to forgetfulness property of exponential distribution t_1' also has exponential distribution with parameter μ . Thus, $f(w/n)$ is a gamma distribution with parameters μ & n+1, with pdf

$$f(w/n) = \frac{\mu(\mu w)^n}{n!} e^{-\mu w} \ ; \ w \geq 0$$

Hence the density function of W is,

$$f(w) = \sum_{n=0}^{\infty} f(w/n)\, \pi_n = \sum_{n=0}^{\infty} \frac{\mu(\mu w)^n}{n!} e^{-\mu w} \rho^n (1-\rho)$$

$$= \mu(1-\rho)\, e^{-\mu w} \sum_{n=0}^{\infty} \frac{(\mu w \rho)^n}{n!}$$

$$= \mu(1-\rho)\, e^{-\mu w}\, e^{\mu w \rho}$$

$$f(w) = \mu(1-\rho) e^{-\mu(1-\rho)w} \ ; \quad w \geq 0 \ , \ \rho < 1$$

Which is an exponential distribution with parameter $\mu(1-\rho)$.

∴ Expected waiting time of the customer in the system: $W_s = \frac{1}{\mu(1-\rho)}$

Similarly, we can obtain the **Waiting time distribution of an arriving customer in the queue.** (i. e. before he receives service)

Let $W^*(t)$ denote the waiting time of a customer in the queue for the customer joining the system at time t. In steady state, $\lim_{t\to\infty} W^*(t) = W^*$.

W^* is the waiting time for the customer joining the Queue at any time.

The Probability distribution of W^* has 2 components:

1) Customer starts receiving the service immediately upon his arrival, if there is no customer in the system.

   Thus $P(W^* = 0) = \pi_0 = 1-\rho = 1-\frac{\lambda}{\mu}$; $w^* = 0$

2) If there are n customers in the system $(n \geq 1)$, when a new customer arrives, following similar line of arguments, the waiting time before the service starts is the sum of n independently exponentially distributed r.vs.

   Clearly, $W^* = t_1' + t_2 + t_3 + \dots + t_n$

   Thus, $f(w^*/n)$ is a gamma distribution with parameters μ & n, with pdf

$$f(w^*/n) = \frac{\mu(\mu w^*)^{n-1}}{(n-1)!} e^{-\mu w^*} \ ; \ w^* > 0$$

Hence the density function of W^* is,

$$f(w^*) = \sum_{n=1}^{\infty} f(w^*/n)\,\pi_n = \sum_{n=1}^{\infty} \frac{\mu(\mu w^*)^{n-1}}{(n-1)!} e^{-\mu w^*}\,\pi_n \quad ; \ w^* > 0$$

$$= \sum_{n=1}^{\infty} \frac{\mu(\mu w^*)^{n-1}}{(n-1)!} e^{-\mu w^*} \rho^n (1-\rho)$$

$$= \mu\rho(1-\rho) e^{-\mu w^*} \sum_{n=1}^{\infty} \frac{(\mu w^* \rho)^{n-1}}{(n-1)!}$$

$$= \mu\rho(1-\rho) e^{-\mu w^*} e^{\mu\rho w^*}$$

$$f(w^*) = \mu\rho(1-\rho)\ e^{-\mu(1-\rho)w^*}$$

$$\therefore \quad f(w^*) = \begin{cases} \mu\rho(1-\rho)\, e^{-\mu(1-\rho)w^*} & ; \ w^* > 0 \\ (1-\rho) & ; \ w^* = 0 \end{cases}$$

By elementary integration methods, we obtain,

Expected waiting time of a customer in the queue: $W_q = \frac{\lambda}{\mu(\mu-\lambda)} = \frac{\rho}{\mu(1-\rho)}$

Problems

1) For a M/M/1 queuing system with an arrival rate 5/hr and mean service time 10 minutes. Find
   (i). Probability that customer directly gets the service.
   (ii). Probability that a customer has to wait.
   (iii). Expected No. of customers in the system, L_s.
   (iv). Expected No. of customers in the queue, L_q.
   (v). Expected waiting time of the customer in the system, W_s
   (vi). Expected waiting time of the customer in the queue, W_q
   (vii). Probability that 5 or more customers in the system at any time.

Solution:

$\lambda = 5$ arrivals per hour, $\mu = 1$ departure per 10 minutes = 6 departures per hour

$\rho = \frac{\lambda}{\mu} = \frac{5}{6} < 1$

$\therefore \pi_n = \rho^n (1 - \rho)$

(i). $\pi_0 = 1 - \rho = \frac{1}{6}$

(ii). $\rho = \frac{5}{6}$

(iii). $L_s = \frac{\rho}{1-\rho} = 5$

(iv). $L_q = \frac{\lambda^2}{\mu(\mu - \lambda)} = \frac{25}{6}$

(v). $W_s = \frac{L_s}{\lambda} = 1$ hr

(vi). $W_q = \frac{L_q}{\lambda} = \frac{5}{6}$ hr $=$ 50 minutes

(vii).
$$P(N \geq 5) = 1 - P(N \leq 4) = 1 - (\pi_0 + \pi_1 + \pi_2 + \pi_3 + \pi_4)$$
$$= 1 - \pi_0\left(1 + \left(\frac{\lambda}{\mu}\right) + \left(\frac{\lambda}{\mu}\right)^2 + \left(\frac{\lambda}{\mu}\right)^3 + \left(\frac{\lambda}{\mu}\right)^4\right)$$
$$= 1 - \frac{1}{6}\left(1 + \frac{5}{6} + \frac{25}{36} + \frac{125}{216} + \frac{625}{1296}\right)$$
$$= \left(\frac{5}{6}\right)^5$$

2) Consider a M/M/1 situation where the mean arrival rate (λ) is 1 customer every 4 minutes and the mean service time $\left(\frac{1}{\mu}\right)$ is 2.5 minutes. Calculate L_q, L_s, W_s & W_q.

Solution:

$\lambda = \frac{1}{4}$ = 0.25 arrivals per minute = 15 arrivals per hour,

$\mu = \frac{1}{2.5} =$ 0.4 departures per minute = 24 departures per hour

$\rho = \frac{\lambda}{\mu} = \frac{15}{24} < 1$

$\therefore \ \pi_n = \rho^n(1-\rho)$

(i). $L_s = \frac{\rho}{1-\rho} = \frac{5}{3} \simeq 2$ customers

(ii). $L_q = \frac{\rho^2}{1-\rho} = \frac{25}{24} \simeq 1$ customer

(iii). $W_s = \frac{L_s}{\lambda} = 6.67$ minutes

(iv). $W_q = \frac{L_q}{\lambda} = 4.16$ minutes

3) Arrival at a telephone booth is considered to be Poisson with an average time of 10 minutes between 2 arrivals. The length of a phone call is assumed to be distributed exponentially with mean 3 minutes.

(i). What is the Probability that person will have to wait?

(ii). The telephone dept. will install a 2nd booth when convinced that an arrival would expect to wait at least 3 minutes for the phone. By how much must the flow of arrivals be increased in order to justify a second booth?

(iii). Find the average No. of customers in the system?

(iv). Estimate the fraction of the day that the phone will be in use.

(v). What is the probability that it will take a person more than 10 minutes to wait for the phone and complete his call?

Solution:

$\lambda = 1$ arrival per 10 minutes = 6 arrivals per hour,

$\mu = \frac{1}{3}$ departures per minute = 20 departures per hour

$\rho = \frac{\lambda}{\mu} = \frac{6}{20} < 1$

$\therefore \pi_n = \rho^n(1 - \rho)$

(i). $P(\text{an arrival has to wait}) = 1 - \pi_0 = \rho = \frac{6}{20}$

(ii). The expected waiting time in the queue for an arrival before he gets the service is

$W_q = \frac{\lambda}{\mu(\mu-\lambda)}$.

Here $\mu = 0.33$. To find the new value of λ, say $\lambda^{|}$, for which $W_q = 3$.

$3 = \frac{\lambda^{|}}{0.33(0.33-\lambda^{|})} \Rightarrow \lambda^{|} = 0.16$ arrivals per minute $\simeq$ 10 arrivals per hour

So we must increase the flow of arrivals from 6 per hr to 10 per hr to justify the installation of a 2nd booth

(iii). $L_s = \frac{\rho}{1-\rho} = \frac{0.3}{0.7} = \frac{3}{7}$

(iv). $\rho = 0.3$

(v). $P(W \geq 10) = \int_{10}^{\infty} f(w)\, dw = \int_{10}^{\infty} \mu(1-\rho)\, e^{-\mu(1-\rho)\, w}\, dw = 0.1$

4) Customers arrive at a ticket window according to a Poisson process with rate 30 per hour. A single person is appointed to serve the customers. The service time is exponential with mean 90 seconds. Find the average waiting time of the customer in the system and in the queue.

Solution:

$\lambda = 30$ arrivals per hour

$\mu = \frac{1}{90}$ departures per second = 40 departures per hour

$\rho = \frac{\lambda}{\mu} = \frac{3}{4} < 1$

$\therefore \pi_n = \rho^n(1-\rho)$

$W_s = \frac{\rho}{\lambda(1-\rho)} = 60 \text{ min} = 360 \text{ seconds}$

$W_q = \frac{\rho^2}{\lambda(1-\rho)} = 270 \text{ seconds}$

5) Arrival at a counter in a bank occur is in accordance with a Poisson process at an Average rate of 8 per hour .The duration of service of a customer has an exponential distribution with a mean of 6minutes. Find the probability that an arriving customer
   (i). has to wait on arrival.
   (ii). finds 4 customers in the queue .
   (iii). has to spend less than 15 minutes in the bank?
   (iv). Also estimate fraction of the total time the counter busy.

Solution:

$\lambda = 8$ arrivals per hour
$\mu = \frac{1}{6}$ departures per minute = 10 departures per hour
$\rho = \frac{\lambda}{\mu} = \frac{4}{5} < 1$
$\therefore \pi_n = \rho^n(1-\rho)$
(i). P(arriving customer has to wait) = $1 - \pi_0 = \rho$ = 0.8
(ii). P(arriving customer finds 4 customers in the queue)
$= P(N=4) = \pi_4 = \rho^4\,\pi_0 = (0.8)^4(1-0.8) = 0.8192$
(iii). $P(W < 15) = 1 - P(W \geq 15) = 1 - \int_{15}^{\infty} (\mu-\lambda).\ e^{-(\mu-\lambda)\,w}\, dw = 0.3935$
(iv). $\rho = 0.8$

6) A TV repairman finds that the time spent on the TVs has an exponential distribution with mean 20mins. If he repairs the sets in the order in which they arrive and if the arrival of the sets is approximately Poisson with an average rate of 10 per 8 hour a day. What is the repairman's expected idle time each day? How many jobs are ahead on an average of the set just brought in?

Solve it!

7) In a post office there is only one window and a stationary employee performs all the services required. The window remains open continuously from 7AM to 1PM. It has been observed that average number of clients arriving is 54 and average service time is 5 minutes per person. Assuming Poisson arrival and Poisson departure determine L_S, L_q, W_S, W_q.

Solve it!

8) Customers arrive at random at a checkout facility at an average rate of 12/hr. The service time has an exponential distribution with parameter μ. If the queing time of atleast 90% of the customers should be less than 4 minutes, show that μ must exceed μ_0, where μ_0 satisfies $\mu_0 e^{4\mu_0} = 2e^{4/5}$.

Solve it!

9) In a railway marshalling yard goods train arrive at a rate 30 trains per day. Assuming that inter-arrival time is exponentially distributed and the service time distribution is also exponential with an average of 36 minutes. Calculate

   (a). Mean queue size.

   (b). Probability that queue size exceeds 10.

If the input of trains exceeds an average of 33 per day, what is the change in (a) and (b) above.

Solve it!

10) At what average rate must a clerk at a supermarket work in order to ensure a probability of 0.9 that customers will not wait longer than 12 minutes. It is assumed that there is one counter at which customers arrive at average rate of 15/hour. Assume that the length of service by the clerk has an exponential distribution.

Solve it!

M | M | s: FIFO | ∞ | ∞ (Multichannel queueing system)

In this model it is assumed that the arrivals and departures occur according to Poisson distribution with rates λ and μ respectively. Equivalently the Inter arrival & service times are exponentially distributed with parameters λ and μ respectively. Further it is assumed that there are `s' parallel servers in the system working independently of each other. Also it is assumed that all the servers offer service at the same rate.

Thus, if the no. of customers in the system be n, then

for $n \geq s$: the combined service rate is $s\mu$

for $n < s$: the combined service rate is $n\mu$ - since no more than n servers will be busy.

To derive the differential-difference equations we use the following assumptions:

1) For an infinitesimal interval of length 'h'

$$P[\text{ one arrival in h }] = \lambda h + O(h)$$

$$P[\text{ no arrival in h }] = 1 - \lambda h + O(h)$$

$$P[\text{ one departure in h }] = n\mu h + O(h) \quad 0 < n < s$$

$$= s\mu h + O(h) \quad n \geq s$$

$$P[\text{ no departure in h }] = 1 - n\mu h + O(h) \quad 0 < n < s$$

$$= 1 - s\mu h + O(h) \quad n \geq s$$

Where $O(h) \to 0$ as $h \to 0$ and $\lim_{h\to 0} \frac{O(h)}{h} = 0$

2) Further, it is assumed that atmost only one event (either an arrival or a departure) can occur in h.

Following the similar line of arguments as in **M|M|1** system we have,

$p_0(t+h) = p_0(t)\left(1-\lambda h+O(h)\right).1 + p_1(t)(1-\lambda h+O(h))(\mu h + O(h)) + O(h)$; n=0

$p_n(t+h) = p_n(t)\left((1-(\lambda+n\mu)h + O(h)\right) + p_{n-1}(t)\left(\lambda h+O(h)\right)$

$+p_{n+1}(t)((n+1)\mu h + O(h)) + O(h)$; n< s [i.e. n=1,2,3,…..(s-1)]

$p_n(t+h) = p_n(t)\left((1-(\lambda+s\mu)h + O(h)\right) + p_{n-1}(t)\left(\lambda h+O(h)\right)$

$+p_{n+1}(t)((s\mu h + O(h)) + O(h)$; $n \geq s$ [i.e. n = s,s+1,s+2…..]

Proceeding further as earlier, we obtain the following differential-difference equations:

$p_0'(t) = -\lambda p_0(t) + \mu p_1(t)$ for n=0 ---------------(1)

$p_n'(t) = -(\lambda+n\mu)p_n(t) + \lambda p_{n-1}(t) + (n+1)\mu p_{n+1}(t)$ for n<s -------------(2)

$p_n'(t) = -(\lambda+s\mu)p_n(t) + \lambda p_{n-1}(t) + s\mu p_{n+1}(t)$ for $n \geq s$ ----------------(3)

Solving (1), (2) and (3) by Laplace transformation technique a solution to $p_n(t)$, the transient state probabilities can be obtained. But the procedure is quite complex.

Due to the complexity involved in solving the above equations and that our interest lies in obtaining an expression for the steady state probabilities, we consider only the steady state analysis.

we obtain the steady state equations by noting that
as $t\to\infty$, $p_n'(t) \to 0$ and $p_n(t) \to \pi_n$ for all n.

Thus we get

$-\lambda\pi_0 + \mu\pi_1 = 0$ for n=0 ---------------- (4)

$-(\lambda+n\mu)\pi_n + \lambda\pi_{n-1} + (n+1)\mu\pi_{n+1} = 0$ for n< s ------------- (5)

$-(\lambda+s\mu)\pi_n + \lambda\pi_{n-1} + s\mu\pi_{n+1} = 0$ for $n \geq s$ --------------- (6)

A solution π_n is obtained by substituting n=1,2.....(s-1) and calling $\rho = \frac{\lambda}{s\mu}$
(the utilization factor for the service facility or the expected fraction of time the individual server is busy)

To obtain the solution π_n :

Here $\pi_0 = \pi_0$ (initially)

$$\pi_1 = \frac{\lambda}{\mu}\,\pi_0$$

$$\pi_2 = \frac{\lambda}{2\mu}\,\pi_1 = \frac{1}{2!}\left(\frac{\lambda}{\mu}\right)^2 \pi_0$$

$$\pi_3 = \frac{\lambda}{3\mu}\,\pi_2 = \frac{1}{3!}\left(\frac{\lambda}{\mu}\right)^3 \pi_0$$

$$\vdots$$

$$\pi_n = \frac{\lambda}{n\mu}\,\pi_{n-1} = \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n \pi_0 \quad \text{for} \quad n < s$$

i.e. $\pi_n = \frac{1}{n!}(s\rho)^n\,\pi_0 \quad \text{for} \quad n < s$

Similarly

$$\pi_s = \frac{\lambda}{s\mu}\,\pi_{s-1} = \frac{1}{s!}\left(\frac{\lambda}{\mu}\right)^s \pi_0$$

$$\pi_{s+1} = \frac{\lambda}{s\mu}\,\pi_s = \frac{1}{s\,s!}\left(\frac{\lambda}{\mu}\right)^{s+1} \pi_0$$

$$\pi_{s+2} = \frac{\lambda}{s\mu}\,\pi_{s+1} = \frac{1}{s^2\,s!}\left(\frac{\lambda}{\mu}\right)^{s+2} \pi_0$$

$$\vdots$$

$$\pi_n = \pi_{s+n-s} = \frac{1}{s^{n-s}\,s!}\left(\frac{\lambda}{\mu}\right)^n \pi_0 \quad \text{for} \quad n \geq s$$

$$\pi_n = \frac{1}{s!}\,s^s\,\rho^n\,\pi_0 \quad \text{for} \quad n \geq s$$

$$\text{Thus } \pi_n = \begin{cases} \frac{1}{n!}(s\rho)^n\,\pi_0\ ; & n < s \\ \frac{1}{s!}\,s^s\,\rho^n\,\pi_0\ ; & n \geq s \end{cases}$$

where π_0 = Prob [having no customers in a multi-channel system]

To obtain π_0 :

Since $\pi_n's$ are Probabilities we have $\sum_{n=0}^{\infty} \pi_n = 1$

i.e. $\pi_0 + \sum_{n=1}^{s-1} \pi_n + \sum_{n=s}^{\infty} \pi_n = 1$

$\Rightarrow \pi_0 = \frac{1}{1 + \sum_{n=1}^{s-1} \frac{(s\rho)^n}{n!} + \sum_{n=s}^{\infty} \frac{s^s \rho^n}{s!}}$

$\Rightarrow \boldsymbol{\pi_0} = \frac{\mathbf{1}}{\sum_{n=0}^{s-1} \frac{(s\rho)^n}{n!} + \frac{s^s \rho^s}{s!} \frac{s\mu}{s\mu - \lambda}}$; which holds only if $\rho < 1$, i.e. $s\mu > \lambda$.

(In all other cases, i.e. $\rho \geq 1 \Rightarrow s\mu \leq \lambda$, the waiting time increases indefinitely)

Now we obtain L_s, L_q, W_s and W_q (as earlier using the Little's formula)

i.e. $L_s = \sum_{n=0}^{\infty} n\pi_n$ & $L_q = \sum_{n=s}^{\infty} (n - s)\pi_n$

$L_s = \sum_{n=0}^{\infty} n\pi_n$, it can be shown that $L_s = \frac{\lambda\mu\left(\frac{\lambda}{\mu}\right)^s}{(s-1)!\,(s\mu-\lambda)^2} \pi_0 + \frac{\lambda}{\mu}$

$L_q = \frac{\lambda\mu\left(\frac{\lambda}{\mu}\right)^s}{(s-1)!\,(s\mu-\lambda)^2} \pi_0$ OR $L_q = \frac{\rho(s\rho)^s}{s!\,(1-\rho)^2} \pi_0$

$W_s = \frac{\mu\left(\frac{\lambda}{\mu}\right)^s}{(s-1)!\,(s\mu-\lambda)^2} \pi_0 + \frac{1}{\mu} = W_q + \frac{1}{\mu}$

& $W_q = \frac{\mu\left(\frac{\lambda}{\mu}\right)^s}{(s-1)!\,(s\mu-\lambda)^2} \pi_0 \Rightarrow W_q = \frac{1}{\lambda} L_q$

P(that an arrival has to wait) = $P(N \geq s) = \sum_{n=s}^{\infty} \pi_n = \frac{\mu\left(\frac{\lambda}{\mu}\right)^s}{(s-1)!\,(s\mu-\lambda)} \pi_0$

& P(that an arrival enters service without wait) = $1 - P(N \geq s)$

Average No. of idle servers = s – Average No. of customers served

Utilization factor $\rho = \frac{\lambda}{s\mu}$

Note: For s = 1, it can be shown that

$\pi_0 = 1 - \rho = 1 - \frac{\lambda}{\mu}$; $\pi_n = \rho^n (1 - \rho) = \left(\frac{\lambda}{\mu}\right)^n \left(1 - \frac{\lambda}{\mu}\right)$, n=0,1,2........

$\therefore L_s = \frac{\lambda}{(\mu - \lambda)}$, $L_q = \frac{\lambda^2}{\mu\,(\mu - \lambda)}$, $W_s = \frac{1}{(\mu - \lambda)}$ and $W_q = \frac{\lambda}{\mu\,(\mu - \lambda)}$

which is nothing but the **M|M|1** system.

Problems

1) Four counters are being run on frontier of a country to check the passports & necessary papers of the tourists. The tourists choose a counter at random, if the arrivals at the frontier is Poisson at the rate λ and the service time is exponential with parameter $\frac{\lambda}{2}$, what is the steady state average queue at each counter?

Solution:

Here we have to find L_q .

we have $L_q = \frac{\lambda\mu\left(\frac{\lambda}{\mu}\right)^s}{(s-1)!\,(s\mu-\lambda)^2}\ \pi_0$

Here $\lambda = \lambda$, $\mu = \frac{\lambda}{2}$, s = 4 $\Rightarrow$ $\rho = \frac{\lambda}{s\mu} = \frac{1}{2}$ $\Rightarrow$ $s\rho = 2$

$$\pi_0 = \left(\sum_{n=0}^{s-1} \frac{(s\rho)^n}{n!} + \frac{(s\rho)^s}{s!}\frac{s\mu}{s\mu-\lambda}\right)^{-1} = \frac{3}{23} = 0.130$$

$$L_q = \frac{\left(s\,\frac{\lambda}{s\mu}\right)^s\left(\frac{\lambda}{s\mu}\right)}{s!\left(1-\frac{\lambda}{s\mu}\right)^2}\pi_0 = \frac{(\rho)(s\rho)^s}{s!(1-\rho)^2}\ \pi_0 = \frac{4}{23} = 0.174$$

2) A telephone company is planning to install telephone booths in a new airport. It has established policy that a person should not have to wait for more than 10 percent of the times he tries to use a phone. The demand for use is estimated to be Poisson with an average of 30 per hour. The average phone call has an exponential distribution with a mean time of 5 minutes. How many phone booths should be installed?

Solution:

Here λ = 30 per hour, $\mu = \frac{60}{5}$ = 12 per hour ($\because \frac{1}{\mu} = 5$ minutes)

Now $\rho = \frac{\lambda}{s\mu}$

If $s = 1 \Rightarrow \rho = \frac{\lambda}{\mu} = \frac{30}{12} = 2.5 \nless 1$

If $s = 2 \Rightarrow \rho = \frac{5}{4} \nless 1$ $\therefore$ queue explodes

If $s = 3 \Rightarrow \rho = \frac{5}{6} < 1$ $\therefore$ minimum value of s is 3

Let us assume that the company decides to install two telephone booths,

i.e. s= 2, then $s\mu = 24$, i.e. $s\mu < \lambda$, the arrival rate for s = 2. Thus the company must have atleast three telephones to meet the demand for service.

Note: Our problem is to decide on s keeping in mind the company planning that probability that a customer will have to wait is < 10% = 0.10.

Now consider s = 5, then

$$\pi_0 = \left(\sum_{n=0}^{5-1} \frac{(s\rho)^n}{n!} + \frac{(s\rho)^s}{s!}\frac{s\mu}{s\mu-\lambda}\right)^{-1} = 0.080$$

Now Probability that a customer on arrival has to wait

$$P(N \geq s) = \frac{\mu\left(\frac{\lambda}{\mu}\right)^s}{(s-1)!\,(s\mu-\lambda)}\pi_0 \quad \text{i.e. } P(N \geq 5) = \frac{12.(2.5)^5}{4!\,(5\times12-30)}(0.080) = 0.13$$

Thus we have for the given problem,

Probability that a customer has to wait = 0.13 > 0.10

Hence installing 5 booths will not meet the company's policy.

Now, if s = 6, then we have $\pi_0 = 0.0816$ and $P(N \geq 6) = 0.047 < 0.10$
Thus an installation of Six phones would meet the company's policy.

3) Given an average arrival rate of 20 per hour, is it better for a customer to get service at a single channel with mean service rate of 22 customers and at one of two channels in parallel with mean service rate of 11 customers for each of the two channels? Assume that both queues are of (M | M | s) type.

Solution:

For a single channel λ = 20 arrival per hour & μ = 22 customers

Here we compare W_q for s = 1 & s = 2

$\pi_0 = 1 - \rho = 1 - \frac{\lambda}{\mu} = 0.09$

$L_s = \frac{\lambda}{(\mu - \lambda)} = 10$ and $W_q = \frac{\lambda}{\mu\,(\mu - \lambda)} = 0.45$

Now when there are 2 parallel channels, we have λ = 20, μ = 11 & s = 2

Here $\pi_0 = \left(\sum_{n=0}^{2-1} \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n + \frac{1}{2!}\left(\frac{\lambda}{\mu}\right)^2 \frac{2\mu}{2\mu-\lambda}\right)^{-1} = 0.16$

and $L_s = \frac{\lambda\mu\left(\frac{\lambda}{\mu}\right)^s}{(s-1)!\,(s\mu-\lambda)^2}\,\pi_0 + \frac{\lambda}{\mu} = 30.9$

Similarly, $L_q = L_s - s\,\rho = 290.09$

$W_q = \frac{\mu\left(\frac{\lambda}{\mu}\right)^s}{(s-1)!\,(s\mu-\lambda)^2}\,\pi_0 = \frac{1}{\lambda}\,L_q = 1.45$

On comparison we see that it is better for a customer to get service at single channel, since in the case of single channel he has to wait for 0.45 hours whereas in the two channels he will have to wait for 1.45 hours.

4) Patients arrive at an OPD of a hospital in accordance with a Poisson process at mean rate of 12/hour and the distribution of time for examination by an attending physician is exponential with a mean of 10 minutes.
   (a). What is minimum number of physicians to be posted for ensuring steady state distribution?
   (b). For this number find $W_{q.}$
   (c). Find Ls.
   (d). How many physicians will remain idle?

Solve it!

5) A supermarket has 2 girls ringing up sales at the counters. If the service time for each customer is exponential with mean 4 minutes and if people arrive in a Poisson fashion at the counter at the rate of 10/hour.

   (a). What is the probability of having to wait for service?

   (b). What is the expected percentage of idle time for each girl?

   Solve it!

6) A bank has 2 tellers working on savings account, the first teller handles withdrawals only while the second teller handles deposits only. It has been found that service time distributions for both deposits and withdrawals are exponential with mean service times 3 minutes per customers. Depositors are found to arrive in Poisson fashion throughout the day at a rate 16/hour. Withdrawals also arrive in Poisson fashion with rate 14/hour. What would be effect on average waiting time for depositors and withdrawers if each teller would handle both withdrawals and deposits? What would be the effect if this could be accomplished by increasing the service time to 3.5 minutes?

   Solve it!

7) Ships arrive at a port at the rate of 1 in every 4 hours with an exponential distribution of inter-arrival times. The time a ship occupies a berth for unloading has an exponential distribution with mean of 10 hours. If the average delay of ships waiting for a berth is to be kept below 14 hours, how many berths should be provided at the port?

   Solve it!

8) A library wants to improve its service facilities in terms of the waiting time of its borrowers. The library has 2 counters at present and borrowers arrive at a rate of 1 every 6 minutes and the service time is 10 minutes per borrower. The library has relaxed its membership rules and a substantial increase in the number of borrowers is expected. Find the number of additional counters to be provided if the arrival rate is expected to be twice the present value and the average waiting time of borrower must be limited to half the present value.

   Solve it!

Non – Poisson Queues

M | G | 1: FIFO | ∞ | ∞ Model

This queueing model has single server and Poisson input process with mean arrival rate λ. However, there is a general distribution for the service time whose mean and variance are $1/\mu$ and σ^2 respectively.

For this system the steady state condition is $\rho = \frac{\lambda}{\mu} < 1$.

The characteristics of this model are given below:

(i) $\pi_0 = 1 - \rho$

(ii) $L_q = \frac{\lambda^2\sigma^2+\rho^2}{2(1-\rho)}$

(iii) $L_s = L_q + \rho$

(iv) $W_q = \frac{L_q}{\lambda}$

(v) $W_s = W_q + \frac{1}{\mu}$

Reliability Theory

Reliability is an old concept and a new discipline. For ages things and people have been called reliable if they had lived upto certain expectations and unreliable otherwise. A reliable person would never (or hardly ever) fail to deliver what he had promised.

The types of expectation to judge reliability have all been related to the performance of some function. The reliability of a device has been considered high if it had repeatedly performed its function with success and low if it had tended to fail in repeated trails. Reliability theory deals with the general methods and procedures to be followed during the process of planning, preparation, acceptance, etc. of manufactured items so as to ensure maximum effectiveness in their usage.

It develops general methods of evaluation of the quality of a system from the known qualities of components used in the system. The theory introduces quantitative indices to the quality of the product.

The history reliability engineering goes back to World War II. Of late, the subject has assumed greater importance with recent advances in the field of electronics, nuclear engineering, computers, aeronautic engineering, etc.

There is a common misunderstanding that quality and reliability are not different, in fact quality and reliability do not imply the same thing. Quality control no doubt contributes significantly to the improvement of the reliability of a product. Quality control is a management function it aims at preventing the manufacture of defectives by exercising control over various factors affecting the manufacturing process. The traditional concept of quality is not time dependent. A product is accepted if it meets certain specifications otherwise it is rejected.

While reliability is usually concerned with failures in time domain.

Definition:

The reliability of a system is the probability that it performs its intended function adequately for a specified time interval under the given operating conditions. Reliability of a system is also defined as the probability of its failure free operation during the period it is intended to be in use.

The failure of a system can be visualized as an event associated with the deviation in the operating characteristics of the system from its permissible limits.

In the above definition, we note four important factors or elements which contribute to reliability of the system.

1. Reliability of a device is expressed in terms of probability.
2. Device is required to perform adequately.
3. Duration of its performance is to be specified.
4. Operating or environmental conditions are to be prescribed.

The basic quantity of interest is the time to failure of the system. It is the time that has elapsed from the start of the operation of the system till its failure for the first time. This failure time is a random variable and is governed by a probability distribution.

Let $T(0 \leq T < \infty)$ be the time to failure of a system which started operating at some time origin. Let f(t) be the pdf of T i.e.

$f(t)dt = P\{t \leq T \leq t + dt\}$

= Probability that failure occurs in the time interval of length dt.

Let F(t) be the cdf of T

$F(t) = P(T \leq t) = \text{Probability that failure takes place at time } \leq t$

i.e. F(t) denotes the probability that the system which started at T=0 will fail before t.

Let R(t)=P{System operates without failure till time t}

R(t) = P{T>t}

= 1-F(t)

= $1-\int_0^t f(x)dx$

R(t) = $\int_t^\infty f(x)dx$

From the properties of the pdf we have, R(0)=1 and $R(\infty) = 0$

Also, we have, R(t)=1-F(t) from which we get f(t) = $\frac{-d}{dt}[R(t)]$

Failure rate:

This is also called hazard rate, instantaneous failure rate or age specific failure rate. The failure rate h(t) may be defined in terms of reliability R(t) and time to failure f(t) as follows:

Let h(t)dt be the conditional probability of time to failure of a system in the interval [t, t+dt] given that the system has not failed at time T=t.

$h(t)dt = P\{(t \leq T \leq t + dt)/(T > t)\}$

$= \frac{P\{(t \leq T \leq t+dt) \cap (T>t)\}}{P(T>t)}$

$= \frac{P(t \leq T \leq t+dt)}{R(t)}$

$= \frac{f(t)dt}{R(t)}$

$\therefore$ h(t) = $\frac{f(t)}{R(t)}$

This quantity h(t), the failure rate is also referred to as the hazard function or instantaneous hazard rate.

One can establish following relationship between h(t), f(t) and R(t)

$$h(t) = \frac{f(t)}{R(t)} = \frac{f(t)}{1-F(t)}$$

$$= \frac{d[-\log(1-F(t))]}{dt}$$

$$\frac{d[\log(1-F(t)]}{dt} = -h(t)$$

$$\log(1 - F(t)) = \int_0^t -h(x)dx$$

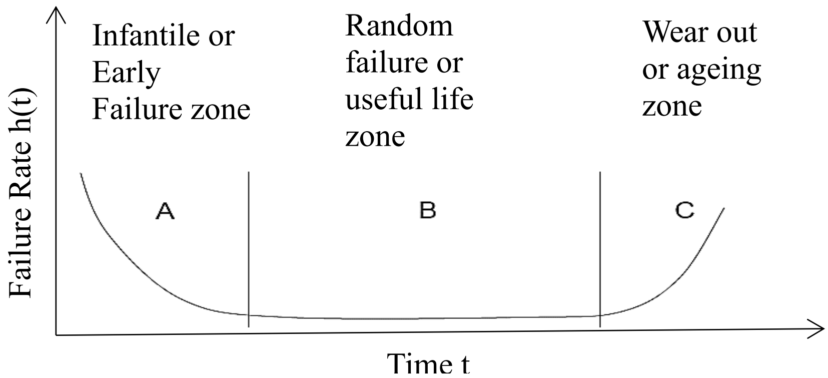
$$1 - F(t) = e^{-\int_0^t h(x)dx}$$

$$R(t) = e^{-\int_0^t h(x)dx}$$

$$f(t) = h(t)e^{-\int_0^t h(x)dx}$$

Note: When $h(t) = \lambda$, a constant, then $f(t) = \lambda e^{-\int_0^t \lambda\, dx} = \lambda e^{-\lambda t}$ i.e. a constant failure rate leads to an exponential distribution.

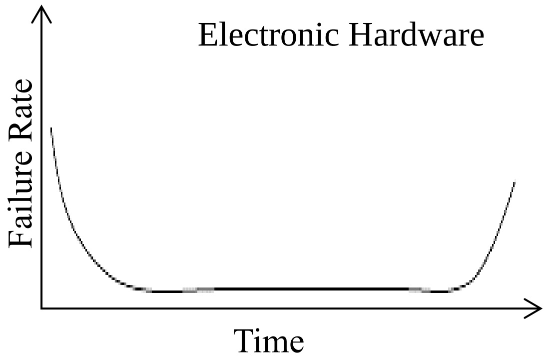
Failure Pattern:

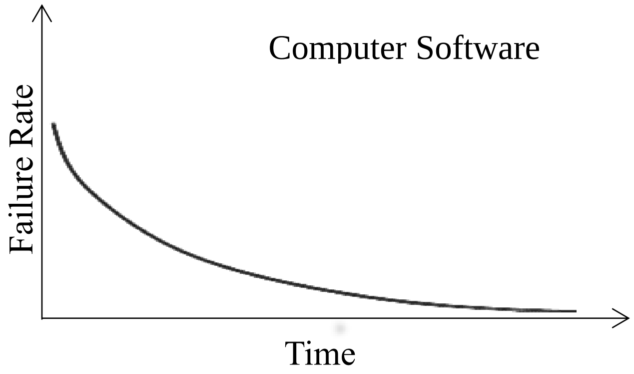


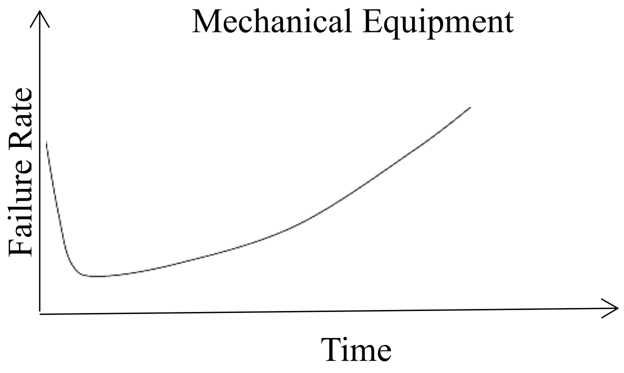
The behavior of h(t), the failure rate, is quite revealing with respect to the causes of failure. Unless a system has redundant components, it will invariably have the general characteristic of a “Bath tub curve” as shown.

At the very beginning the possibility of defective design or manufacturing or assembly or poor quality of component parts etc. produce a very high hazard on the system and is the significant cause of failures. Thus h(t) has a very high value at the beginning. This region is referred to as infantile or early failure zone. h(t) decreases rapidly as the items with such above mentioned defects are eliminated. Those which survive the infant mortality attain a constant failure rate, where failures due to chance is predominant. This persists for some time. This region is referred to as useful life zone and failures during this period of time are random failures. Possible causes may be due to external loading of the system or its components. In the final stage h(t) starts rising as items fail due to aging and wear and tear. This region is referred as ageing zone.

It should be noted that no single failure distribution has a failure rate which satisfies the bathtub model but one of the three phases may be predominant for a particular class of system.







Random failures constitute the most widely used model for describing reliability phenomena. They are defined by the assumption that the rate of failure of a system is independent of its age and other characteristics of its operating history, which in turn implies a constant failure rate. The constant failure rate approximation is often quite adequate even though a system or some of its components exhibit moderate, early failures or ageing effects. Here, we shall consider the exponential distribution that is employed when constant failure rates adequately describe behavior of systems.

The constant failure rate model for continuously operating systems leads to an exponential distribution, i.e. with $h(t) = \lambda$ we have,

$f(t) = \lambda e^{-\int_0^t \lambda \, dx}$

$f(t) = \lambda e^{-\lambda t}$; $\lambda>0$ $t>0$

$\therefore F(t) = 1 - e^{-\lambda t}$

$R(t) = 1 - F(t)$

$R(t)=e^{-\lambda t}$

Mean failure time = $1/\lambda$

Since $h(t)=\lambda$, a constant, this implies that aging has no impact on the component. Thus, the probability that it will fail during some period of time in the future is independent of its age, which is the memory less property of exponential distribution.

Mean Time To Failure (MTTF):

This is just the expected value or mean value of the rv T and is given by,

$$\begin{aligned} MTTF &= \int_0^\infty t f(t) dt \\ &= \int_0^\infty t \left[-\frac{d}{dt}[R(t)]\right] dt \\ &= -tR(t)\Big|_0^\infty + \int_0^\infty R(t) dt \end{aligned}$$

Since $R(0)=1$ and $R(\infty) = 0$ the 1st term vanishes

$\therefore MTTF=\int_0^\infty R(t) dt$

Or $\equiv MTTF = \int_0^\infty e^{-\int_0^t h(x) dx} dt$

When, $h(t)=\lambda$ i.e. for an exponential distribution

$MTTF=\int_0^\infty e^{-\lambda t} \, dt$

$MTTF=1/\lambda$.

Further, more (greater value) is MTTF more reliable is the system/device.

Note: Strictly speaking, MTTF should be used in the case of simple components which are not repaired when they fail, but are replaced by good components. Similarly MTBF should be used with repairable equipment or systems.

It has become customary however, to use MTTF for both non-repairable and repairable equipments. In any case, it represents the same statistical concept of mean time at which failure occurs.
This mean must be known in order to make probabilistic calculations which are necessary for the evaluation of reliability of components and systems. The MTTF is usually considered as mean time for first failure. While MTBF is normally conceived as being mean time between n^{th} and $(n+1)^{th}$ failure in a system, where n is relatively large.

Improvement in Reliability:

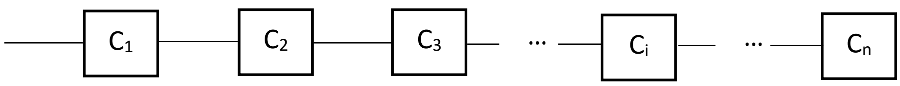
Many systems generally consist of a number of components. The reliability of entire system is a function of reliability of its components. These components may be connected in series or parallel. The system configuration could be a complex one also or a combination of series and parallel structure.
The reliability of a system can be improved in any one of the following ways.

i. Simplification of circuit design
ii. Increasing component reliability
iii. By maintenance of components and devices
iv. Carrying-out repair work of failed units.
v. By having standby units
vi. Using skilled operators

Series Configuration:

A series system is one in which all components are connected in such a way that the system fails if any one of its components fails.

Consider a system having n components connected in series as shown.



Let $R_i(t)$ be the reliability of i^{th} component i=1,2,…,n
It is assumed that each component functions independently of other components.
Then, the system reliability $R_s(t)$, is given by

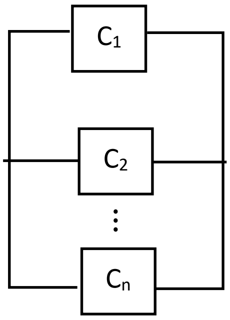
$R_s(t)$ = P(System works)=P(All components work)
= P(C_1 works and C_2 works and …and C_n works)
= P(C_1 works) P(C_2 works/C_1 works)… P(C_n works/C_{n-1} works)
= P(C_1 works) P(C_2 works) … P(C_n works)(since they are independent)
= $R_1(t)$. $R_2(t)$… $R_n(t)$

$$R_s(t) = \prod_{i=1}^{n} R_i(t)$$

If $R_i(t) = e^{-\alpha_i t}$ with constant failure rate α_i.
$\therefore\ R_s(t) = e^{-\alpha t}$; $\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_n = \sum_{i=1}^{n} \alpha_i$
$\Rightarrow$ α is failure rate of the system
Thus the MTTF of the system is given by
$MTTF_S = \int_0^\infty R_s(t)\, dt = \frac{1}{\alpha}$
It can be shown that
$\frac{1}{MTTF_S} = \frac{1}{(MTTF)_1} + \frac{1}{(MTTF)_2} + \cdots + \frac{1}{(MTTF)_n}$

Parallel Configuration:

A parallel system is one in which all components are connected in such a way that atleast one component must function for system to function.
Consider a system consisting of n components as shown

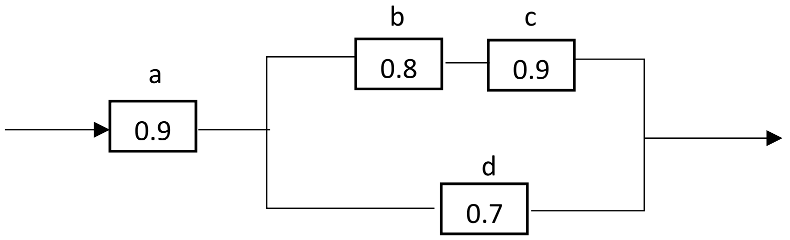


Let $R_i(t)$ be reliability of i^{th} component. It is assumed that each component functions independent of other components.
Then the system reliability $R_s(t)$ is given by,
$R_s(t)$ = P(System works) = P(atleast one component works)
= 1-P(System does not work)
= 1-P(all components fail) = 1-P(C_1 fails and C_2 fails and … and C_n fails)
= 1- P(C_1 fails) P(C_2 fails/C_1 fails)… P(C_n fails/C_{n-1} fails)
= 1- P(C_1 fails) P(C_2 fails) … P(C_n fails) (since all are independent)
= 1-[(1- $R_1(t)$)(1-$R_2(t)$)…(1-$R_n(t)$)]
$R_s(t) = 1 - \prod_{i=1}^{n}(1 - R_i(t))$

If $R_i(t) = e^{-\alpha_i t}$ with constant failure rate α_i.
$R_s(t)$ = 1-[(1-$e^{-\alpha_1 t}$) (1-$e^{-\alpha_2 t}$) (1-$e^{-\alpha_3 t}$)… (1-$e^{-\alpha_n t}$)]
If $\alpha_i = \alpha\ \forall$ i then we have
$R_s(t) = 1 - (1 - e^{-\alpha t})^n$
Thus, the MTTF of the system is given by
$MTTF_S = \int_0^\infty R_S(t)dt$
$= \int_0^\infty 1 - (1 - e^{-\alpha t})^n\, dt$
$= \sum_{j=1}^{n} (-1)^{j-1} \binom{n}{j} \frac{1}{j\alpha}$
$\Rightarrow MTTF_S = \frac{1}{\alpha}\left[1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}\right]$

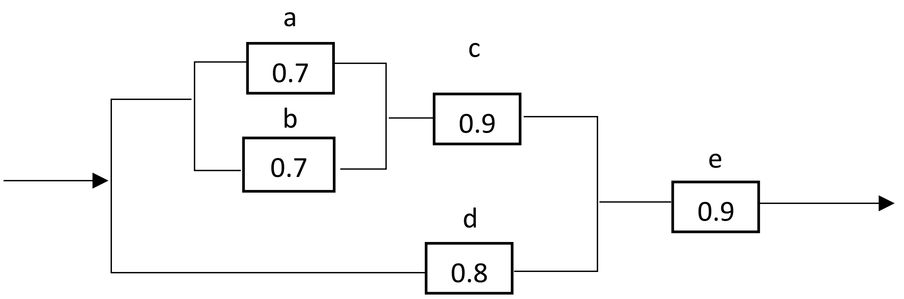
Problems

1. Calculate the system reliability for the units connected as shown



Solve it!

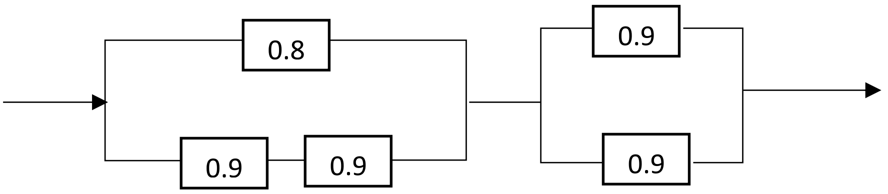
2. Obtain the system reliability



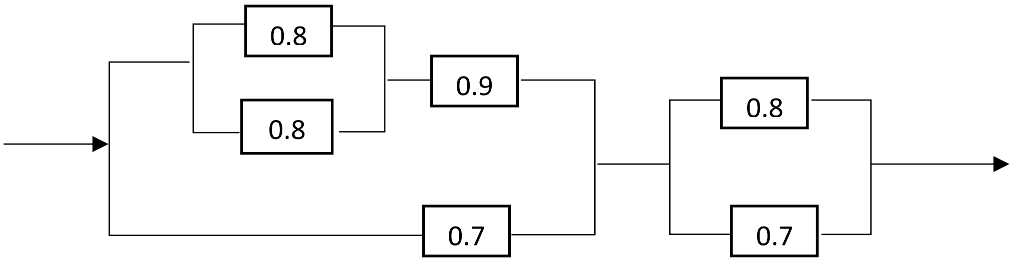
Solve it!

3. Calculate system reliability R_s for following configuration

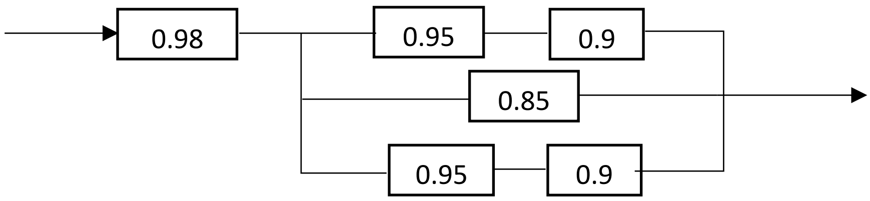
(a).



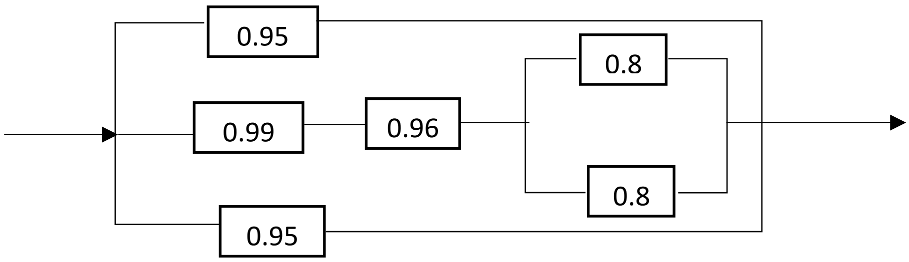
(b).



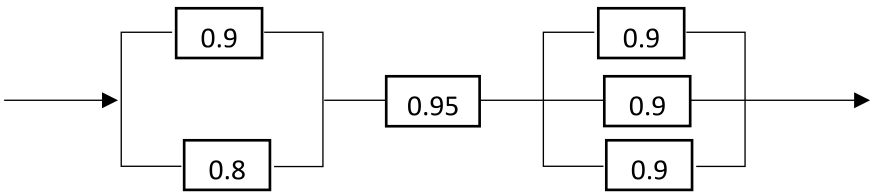
(c).



(d).



(e).

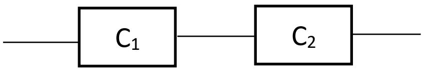


Solve all the above!

4. An equipment consists of 100 parts of which 20 parts are tubes connected functionally in series (Branch A). This branch is in turn connected in series to 2 parallel branches of 20 parts (Branch B) and 60 parts (Branch C). The parts which comprise these branches are connected in series. The reliability of each tube in A is 0.99 while that in B is 0.98 and that in C is 0.94. Calculate the Reliability R_S of the system.

   Solve it!

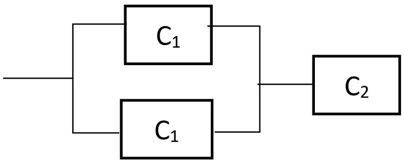
Consider two components C_1 and C_2 with reliabilities R_1 and R_2 respectively connected as shown below



…………… Design 1

then $R_S=R_1R_2$

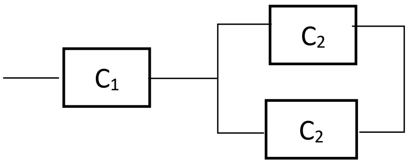
If we replace component C_1 by two parallel components of C_1 as below



…………… Design 2

then $R_S = (1-(1-R_1)^2)R_2$
$= (2-R_1)R_1R_2$

Similarly, if we replace component C_2 by two parallel components of C_2 as below



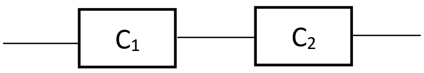
…………… Design 3

then $R_S = (1-(1-R_2)^2)R_1$
$= (2-R_2)R_1R_2$

Which design is to be preferred?
If $R_1>R_2$ then Design 3 is preferred.
If $R_2>R_1$ then Design 2 is preferred.

Problems (Contd…)

5. Consider a system with 2 components say C_1 and C_2, having the same cost but with reliabilities $R_1=0.7$ and $R_2=0.95$ as shown below. If we are allowed to add 2 components to the system, would it be preferable to replace component C_1 by 3 components of C_1 in parallel or to replace the components C_1 and C_2 each by simple parallel components.



Solve it!

6. An engineer approximates the reliability of a cutting assembly by

$$R(t) = \begin{cases} \left(1 - \frac{t}{t_0}\right)^2 ; & 0 \le t < t_0 \\ 0 ; & t \ge t_0 \end{cases}$$

Determine the failure rate. Does it increase or decrease with time? Also find the MTTF.

Solution:

We have $h(t) = \dfrac{f(t)}{R(t)}$

where $f(t) = -\frac{d}{dt}[R(t)] = -\frac{d}{dt}\left(1 - \frac{t}{t_0}\right)^2$

$= \frac{2}{t_0}\left(1 - \frac{t}{t_0}\right) ; 0 \le t \le t_0$

$\therefore \quad h(t) = \dfrac{\frac{2}{t_0}\left(1-\frac{t}{t_0}\right)}{\left(1-\frac{t}{t_0}\right)^2}$

$$h(t) = \begin{cases} \frac{2}{t_0}\left(1 - \frac{t}{t_0}\right) ; & 0 \le t \le t_0 \\ 0 & ; t \ge t_0 \end{cases}$$

Now as time increases the failure rate h(t) increases

since $h(t) = \frac{2}{t_0} ; t = 0$

and $h(t) = \infty ; t = t_0$

Also, MTTF$=\int_0^\infty R(t)dt = \int_0^{t_0} \left(1 - \frac{t}{t_0}\right)^2 dt = ?$

Obtain it!

7. A device has constant failure rate of 0.02 per hour. Obtain the probability that it will fail during first 10 hours of operations. Suppose the device has been successfully operated for 100 hours, what is the probability that it will fail during next 10 hours of operation? Also give comments.

Solve it!

8. A critical measuring instrument consists of 2 subsystems A and B connected in series having reliabilities 0.9 and 0.92 respectively, for a certain operating time. It is necessary that reliability of instrument be raised to a minimum value of 0.917 by using parallel subsystems of A alone. Determine how many units of A should be used with one unit of B to get minimum reliability of 0.918. What is the actual value of the reliability obtained? Can you use 2 units of A and 2 units of B to achieve the desired result? What is the reliability of the system now?

Solve it!

9. The life of an electronic tube is exponentially distributed. It is known that reliability of device for 100 hours period is 0.9. How many hours of operation have considered to achieve a reliability of 0.95?

   Solve it!

10. Suppose that there are 3 components with constant failure rate λ=0.01 connected in parallel. Estimate the improvement in reliability over a period of 10 hours.

    Solve it!

11. Consider a unit having a constant failure rate of 0.3% per 1000 hours. Obtain the MTTF. What are the probabilities of its successful completion of missions of 10,000 hours and 1,00,000 hours.

    Solve it!

12. An airborne electronic system has a radar, computer and an auxiliary unit with MTTFs of 83 hours, 167 hours and 500 hours and equivalent failure rates (in failures per 1000 hours) of 12, 6 and 2 respectively. Find the system reliability and the MTTF for 5 hours operating time. Also find the reliability of subsystems.

    Solve it!

13. Suppose that 2 independently functioning components with failure rates α_1 and α_2 are connected in parallel. Let T be the time to failure of resulting system. Obtain mgf of T.

    Solve it!

DATA ANALYSIS

Classification of data:

Usually any data which is available as raw data may be written in the form of a table, for better understanding of the data, by adopting the following steps:

(i). Grouping of data using tally mark.
(ii). Formation of classes - width of the class is called as class interval.
(iii). Writing down the Frequencies

The table showing the classes and the corresponding frequencies is called the frequency table.

Thus the set of raw data summarized by distributing it into a No. of classes along with their frequencies is known as a frequency distribution.

Averages or Measures of Central Tendency:

A frequency distribution, in general shows clustering of data around some central value. An average is the central value of the frequency distribution which is the most representative value of the entire distribution.

The following or the measures of central tendency:
i) Mean ii) Median iii) Mode iv) Geometric Mean v) Harmonic Mean

Arithmetic Mean or Mean:

The Arithmetic Mean or Mean of a set of observations is their sum divided by the No. of observations. i.e. The Arithmetic Mean, denoted by $\bar{x}$, of n observations $x_1, x_2, \dots, x_n$ is

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

In the case of a frequency distribution, where f_i is the frequency of x_i ; i=1,2,…,n, we have

$$\bar{x} = \frac{f_1x_1 + f_2x_2 + \cdots + f_nx_n}{f_1 + f_2 + \cdots + f_n} = \frac{\sum f_i x_i}{\sum f_i}$$

The same formula will hold for a grouped data, where x_i will then be the mid-points of the clases.

Calculation of Mean:

(i). By taking the deviations of the values of x_i from an arbitrary point A:
Let $d_i = x_i - A$
$\Rightarrow \bar{x} = A + \frac{\sum f_i d_i}{\sum f_i}$

(ii). In case of grouped data:
Let $d_i = \frac{x_i - A}{h}$; where h is the width of the class interval
$\Rightarrow \bar{x} = A + h\ \frac{\sum f_i d_i}{\sum f_i}$

Measures of Dispersion:

The measures of Central Tendency gives us an idea of the concentration of the observations about a central value but they are inadequate in telling us the complete distribution. i.e. how an individual value differs from this central value. Thus, they must be supported and supplemented by some other measures. One such measure is Dispersion.

The following are the measures of dispersion:
(i) Range (ii) Quantile deviation (iii) Mean deviation (iv) Standard Deviation

The square of the standard deviation is known as the Variance

Standard Deviation (or Variance)::

The standard deviation (σ) is the square root of the mean of the squares of the deviations of the given values from their mean and is given by

$\sigma = \sqrt{\frac{1}{n}\sum(x_i - \bar{x})^2}$; where n is the No. of observations

or, $\sigma = \sqrt{\frac{1}{N}\sum f_i(x_i - \bar{x})^2}$; where $N = \sum f_i$ is the total frequency

The square of standard deviation is the variance.

Calculation of Standard Deviation (or Variance):

$$\sigma = \sqrt{\frac{1}{n}\sum(x_i - \bar{x})^2}$$

$$\sigma = \sqrt{\frac{1}{n}\left(\sum x_i^2\right) - \bar{x}^2}$$

and $\sigma = \sqrt{\frac{1}{N}\sum f_i(x_i - \bar{x})^2}$

$$= \sqrt{\frac{1}{N}\sum f_i x_i^2 - \left(\frac{1}{N}\sum f_i x_i\right)^2}$$

(i). If $d_i = x_i - A$ so that $\bar{x} = A + \frac{1}{N}\sum f_i d_i$

$\sigma = \sqrt{\frac{1}{N}\sum f_i d_i^2 - \left(\frac{1}{N}\sum f_i d_i\right)^2}$ $\Rightarrow \sigma^2$ is independent of change of origin.

(ii). If $d_i = \frac{x_i - A}{h}$ so that $\bar{x} = A + \frac{h}{N}\sum f_i d_i$

$\sigma = h\sqrt{\frac{1}{N}\sum f_i d_i^2 - \left(\frac{1}{N}\sum f_i d_i\right)^2}$ $\Rightarrow \sigma^2$ is not independent of change of scale.

Example: Find the Mean and the Standard Deviation for the following data

Age Group (Class Intervals)	No. of members (frequency) f	Mid value x	$d = \frac{x-55}{10}$	fd	fd^2
20-30	03	25	-3	-9	27
30-40	61	35	-2	-122	244
40-50	132	45	-1	-132	132
50-60	153	55	0	0	0
60-70	140	65	1	140	140
70-80	51	75	2	102	204
80-90	02	85	3	06	18
Total	N=542			-15	765

$\therefore$ Mean $= \bar{x} = A + h\frac{\sum fd}{N} = 55 + \frac{10*(-15)}{542}$ = 55-0.28 = 54.72

Variance = $\sigma^2 = h^2\left(\frac{1}{N}\sum fd^2 - \left(\frac{1}{N}\sum fd\right)^2\right) = 100\left(\frac{765}{542} - (0.0285)\right)$

$= 100 * 1.4107$

=141.07

$\Rightarrow$ Standard Deviation $= \sigma = 11.9$

Correlation Coefficient:

The Correlation Coefficient is a measure of the degree of linear relationship between two variables.
Let X and Y be the two variables taking values $(x_1, y_1), (x_2, y_2) \dots (x_n, y_n)$. Then the correlation coefficient r_{xy} or 'r' between X and Y, is defined as

$$r = \frac{\frac{1}{n}\sum(x_i - \bar{x})(y_i - \bar{y})}{\sigma_x \sigma_y} = \frac{\mathrm{cov}(x,y)}{\sigma_x \sigma_y}$$

Calculation of r:

$$\text{I). } r = \frac{n\sum xy - [(\sum x)(\sum y)]}{\sqrt{[\{n\sum x^2 - (\sum x)^2\}\{n\sum y^2 - (\sum y)^2\}]}} = \frac{\frac{1}{n}\sum xy - \bar{x}\bar{y}}{\sqrt{\left(\frac{1}{n}\sum x^2 - \bar{x}^2\right)\left(\frac{1}{n}\sum y^2 - \bar{y}^2\right)}}$$

II). If $d_x = x - A$ and $d_y = y - B$, where A and B are arbitrary origins then

$$\bar{x} = A + \frac{\sum d_x}{n} \quad \text{and} \quad \bar{y} = B + \frac{\sum d_y}{n}$$

$$\sigma_x^2 = \frac{\sum d_x^2}{n} - \left(\frac{\sum d_x}{n}\right)^2 \text{ and } \sigma_y^2 = \frac{\sum d_y^2}{n} - \left(\frac{\sum d_y}{n}\right)^2$$

$$r = \frac{n\sum d_x.d_y - (\sum d_x)(\sum d_y)}{\sqrt{[n\sum d_x^2 - (\sum d_x)^2]\,[n\sum d_y^2 - (\sum d_y)^2]}}$$

III). If $U = x - \bar{x}$ and $V = y - \bar{y}$
then $\sigma_x^2 = \frac{\sum U^2}{n}$ and $\sigma_y^2 = \frac{\sum V^2}{n}$

$$r = \frac{\sum UV}{\sqrt{\sum U^2 \sum V^2}};$$

Note: r lies between -1 and +1 i.e. $-1 \le r \le +1$

Example: Find the Correlation coefficient for the following data.

Student	Intelligence ratio		Eng. ratio		U^2	V^2	UV
	x	$(x-\bar{x})=U$	y	$(y-\bar{y})=V$			
A	105	6	101	3	36	9	18
B	104	5	103	5	25	25	25
C	102	3	100	2	9	4	6
D	101	2	98	0	4	0	0
E	100	1	95	-3	1	9	-3
F	99	0	96	-2	0	4	0
G	98	-1	104	6	1	36	-6
H	96	-3	92	-6	9	36	18
I	93	-6	97	-1	36	1	6
J	92	-7	94	-4	49	16	28
Total	990	0	980	0	170	140	92

where $\bar{x} = \frac{990}{10} = 99$ and $\bar{y} = \frac{980}{10} = 98$

∴ Correlation coefficient is given by,

$$r = \frac{\sum UV}{\sqrt{\sum U^2 \sum V^2}} = \frac{92}{\sqrt{170*140}} = \frac{92}{154.3} = 0.596$$

Result:

To show that r_{xy} is $\frac{\sigma_X^2+\sigma_Y^2-\sigma_{X-Y}^2}{2\sigma_X\sigma_Y}$

Proof:

$\sigma_{X-Y}^2 = V(X - Y)$

$\sigma_{X-Y}^2 = V(X) + V(Y) - 2Cov(X, Y)$

$\sigma_{X-Y}^2 = \sigma_X^2 + \sigma_Y^2 - 2r\sigma_X\sigma_Y$

$$\Rightarrow r = \frac{\sigma_X^2+\sigma_Y^2-\sigma_{X-Y}^2}{2\sigma_X\sigma_Y}$$

Aliter:

Let $Z = X - Y$

$\equiv Z_i = X_i - Y_i$

$\bar{Z} = \bar{X} - \bar{Y}$

$\therefore\ Z_i - \bar{Z} = (X_i - \bar{X}) - (Y_i - \bar{Y})$

$\Rightarrow (Z_i - \bar{Z})^2 = [(X_i - \bar{X}) - (Y_i - \bar{Y})]^2$

Taking summation on both sides we get

$\frac{1}{n}\sum(Z_i - \bar{Z})^2 = \frac{1}{n}\sum[(X_i - \bar{X}) - (Y_i - \bar{Y})]^2$

$\sigma_Z^2 = \frac{1}{n}\sum(X_i - \bar{X})^2 + \frac{1}{n}(Y_i - \bar{Y})^2 - 2\frac{1}{n}\sum(X_i - \bar{X})(Y_i - \bar{Y})$

$\sigma_{X-Y}^2 = \sigma_X^2 + \sigma_Y^2 - 2r\sigma_X\sigma_Y \qquad \left[\because r = \frac{\frac{1}{n}\sum(X_i-\bar{X})(Y_i-\bar{Y})}{\sigma_X\sigma_Y}\right]$

On simplification we get

$$r = \frac{\sigma_X^2+\sigma_Y^2-\sigma_{X-Y}^2}{2\sigma_X\sigma_Y}$$

Regression Analysis:

If the variables X and Y in a bivariate distribution are related we will find that the points in the scatter diagram will cluster around some curve called the Curve of Regression. If the curve is a straight line then it is called the Line of Regression and there is said to be a linear regression between the variables. The line of regression is the line which gives the best estimate to the value of one variable for any specific value of the other. In fact, there are two such lines one giving the best possible mean values of Y for each specified value of X and other giving the best possible mean values of X for each specified value of Y. The former is known as the line of regression of Y on X and the latter is known as the line of regression of X on Y. Thus the line of regression is the line of best fit and is obtained using the principles of least squares.

Note: The principle of least squares consists in minimizing the sum of squares of the deviations of the actual values of Y from its estimated values as given by the line of best fit.

Thus, the line of regression of Y on X is given by;

$$(\mathrm{Y} - \overline{\mathrm{Y}}) = \mathrm{r}\frac{\sigma_{\mathrm{Y}}}{\sigma_{\mathrm{X}}}(\mathrm{X} - \overline{\mathrm{X}}) \longrightarrow \qquad (A)$$

and line of regression of X on Y is given by;

$$(\mathrm{X} - \overline{\mathrm{X}}) = \mathrm{r}\frac{\sigma_{\mathrm{X}}}{\sigma_{\mathrm{Y}}}(\mathrm{Y} - \overline{\mathrm{Y}}) \longrightarrow \qquad (B)$$

where ‘r’ is the sample correlation coefficient and σ_{X} and σ_{Y} are the standard deviations of X and Y respectively

We denote the factors:

$\mathrm{b}_{\mathrm{YX}} = \mathrm{r}\frac{\sigma_{\mathrm{Y}}}{\sigma_{\mathrm{X}}}$ and is called the regression coefficient of Y on X.

$\mathrm{b}_{\mathrm{XY}} = \mathrm{r}\frac{\sigma_{\mathrm{X}}}{\sigma_{\mathrm{Y}}}$ and is called the regression coefficient of X on Y.

NOTE:

1. Whenever we have to estimate Y for a given value of X, i.e., Y is dependent and X is independent, we then use equation (A) otherwise we use (B).
2. In the case of perfect correlation i.e., $\mathrm{r} = \pm 1$, the two lines of regression coincide. Thus we have only one line.
3. The correlation coefficient is obtained as the geometric mean of the two regression coefficients. Thus $\mathrm{r} = \pm\sqrt{\mathrm{b}_{\mathrm{XY}}\mathrm{b}_{\mathrm{YX}}}$
4. Both the lines of regression pass through the point $(\bar{\mathrm{x}}, \bar{\mathrm{y}})$, the sample mean.

Line of regression — The line of best fit — Principle of Least squares :

Consider the line of regression of Y on X. Let this straight line be given by $y = a + bx$ —— ①

Now we have to determine the constants a & b s.t ① gives, for each value of x the best value (estimate for the average value) of y. This is done using the Least square Principle.

Accordingly,

Let $S_i = y_i - y$

$\Rightarrow S_i = y_i - (a + bx_i)$

$\therefore S = \Sigma S_i^2 = \Sigma (y_i - (a + bx_i))^2$

$\Rightarrow \frac{\partial S}{\partial a} = 0 \Rightarrow -2\Sigma(y - (a+bx)) = 0$

& $\frac{\partial S}{\partial b} = 0 \Rightarrow -2\Sigma x(y - (a+bx)) = 0$

From which we get the normal equations for a & b as,

$$\Sigma y = na + b\Sigma x \quad \text{—②}$$

$$\Sigma xy = a\Sigma x + b\Sigma x^2 \quad \text{—③}$$

$(\div)$ing ② by $n \Rightarrow \bar{y} = a + b\bar{x}$

$\Rightarrow$ Thus $(\bar{x}, \bar{y})$ i.e. means of x & y lie on ①

Shifting the origin to $(\bar{x}, \bar{y})$, ③ takes the form

$$\Sigma(x - \bar{x})(y - \bar{y}) = a\Sigma(x - \bar{x}) + b\Sigma(x - \bar{x})^2$$

But $a\Sigma(x - \bar{x}) = 0$ (why?)

$$\therefore \Rightarrow b = \frac{\Sigma(x - \bar{x})(y - \bar{y})}{\Sigma(x - \bar{x})^2}$$

or $b = \dfrac{\text{cov}(x,y)}{\sigma_x^2}$

$b = r\dfrac{\sigma_y}{\sigma_x}$ $\quad \left(\because r = \dfrac{\text{cov}(x,y)}{\sigma_x \cdot \sigma_y}\right)$

which is the slope of the line of regression of y on x

Thus the line of best fit becomes (since the line of regression passes through the point $(\bar{x}, \bar{y})$ with slope $b = r\dfrac{\sigma_y}{\sigma_x}$)

$$\Rightarrow y - \bar{y} = r\frac{\sigma_y}{\sigma_x}(x - \bar{x})$$

is the equation of the line of regression of y on x.

Its slope is called the coefficient of regression of y on x and is denoted by b_{yx}.

Interchanging x & y, i.e. starting with the equation $x = a + by$ and proceeding similarly we obtain the equation to the line of regression of x on y as

$$x - \bar{x} = r\frac{\sigma_x}{\sigma_y}(y - \bar{y})$$

Problems

1. Obtain the line of regression of Y on X for the following data and estimate the most probable value of Y when X is 70.

Item No	x	y
1	40	2.5
2	70	6.0
3	50	4.5
4	60	5.0
5	80	4.5
6	50	2.0
7	90	5.5
8	40	3.0
9	60	4.5
10	60	3.0

Solution:

The line of regression of Y on X is given by:

$$(\mathrm{Y}-\overline{\mathrm{Y}}) = \mathrm{r_{xy}}\frac{\sigma_\mathrm{Y}}{\sigma_\mathrm{X}}(\mathrm{X}-\overline{\mathrm{X}})$$

Item No	X	Y	d_x=x-60	d_y=y-4.5	d_x^2	d_xd_y
1	40	2.5	-20	-2	400	40
2	70	6.0	10	1.5	100	15
3	50	4.5	-10	0	100	0
4	60	5.0	0	0.5	0	0
5	80	4.5	20	0	400	0
6	50	2.0	-10	-2.5	100	25
7	90	5.5	30	1.0	400	30
8	40	3.0	-20	-1.5	400	300
9	60	4.5	0	0	0	0
10	60	3.0	0	-1.5	0	0
			$\sum d_x = 0$	$\sum d_y = -4.5$	$\sum d_x^2 = 2400$	$\sum d_xd_y = 140$

With A=60 and B=4.5 we have,

$$\bar{x} = A + \frac{\sum dx}{n} = 60 + 0 = 60 \text{ and } \bar{y} = B + \frac{\sum dy}{n} = 4.5 + \left(\frac{-4.5}{10}\right) = 4.05$$

$$\sigma_x^2 = \frac{\sum d_x^2}{n} - \left(\frac{\sum d_x}{n}\right)^2 \text{ and } \sigma_y^2 = \frac{\sum d_y^2}{n} - \left(\frac{\sum d_y}{n}\right)^2$$

$$r = \frac{n\sum d_x.d_y - (\sum d_x)(\sum d_y)}{\sqrt{[n\sum d_x^2 - (\sum d_x)^2]\,[n\sum d_y^2 - \left(\sum d_y\right)^2]}}$$

$$r.\frac{\sigma_y}{\sigma_x} = \frac{\sum dxdy - \frac{\sum dx \sum dy}{n}}{\sum d^2x - \frac{(\sum dx)^2}{n}}$$

$$= \frac{140-0}{2400-0} = \frac{140}{2400}$$

= 0.06

Thus, the required line of regression of Y on X is given by:

$$(Y - \bar{Y}) = r_{xy}\frac{\sigma_Y}{\sigma_X}(X - \bar{X})$$

$$Y - 4.05 = 0.06(X - 60)$$

$$\Rightarrow Y = 0.06X + 0.45$$

Which is the line of the equation of Y on X.

Now, when X = 70, we have

Y = 0.06*70+0.45 = 4.65

Thus Y=4.65 is the most probable estimated value of Y when X=70.

2. Obtain the line of regression of X and Y for the above data

Solve it!

3. The regression equations of two variables X and Y are X=0.7Y+5.2 and Y=0.3X+2.8. Find the means of the variables and the correlation coefficient.

Solution:

Since both the lines of regression passes through the point $(\overline{X}, \overline{Y})$ we have,

$\overline{X} = 0.7\overline{Y} + 5.2$ and $\overline{Y} = 0.3\overline{X} + 2.8$

Solving we get,

$\overline{X} = 9.06$

$\overline{Y} = 5.518$

Now, regression coefficient of Y on X is: b_{yx}=0.3

regression coefficient of X on Y is: b_{xy}=0.7

$\therefore\ r = \sqrt{b_{yx}b_{xy}} = \sqrt{0.3 * 0.7} = \sqrt{0.21} = 0.46$

4. The heights of 12 fathers and sons are given below:

Height of father x	165	160	170	163	173	158	178	168	173	170	175	180
Height of son y	173	168	173	165	175	168	173	165	180	170	173	178

Obtain the two regression lines and hence obtain r, the correlation coefficient

Solve it!

5. The equations of two lines of regression are 4X+3Y+7=0 and 3X+4Y+8=0.
Find (i) the means of the variables X and Y
(ii) the Regression coefficients b_{yx} and b_{xy}
(iii) the Correlation coefficient r, between X and Y.

Solve it!

Multiple-Linear Regression:

In many of the real life situations it may happen that the dependent variable Y can more adequately be predicted when there are more than one independent variable, say, $X_1, X_2, X_3 \ldots X_k$. Typically one may have a relationship of the type:
$Y = a + b_1X_1 + b_2X_2 + \cdots + b_kX_k$
One can obtain the coefficients a, b_1, b_2, …, b_k, as earlier, by using the Principle Least Squares and hence obtain the line of regression. This is called the multiple regression.

Non-linear Regression:

Quite often it is observed while using linear regression that the estimated (predicted) values of the dependent variable produce poor results. One reason behind this could be due to the fact that the variables may be far from being linearly related but a Curvi-linear relationship may be more appropriate.
For example, instead of $Y = a + bX$, a relationship of second degree, say of the type,
$Y = a + bX + cX^2$ may be more appropriate. By using Principle of Least Squares, one may obtain the coefficients a, b and c and hence arrived the regression equation which is non-linear.