Dr.R.RAVIKUMAR is an Associate Professor in Mathematics at St. Joseph's College of Engineering and Technology, Thanjavur. He received his Ph.D degree from Anna University, Chennai. He has more than 25 years of teaching experience in Engineering colleges. He already authored three books for B.E/B.Tech Students. Also he has published more than 15 papers in reputed national and international journals and conferences.

R.MALARKODI is the Head of the Department of Science and Humanities at St. Joseph's College of Engineering and Technology, Thanjavur. She received his Master degree from Bharathidasan University, Trichy. She has more than 16 years of teaching experience in Engineering colleges. She already authored two books for B.E/B.Tech Students. She has organized many FDPs and Conferences.

Dr.N.LATHA is an Assitant Professor in Mathematics at St. Joseph's College of Engineering and Technology, Thanjavur. She has more than 23 years of teaching experience in various institutions which includes 14 years in Engineering colleges only. She has Published 7 research papers in Mathematical Modelling in Conferences and international journals.

Key Features of the Book : Based on the authors combined experience in teaching, the book has the following unique features :

* Introduces each topics in an easy way
* Sufficient number of examples have been given
* Taken care to solve university examination questions
* Recent Anna University examination question papers are included.

ISBN 978-81-981760-0-4

9 788198 176004

PUBLISHER

ST. JOSEPH'S COLLEGE OF ENGINEERING AND TECHNOLOGY

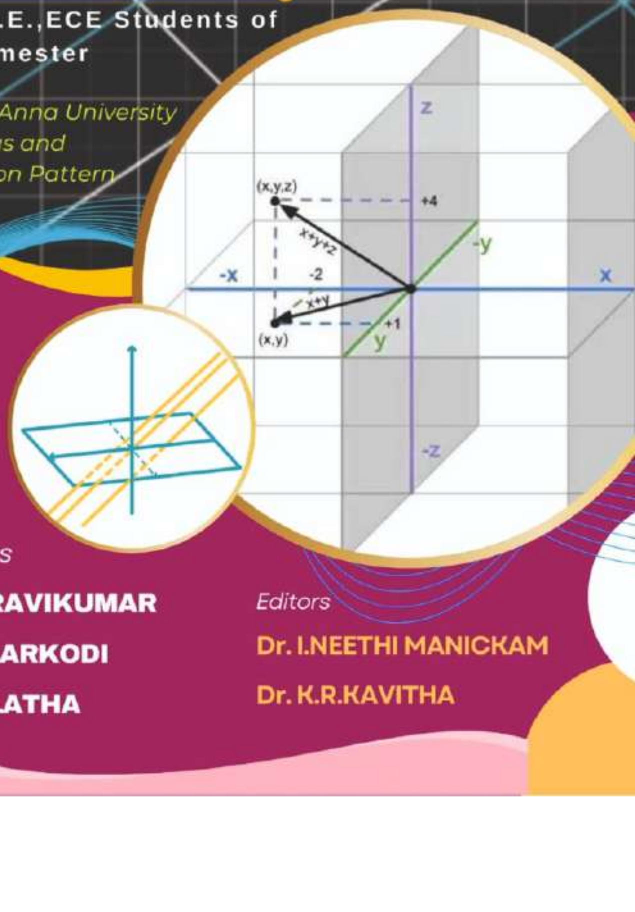
ELUPATTI, THANJAVUR – 613 403.

RANDOM PROCESSES AND LINEAR ALGEBRA

(MA 3355 REGULATIONS 2021)

For B.E.,ECE Students of III Semester

As Per Anna University Syllabus and Question Pattern



Authors

Dr. R.RAVIKUMAR

R.MALARKODI

Dr. N.LATHA

Editors

Dr. I.NEETHI MANICKAM

Dr. K.R.KAVITHA

Unit-I Probability and Random Variables

INTRODUCTION

Most of the decision making problems are related to uncertainty. The notion of uncertainty or chance is so common in everybody's life that it becomes difficult to define it. In our day to day life, every thing happening is a matter of chance. We talk about chances of one's winning the election, getting a handsome job etc.

The term probability means likelihood or chance or possibility. If an event is likely to occur, we say that it is probable. Some people would prefer to call it luck. Otherwise it is improbable or failure event. Some of the business situations like investment problem, introducing new product, stocking decisions, individual investor are characterized by uncertainty.

Probability theory and stochastic differential equations (SDEs) are essential tools for analyzing randomness and uncertainty in electrical and communication systems. Probability theory provides a mathematical framework for quantifying uncertainty and randomness in engineering systems.

DEFINITIONS

SAMPLE SPACE : The set of all possible outcomes of some given experiment is called the sample spaces.

EXAMPLE :

(i) In tossing a coin, sample space $S=\{H, T\}$

(ii) In rolling a die, $S=\{1, 2, 3, 4, 5, 6\}$

(iii) If two fair coins are thrown simultaneously, $S=\{HH, TT, TH, HT\}$

EVENT : An event A is a set of outcomes (or) a subset of the sample space S.

NOTE:

(i) The event $A=\{a\}$ consisting of a single element $a \in S$ is called an elementary event.

(ii) ϕ and S are also events. ϕ – impossible event ; S – sure event

MUTUALLY EXCLUSIVE EVENTS : Two events A and B are called mutually exclusive if A and B are disjoint.

(i.e) if $A \cap B = \phi$

(i.e) if A and B cannot appear simultaneously.

Note: If events A and B are mutually exclusive, then $P\left(A \cap \overline{B}\right) = P(A)$

INDEPENDENT EVENTS : If A and B are independent events then $P(A \cap B) = P(A) \cdot P(B)$

https://doi.org/10.5281/zenodo.15089962

AXIOMS OF PROBABILITY :

Let S be the sample space, let Σ be the class of events and let P be a real valued function defined on Σ. Then P is called a probability function and $P(A)$ is called the probability of the event if the following axioms hold.

(i) For any event A, $0 \le P(A) \le 1$.

(ii) $P(S) = 1$

(iii) If A and B are mutually exclusive events then $P(A \cup B) = P(A) + P(B)$.

THEOREMS ON PROBABILITY :

Theorem 1 : If ϕ is the empty set, $P(\phi) = 0$. Clearly, $A = A \cup \phi$. A, ϕ are mutually exclusive events, since $A \cap \phi = \phi$. Now $P(A) = P(A \cup \phi) = P(A) + P(\phi)$ $P(\phi) = 0$.	**Theorem 2 :** If $\overline{A}$ is the complement of an event A, then $P(\overline{A}) = 1 - P(A)$. Since A and $\overline{A}$ are complement, they are mutually exclusive, hence $A \cap \overline{A} = \phi$. We know that $S = A \cup \overline{A}$ $P(S) = P(A \cup \overline{A})$ $1 = P(A) + P(\overline{A})$ $P(\overline{A}) = 1 - P(A)$

.

Theorem 3 : If $A \subset B$ then $P(A) \le P(B)$ and $P(B - A) = P(B) - P(A)$. Clearly, $B = A \cup (B - A)$ $P(B) = P[A \cup (B - A)]$ But A and $(B - A)$ are mutually exclusive events. Therefore $P(B) = P(A) + P(B - A)$ $P(B) - P(A) \ge P(B - A)$ $P(B) - P(A) \ge 0$ $P(B) \ge P(A)$	**Theorem 4 :** If A and B are any two events, $P(A - B) = P(A) - P(A \cap B)$. Clearly, $A = (A - B) \cup (A \cap B)$ $P(A) = P[(A - B) \cup (A \cap B)]$ But $A - B$ and $(A \cap B)$ are mutually exclusive events. Therefore $P(A) = P(A - B) + P(A \cap B)$ $P(A - B) = P(A) - P(A \cap B)$

.

Theorem 5: (Addition Theorem) If A and B are any two events, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. Clearly, $A \cup B = A \cup (B - A)$ $P(A \cup B) = P(A \cup (B - A))$	Also $B = (B - A) \cup (A \cap B)$ $P(B) = P[(B - A) \cup (A \cap B)]$

But A and $B-A$ are mutually exclusive events. Therefore $$P(A\cup B)=P(A)+P(B-A)\ldots.(1)$$	But $A\cap B$ and $B-A$ are mutually exclusive events. Therefore $$P(B)=P(B-A)+P(A\cap B)$$ $$P(B-A)=P(B)-P(A\cap B)\ldots.(2)$$
Substitute (2) in (1), we have $P(A\cup B)=P(A)+P(B)-P(A\cap B)$ **Corollary:** If A, B and C are any three events, $P(A\cup B\cup C)=P(A)+P(B)+P(C)-P(A\cap B)-P(B\cap C)-P(C\cap A)+P(A\cap B\cap C)$.	

Theorem 6 : If A and B are independent events then (i) A and $\overline{B}$ are independent (ii) $\overline{A}$ and B are independent (iii) $\overline{A}$ and $\overline{B}$ are independent

Given that A and B are independent. Therefore $P(A\cap B)=P(A)\cdot P(B)\ldots..(1)$

We know that $(A\cap\overline{B})=A-(A\cap B)$ Therefore $P(A\cap\overline{B})=P(A)-P(A\cap B)$ $=P(A)-P(A)\cdot P(B)$ $=P(A)[1-P(B)]$ $=P(A)\cdot P(\overline{B})$ Hence A and $\overline{B}$ are independent	We know that $(\overline{A}\cap B)=B-(A\cap B)$ Therefore $P(\overline{A}\cap B)=P(B)-P(A\cap B)$ $=P(B)-P(A)\cdot P(B)$ $=P(B)[1-P(A)]$ $=P(B)\cdot P(\overline{A})$ Hence $\overline{A}$ and B are independent
We know that $\overline{(A\cup B)}=(\overline{A})\cap(\overline{B})$ $P(\overline{A}\cap\overline{B})=P\overline{(A\cup B)}$, {De Morgan's law} $=1-P(A\cup B)$ $=1-[P(A)+P(B)-P(A\cap B)]$ $=1-P(A)-P(B)+P(A\cap B)$ $=1-P(A)-P(B)+P(A)P(B)$ $=[1-P(A)]-P(B)[1-P(A)]$ $=[1-P(A)]\cdot[1-P(B)]$ $=P(\overline{A})\cdot P(\overline{B})$ Hence, $\overline{A}$ and $\overline{B}$ are independent.	

Theorem 8 : If A, B, C are random events in a sample space S and if they are pair wise independent and A is independent of $B\cup C$, then A, B, C are mutually independent.

Given that A, B, C are pair wise independent.
Therefore $P(A\cap B)=P(A)\cdot P(B)$, $P(B\cap C)=P(B)\cdot P(C)$, $P(C\cap A)=P(C)\cdot P(A)\ldots..(1)$

Also A is independent of $B\cup C$. Then $P(A\cap(B\cup C))=P(A)\cdot P(B\cup C)\ldots..(2)$

$$P\left[(A\cap B)\cup(A\cap C)\right]=P(A)\cdot\left[P(B)+P(C)-P(B\cap C)\right]$$

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

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

$$P(A\cap B\cap C)=P(A)\cdot P(B)\cdot P(C)$$

Theorem 9 : For any two events A and B, $P(A\cap B)\le P(A)\le P(A\cup B)\le P(A)+P(B)$

We know that $(A\cap B)\subseteq A$ Also $A\subseteq A\cup B$.

$\therefore\ P(A\cap B)\le P(A)$, by a theorem. $\therefore\ P(A)\le P(A\cup B)$, by a theorem.

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

$P(A\cup B)\le P(A)+P(B)$, since $P(A\cap B)\ge 0$.

Combining the three results, we have $P(A\cap B)\le P(A)\le P(A\cup B)\le P(A)+P(B)$

DEMORGAN'S LAW : (i) $P\left(\overline{A\cap B}\right)=P\left(\overline{A}\right)\cup P\left(\overline{B}\right)$ and (ii) $P\left(\overline{A\cup B}\right)=P\left(\overline{A}\right)\cap P\left(\overline{B}\right)$

DEFINITION : (CONDITIONAL PROBABILITY)

Let A and B are any two events then, $P(B/A)$ is called the conditional probability of occurrence of B when the event A has already happened and $P(A/B)$ is the conditional probability of happening of A when the event B has already happened.

THEOREM 7 : (MULTIPLICATION THEOREM)

For two events A and B, $P(A\cap B)=P(A)\cdot P(B/A),\ P(A)>0$ or

$$P(A\cap B)=P(B)\cdot P(A/B),\ P(B)>0$$

Let N be the size of the sample space and nA be the favourable events for A and nB be the number of events favourable to B and let nAB be the number of favourable events for the compound event $A\cap B$.

Then the unconditional probabilities are, $P(A)=\dfrac{nA}{N},\ P(B)=\dfrac{nB}{N},\ P(A\cap B)=\dfrac{nAB}{N}$.

Now the conditional probability $P(A/B)$ refer to the sample space of nB occurrences, out of which nAB occurrences associate to occurrence of A i.e. when B has already happened.

$\therefore\ P(A/B)=\dfrac{nAB}{nB}$. Similarly $P(B/A)=\dfrac{nAB}{nA}$

But $P(A\cap B)=\dfrac{nAB}{N}$

$$=\frac{nAB}{nA}\cdot\frac{nA}{N}$$

$$=P(B/A)\cdot P(A)$$

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

Also $P(A\cap B)=\dfrac{nAB}{N}$

$$=\frac{nAB}{nB}\cdot\frac{nB}{N}$$

$$=P(A/B)\cdot P(B)$$

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

Hence the conditional probabilities, $P(A/B)$ and $P(B/A)$ are defined iff $P(A)>0$ & $P(A)>0$

Example 1 : If $P(A)=0.35$, $P(B)=0.75$ and $P(A\cap B)=0.15$, find $P(\overline{A}\cup\overline{B})$. $P(\overline{A}\cup\overline{B})=P(\overline{A\cap B})$, By De Morgan's theorem $=1-P(A\cap B)$ $=1-0.15$ $=0.85$	**Example 2 :** If A and B are two mutually exclusive events, why the following assignment of probabilities is not permissible. $P(A)=0.3$, $P(A\cap\overline{B})=0.7$. We know that if events A and B are mutually exclusive, then $P(A\cap\overline{B})=P(A)$. But here they are not equal.

.

Example 3: If $P(A\cup B)=0.8$, $P(A\cap B)=0.4$ and $P(\overline{B})=0.5$ determine $P(A)$ and $P(B)$. $P(B)=1-P(\overline{B})=1-0.5=0.5$ $P(A\cup B)=P(A)+P(B)-P(A\cap B)$ $P(A)=P(A\cap B)+P(A\cup B)-P(B)$ $P(A)=0.4+0.8-0.5$ $=0.7$	**Example 4:** If $P(A)=0.4$, $P(B)=0.6$ and $P(A\cap B)=0.2$ determine $P(A\cap\overline{B})$. We know that $P(A\cap\overline{B})=P(A)-P(A\cap B)$ $=0.4-0.2$ $=0.2$

Example 5 : If two dice are tossed simultaneously what is the probability of getting 4 as the sum of the resultant faces?

$$\text{Sample space } S=\left\{\begin{matrix} (1,1), & (1,2), & (1,3), & ...., & (1,6), \\ (2,1), & (2,2), & (2,3), & ...., & (2,6), \\ (3,1), & (3,2), & (3,3), & ...., & (3,6), \\ \vdots & \vdots & \vdots & ...., & \vdots \\ (6,1), & (6,2), & (6,3), & ...., & (6,6) \end{matrix}\right\}$$

Let A be the Event of getting 4 as the sum. Then $A=\{(1,3),\ (3,1),\ (2,2)\}$

Therefore probability of getting 4 as the sum of the resultant faces is

$$P(A) = \frac{n(A)}{n(S)} = \frac{3}{36}.$$

Example 6: What is the probability of getting at least one head when two coins are tossed?

Sample space $S = \{HH, TT, TH, HT\}$

Let A be the event of getting at least one head . Then $A = \{HH, TH, HT\}$

$\therefore$ Probability of getting at least one head is $P(A) = \dfrac{n(A)}{n(S)} = \dfrac{3}{4}$

Example 7: If 4 balls are drawn at random from a bag containing 7 white and 6 black balls what is the probability that 3 are white?

Total number of balls is $7+6=13$
Total number of ways of getting 4 balls $= 13C4$
Total number of ways of getting 3 white balls $= 7C3$
Total number of ways of getting 1 black ball $= 6C1$

Let A be the event of getting 4 balls in which 3 is white and one is black ball.
The number of ways of selecting 4 balls (3W+1B) $= 7C3 \times 6C1$

$\therefore$ Probability of getting 3 white balls while taking 4 balls $P(A) = \dfrac{n(A)}{n(S)} = \dfrac{7C3 \times 6C1}{13C4} = \dfrac{42}{143}$

Example 7: Two dice are thrown. What is the probability that the sum is (a) greater than 9 (b) neither 8 nor 11

Let S denotes the sample space when two dice are thrown. Then $n(S) = 36$

(a) Let A be the event of getting the sum of faces greater than 9. i.e. sum = 10 or 11 or 12.

Let A_1 be the event of getting the sum of faces =10, which is possible if $A_1 = \{(4,6), (6,4), (5,5)\}$.

Therefore $P(A_1) = \dfrac{3}{36}$

Let A_2 be the event of getting the sum of faces =11, which is possible if $A_2 = \{(5,6), (6,5)\}$.

Therefore $P(A_2) = \dfrac{2}{36}$

Let A_3 be the event of getting the sum of faces =12, which is possible if $A_3 = \{(6,6)\}$.

Therefore $P(A_3) = \dfrac{1}{36}$

By addition theorem, $P(A) = P(A_1) + P(A_2) + P(A_3)$

$$= \frac{3}{36} + \frac{2}{36} + \frac{1}{36}$$

$$= \frac{1}{6}$$

(b) Let A be the event of getting the sum of faces is neither 8 nor 11.

Let B be the event of getting sum of two dice value is 8 and hence

$B=\{(2,6), (6,2), (3,5), (5,3), (4,4)\}$. Also $P(B)=\frac{5}{36}$

Let C be the event of getting sum of two dice value is 11 and hence $C=\{(5,6), (6,5)\}$. Also $P(C)=\frac{2}{36}$

$\therefore$ The required probability
$$\begin{aligned}
P(A) &= P\left(\overline{B}\cap\overline{C}\right)\\
&= P\left(\overline{B\cup C}\right)\\
&= 1-P(B\cup C)\\
&= 1-\left[P(B)+P(C)\right], \ (B, C \text{ are m.e.})\\
&= 1-\left[\frac{5}{36}+\frac{2}{36}\right]\\
&= \frac{29}{36}
\end{aligned}$$

Example 8 : A problem of statistics is given to three students A, B and C whose chances of solving it are $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}$ respectively. What is the probability that

i) no one will solve the problem
ii) only one will solve the problem
iii) at least one will solve the problem (or) the problem will be solved?

Let A be the event of solving the problem by the student A. Given that $P(A)=\frac{1}{2}$.	Let $\overline{A}$ be the event of not solving the problem by the student A. Therefore $P\left(\overline{A}\right)=\frac{1}{2}$.
Let B be the event of solving the problem by the student B. Given that $P(B)=\frac{1}{3}$.	Let $\overline{B}$ be the event of not solving the problem by the student B. Therefore $P\left(\overline{B}\right)=\frac{2}{3}$.
Let C be the event of solving the problem by the student C. Given that $P(C)=\frac{1}{4}$.	Let $\overline{C}$ be the event of not solving the problem by the student C. Therefore $P\left(\overline{C}\right)=\frac{3}{4}$.
Here the events A, B and C are independent.	Here the events $\overline{A}$, $\overline{B}$ and $\overline{C}$ are independent.

(i) probability that no one will solve the problem is $=P\left(\overline{A}\cap\overline{B}\cap\overline{C}\right)$

$$= P\left(\overline{A}\right)\cdot P\left(\overline{B}\right)\cdot P\left(\overline{C}\right)$$

$$= \frac{1}{2}\cdot\frac{2}{3}\cdot\frac{3}{4}$$

$$= \frac{1}{4}$$

(ii) probability that only one will solve the problem

$$= P\left(A\cap\overline{B}\cap\overline{C}\right)+P\left(\overline{A}\cap B\cap\overline{C}\right)+P\left(\overline{A}\cap\overline{B}\cap C\right)$$

$$= P(A)\cdot P\left(\overline{B}\right)\cdot P\left(\overline{C}\right)+P\left(\overline{A}\right)\cdot P(B)\cdot P\left(\overline{C}\right)+P\left(\overline{A}\right)\cdot P\left(\overline{B}\right)\cdot P(C)$$

$$= \frac{1}{2}\cdot\frac{2}{3}\cdot\frac{3}{4}+\frac{1}{2}\cdot\frac{1}{3}\cdot\frac{3}{4}+\frac{1}{2}\cdot\frac{2}{3}\cdot\frac{1}{4}$$

$$= \frac{11}{24}$$

(iii) probability that at least one will solve the problem(the problem will be solved)

= 1 – probability that no one will solve the problem

$$= 1-\frac{1}{4}$$

$$= \frac{3}{4}$$

Example 9 : A and B throw alternately a pair of dice. A wins if he throws 6 before B throws 7 and B wins if he throws 7 before A throws 6. If A throws first what is his chance of winning?

Let S denotes the sample space when two dice are thrown. Then $n(S)=36$

Let A be the event of getting sum 6 (i.e. A wins)

Sample space for the event $A=\left\{(1,5),\ (5,1),\ (2,4),\ (4,2),\ (3,3)\right\}$. Therefore $P(A)=\frac{5}{36}$

Let $\overline{A}$ be the event of not getting sum 6 (i.e. A fails)

Therefore $P(\overline{A})=1-\frac{5}{36}=\frac{31}{36}$

Let B be the event of getting sum 7 (i.e. B wins)

Sample space for the event $B=\left\{(2,5),\ (5,2),\ (3,4),\ (4,3),\ (6,1),\ (1,6)\right\}$.

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

Let $\overline{B}$ be the event of not getting sum 7 (i.e. B fails)

Therefore $P(\overline{B})=1-\frac{1}{6}=\frac{5}{6}$

Now A throw first. The possibilities of his winning are as follows:

Events	**Probability**
A wins	$P(A)$
or	
A fails, B fails and A wins	$P\left(\overline{A}\cap\overline{B}\cap A\right)$
or	
A fails, B fails, A fails, B fails and A wins	$P\left(\overline{A}\cap\overline{B}\cap\overline{A}\cap\overline{B}\cap A\right)$
and so on	

$$\therefore \text{ probability for } A\text{'s win} = P(A)+P\left(\overline{A}\cap\overline{B}\cap A\right)+P\left(\overline{A}\cap\overline{B}\cap\overline{A}\cap\overline{B}\cap A\right)+\ldots\ldots.$$

$$= P(A)+P\left(\overline{A}\right)\cdot P\left(\overline{B}\right)\cdot P(A)+P\left(\overline{A}\right)\cdot P\left(\overline{B}\right)\cdot P\left(\overline{A}\right)\cdot P\left(\overline{B}\right)\cdot P(A)+\ldots.$$

$$= P(A)\left[1+\left(P\left(\overline{A}\right)\cdot P\left(\overline{B}\right)\right)+\left(P\left(\overline{A}\right)\cdot P\left(\overline{B}\right)\right)^2+\ldots.\right]$$

$$= P(A)\left[1-\left(P\left(\overline{A}\right)\cdot P\left(\overline{B}\right)\right)\right]^{-1}$$

$$= \frac{5}{36}\left[1-\left(\frac{31}{36}\cdot\frac{5}{6}\right)\right]^{-1}$$

$$= \frac{30}{61}$$

Example 9 : A student takes his examination in four subjects A, B, C and D. He estimates his chance of passing the subjects is $\frac{4}{5}$, $\frac{3}{4}$, $\frac{5}{6}$ and $\frac{2}{3}$ respectively. To qualify he must pass in A and at least in two other subjects. What is the probability that he qualifies?

Let E_1, E_2, E_3, E_4 be the event of passing in the subjects A, B, C and D respectively. Given that $P(E_1)=\frac{4}{5}, P(E_2)=\frac{3}{4}, P(E_3)=\frac{5}{6}, P(E_4)=\frac{2}{3}$	Let $\overline{E_1}, \overline{E_2}, \overline{E_3}, \overline{E_4}$ be the event of not passing in the subjects A, B, C and D respectively. Therefore $P\left(\overline{E_1}\right)=\frac{1}{5}, P\left(\overline{E_2}\right)=\frac{1}{4}, P\left(\overline{E_3}\right)=\frac{1}{6}, P\left(\overline{E_4}\right)=\frac{1}{3}$

Given that to qualify, he must pass in the subject A and at least in two other subjects.
$\therefore$ For this the mutually exclusive possibilities are
{Pass in A, B, C and fail in D} or {Pass in A, B, D and fail in C} or {Pass in A, D, C and fail in B} or {Pass in A, B, C and D}

Required probability

$$= P(E_1\cap E_2\cap E_3\cap E_4)+P\left(E_1\cap E_2\cap E_3\cap\overline{E_4}\right)+P\left(E_1\cap E_2\cap\overline{E_3}\cap E_4\right)+P\left(E_1\cap\overline{E_2}\cap E_3\cap E_4\right)$$

$$= P(E_1)\cdot P(E_2)\cdot P(E_3)\cdot P(E_4)+P(E_1)\cdot P(E_2)\cdot P(E_3)\cdot P\left(\overline{E_4}\right)+P(E_1)\cdot P(E_2)\cdot P\left(\overline{E_3}\right)\cdot P(E_4)$$
$$+P(E_1)\cdot P\left(\overline{E_2}\right)\cdot P(E_3)\cdot P(E_4)$$

$$=\frac{4}{5}\cdot\frac{3}{4}\cdot\frac{5}{6}\cdot\frac{2}{3}+\frac{4}{5}\cdot\frac{3}{4}\cdot\frac{5}{6}\cdot\frac{1}{3}+\frac{4}{5}\cdot\frac{3}{4}\cdot\frac{1}{6}\cdot\frac{2}{3}+\frac{4}{5}\cdot\frac{1}{4}\cdot\frac{5}{6}\cdot\frac{2}{3}$$

$$=\frac{1}{3}+\frac{1}{6}+\frac{1}{15}+\frac{1}{9}$$

$$=\frac{61}{90}$$

Example 10 : A coin is biased so that a head is twice as likely to occur as a tail. If the coin is tossed three times what is the probability of getting exactly two tails?

Let H, T be the event of getting head and tail respectively.
Given that the coin is biased so that a head is twice as likely to occur as a tail.

$$\therefore P(H)=\frac{2}{3} \text{ and } P(T)=\frac{1}{3}$$

The chances of getting exactly two tails $=\{TTH \text{ or } THT \text{ or } HTT\}$, which are all mutually exclusive.

Now required probability $=P(T\cap T\cap H)+P(H\cap T\cap T)+P(T\cap H\cap T)$

$$=P(T)\cdot P(T)\cdot P(H)+P(H)\cdot P(T)\cdot P(T)+P(T)\cdot P(H)\cdot P(T)$$

$$=3\times P(T)\cdot P(T)\cdot P(H)$$

$$=3\times\frac{1}{3}\cdot\frac{1}{3}\cdot\frac{2}{3}$$

$$=\frac{2}{9}$$

Example 10 : Let an urn contains 4 tickets numbered 1 , 2 , 3 , 4 and another urn contains 6 tickets numbered 2 , 4 , 6 , 7 , 8 , 9 . If one of the two urns is chosen at random and a ticket is drawn at random from the chosen urn, find the probabilities that the ticket drawn bears the number (i) 2 or 4 (ii) 3.

Given that one of the two urns is chosen at random and a ticket is drawn at random from the chosen urn.
There are two mutually exclusive possibilities for this. They are
I : First urn is chosen and the ticket is drawn.
II: second urn is chosen and the ticket is drawn.

Let U_1 and U_2 be the event of selecting urn I and II respectively.
Then $P(U_1)=\frac{1}{2}$ and $P(U_2)=\frac{1}{2}$
(i) Let A be the event of selecting the ticket with number 2.
Then $P(A/U_1)$ and P $P(A/U_2)$ are the event of getting the ticket with number 2 after selecting the urn I and II respectively.

Then $P(A/U_1)=\frac{1}{4}$ and $P(A/U_2)=\frac{1}{6}$
$\therefore$ Probability of getting the ticket with number 2 $=P(A)$

$$=P(U_1)\times P(A/U_1)+P(U_2)\times P(A/U_2)$$

$$=\frac{1}{2}\times\frac{1}{4}+\frac{1}{2}\times\frac{1}{6}$$
$$=\frac{5}{24}$$

Similarly, let B be the event of selecting the ticket with number 4.

Now $A\cup B$ is the event of getting the ticket with number 2 or 4

Then $P(A\cup B)=P(A)+P(B)$, { Since A and B are mutually exclusive events }

$$=\frac{5}{24}+\frac{5}{24}$$
$$=\frac{5}{12}$$

(ii) Let C be the event of getting the ticket with number 3

Then $P(C/U_1)$ and P $P(C/U_2)$ are the event of getting the ticket with number 3 after selecting the urn I and II respectively.

Then $P(C/U_1)=\frac{1}{4}$ and $P(C/U_2)=\frac{0}{6}=0$

$\therefore$ Probability of getting the ticket with number 3 $=P(C)$

$$=P(U_1)\times P(C/U_1)+P(U_2)\times P(C/U_2)$$
$$=\frac{1}{2}\times\frac{1}{4}+\frac{1}{2}\times 0$$
$$=\frac{1}{8}$$

Example 10 : A husband and wife appear an interview for two vacancies in the same post. The probability of husband's selection is $\frac{1}{7}$ and the wife is $\frac{1}{5}$. What is the probability that

(i) both of them will be selected.
(ii) only one of them will be selected.
(iii) none of them will be selected.

Let H and W be the event of selecting the husband and wife respectively.

$\therefore$ $P(H)=\frac{1}{7}$ and $P(W)=\frac{1}{5}$

Let $\overline{\mathrm{H}}$ and $\overline{\mathrm{W}}$ be the event of not selecting the husband and wife respectively.

$\therefore$ $P(\overline{H})=\frac{6}{7}$ and $P(\overline{W})=\frac{4}{5}$

(i) Let $H\cap W$ be the event of selecting both husband and wife

$P(H\cap W)=P(H)\times P(W)$ { Since H and W are independent events }

$$= \frac{1}{7} \cdot \frac{1}{5}$$

$$= \frac{1}{35}$$

ii) Let A be the event of selecting only one (either husband or wife)

$A = H \cap \overline{W}$ or $\overline{H} \cap W$ which are mutually exclusive events

$$\therefore \; P(A) = P\left(H \cap \overline{W}\right) + P\left(\overline{H} \cap W\right)$$

$$= P(H) \cdot P\left(\overline{W}\right) + P\left(\overline{H}\right) \cdot P(W) \quad \{\text{Since } H \text{ and } W \text{ are independent events}\}$$

$$= \frac{1}{7} \cdot \frac{4}{5} + \frac{6}{7} \cdot \frac{1}{5}$$

$$= \frac{2}{7}$$

(iii) Let B be the event of selecting none of them. i.e. $B = \overline{H} \cap \overline{W}$

$$\therefore \; P(B) = P\left(\overline{H} \cap \overline{W}\right)$$

$$= P\left(\overline{H}\right) \times P\left(\overline{W}\right)$$

$$= \frac{6}{7} \cdot \frac{4}{5}$$

$$= \frac{24}{35}$$

Example 10 : Probability that a student passes in statistics examination is $\frac{2}{3}$ and the probability that the student will not pass in mathematics examination is $\frac{5}{9}$. The probability that a student will pass in at least one of the examination is $\frac{4}{5}$. Find the probability that the student will pass in both examinations.

Let S and M be the event of passing in statistics and mathematics exam respectively and $\overline{S}$ and $\overline{M}$ are the event of not passing in statistics and mathematics exam respectively.

Given that $P(S) = \frac{2}{3}$, $P\left(\overline{\mathrm{M}}\right) = \frac{5}{9}$

Therefore $P(\mathrm{M}) = 1 - \mathrm{P}\left(\overline{\mathrm{M}}\right) = 1 - \frac{5}{9} = \frac{4}{9}$

$S \cup M$ = event of passing in at least one exam and also given that $P(S \cup \mathrm{M}) = \frac{4}{5}$

$S \cap M$ = event of passing in both exams

Probability of passing in both exams $= P(S \cap M)$

$$= P(S) + P(M) - P(S \cup M) \quad \{\text{By addition theorem}\}$$

$$= \frac{4}{9} + \frac{2}{3} - \frac{4}{5}$$

$$= \frac{14}{45}$$

BAYE'S THEORM: Let $E_1, E_2, \ldots\ldots, E_n$ be n mutually disjoint events with $P(E_i) \neq 0$ for all i, and for any event A which is a subset of $\bigcup_{i=1}^{n} E_i$ such that $P(A) > 0$, then $P(E_i / A) = \dfrac{P(E_i) \cdot P(A/E_i)}{\sum_{i=1}^{n} P(E_i) \cdot P(A/E_i)}$. .

Since A is a subset of $\bigcup_{i=1}^{n} E_i$, we have

$$A = A \cap \left[\bigcup_{i=1}^{n} E_i \right]$$

$$= \bigcup_{i=1}^{n} (A \cap E_i)$$

$$P(A) = P\left[\bigcup_{i=1}^{n} (A \cap E_i) \right]$$

$$= P\left[\bigcup_{i=1}^{n} (A \cap E_i) \right]$$

Since, $(A \cap E_i)$ is a subset of n mutually disjoint events, by addition theorem of probability, we have

$$P(A) = \sum_{i=1}^{n} P(A \cap E_i)$$

Then by compound theorem of probability

$$P(A) = \sum_{i=1}^{n} P(E_i) \cdot P(A/E_i)$$

Also, $P(A \cap E_i) = P(A) \cdot P(E_i / A)$

$$P(E_i / A) = \frac{P(A \cap E_i)}{P(A)}$$

$$P(E_i / A) = \frac{P(E_i) \cdot P(A/E_i)}{\sum_{i=1}^{n} P(E_i) \cdot P(A/E_i)}$$

Example 1: A company producing electric relays has three manufacturing plants producing 50, 30 and 20 percent respectively of its product. Suppose that the probabilities that a relay manufactured by these plants is defective are 0.02, 0.05 and 0.01 respectively.

If a relay selected at random is found to be defective, what is the probability that it was manufactured by plant 2?.

Let E_1, E_2 and E_3 be the event of selecting the electric relay which was manufactured by Plant1, Plant2 and Plant3 respectively.

Given that $P(E_1) = 50\% = 0.5$, $P(E_2) = 30\% = 0.3$, $P(E_3) = 20\% = 0.2$

Let A be the event of selecting defective electric relay. Also given that

$P(A/E_1) = 0.02$, $P(A/E_2) = 0.05$, $P(A/E_3) = 0.01$

Required probability = Probability of selecting defective relay which was manufactured by Plant2.

$$P(E_2/A) = \frac{P(E_2)\cdot P(A/E_2)}{P(E_1)\cdot P(A/E_1) + P(E_2)\cdot P(A/E_2) + P(E_3)\cdot P(A/E_3)}$$

$$= \frac{(0.35)(0.04)}{(0.25)(0.05) + (0.35)(0.04) + (0.04)(0.02)}$$

$$= 0.5128$$

Example 2: Suppose that coloured balls are distributed in three indistinguishable boxes as follows:

	Box 1	Box 2	Box 3
Red	2	4	3
White	3	1	4
Blue	5	3	5

A box is selected at random from which a ball is selected at random and it is observed to be red. What is the probability that the box 3 was selected?

Let E_1, E_2 and E_3 be the events of selecting the boxes A, B and C respectively.

$\therefore P(E_1) = \frac{1}{3}$, $P(E_2) = \frac{1}{3}$, $P(E_3) = \frac{1}{3}$

Now let A be the event of selecting the red ball.

Then $P(A/E_1) = \frac{2}{10} = \frac{1}{5}$, $P(A/E_2) = \frac{4}{8} = \frac{1}{2}$, $P(A/E_3) = \frac{3}{12} = \frac{1}{4}$

Required probability = Probability of selecting the red ball which came from box 3

$$P(E_3/A) = \frac{P(E_3)\cdot P(A/E_3)}{P(E_1)\cdot P(A/E_1) + P(E_2)\cdot P(A/E_2) + P(E_3)\cdot P(A/E_3)}$$

$$= \frac{\left(\frac{1}{3}\right)\cdot\left(\frac{1}{4}\right)}{\left(\frac{1}{3}\right)\cdot\left(\frac{1}{5}\right) + \left(\frac{1}{3}\right)\cdot\left(\frac{1}{2}\right) + \left(\frac{1}{3}\right)\cdot\left(\frac{1}{4}\right)} = \frac{5}{19}$$

Example 3: Three machines A, B and C are producing 20,000, 40,000 and 60,000 bolts per shift. They are known to produce 4% , 3% and 2% defective bolts respectively. If a bolt is chosen at random from bolts produced in a shift and was found to be defective . What is the probability that it was produced by B ?

Let E_1, E_2 and E_3 be the event of selecting the bolts which was manufactured by A, B and C respectively.

$$\therefore\ P(E_1)=\frac{20.000}{1,20,000}=\frac{1}{6},\qquad P(E_2)=\frac{40.000}{1,20,000}=\frac{1}{3},\qquad P(E_3)=\frac{60.000}{1,20,000}=\frac{1}{2}$$

Let A be the event of selecting the defective bolt.

$$P(A/E_1)=4\%=\frac{4}{100},\quad P(A/E_2)=3\%=\frac{3}{100},\quad P(A/E_3)=2\%=\frac{2}{100}$$

Required probability = Probability of getting defective bolt which was produced by B .

$$P(E_2/A)=\frac{P(E_2)\cdot P(A/E_2)}{P(E_1)\cdot P(A/E_1)+P(E_2)\cdot P(A/E_2)+P(E_3)\cdot P(A/E_3)}$$

$$=\frac{\frac{1}{3}\times\frac{3}{100}}{\frac{1}{6}\times\frac{4}{100}+\frac{1}{3}\times\frac{3}{100}+\frac{1}{2}\times\frac{2}{100}}$$

$$=\frac{3}{8}$$

Example 4: A has a scooter and a car. About three fourths of the time he uses the scooter and uses the car otherwise. When he uses the scooter he comes to the office on time about 75% of the time . If he uses the car he gets to his office on time about 60% of the time. On a given, day he was late to the office. What is the probability that he came to the office by scooter.

Let E_1, E_2 be the events of using the scooter and car to come to the office respectively.

Given that $P(E_1)=\frac{3}{4},\ P(E_2)=\frac{1}{4}.$

Let A be the event of coming to the office by late.

When he uses the scooter he comes to the office on time about 75% of the time. If he uses the car he gets to his office on time about 60% of the time.

Hence, when he uses the scooter he comes to the office by late about 25% of the time. If he uses the car he gets to his office by late about 40% of the time.

$$\therefore\ P(A/E_1)=25\%=\frac{25}{100}=\frac{1}{4},\quad P(A/E_2)=40\%=\frac{40}{100}=\frac{2}{5}$$

Required probability = Probability of coming to the office by late when he came to the office by scooter.

$$P(E_1/A) = \frac{P(E_1)\cdot P(A/E_1)}{P(E_1)\cdot P(A/E_1) + P(E_2)\cdot P(A/E_2)}$$

$$= \frac{\frac{3}{4}\left(\frac{1}{4}\right)}{\frac{3}{4}\left(\frac{1}{4}\right) + \frac{1}{4}\left(\frac{2}{5}\right)}$$

$$= \frac{15}{23}$$

Random Variable

Let S be the sample space of an experiment. A random variable X is a real valued function defined on S. i.e. for each $s \in S$ there is a real number $X(s) = p$.

Example: Suppose a coin is tossed twice. The sample space is $S = \{HH, TT, TH, HT\}$. Let X denote the 'number of heads' appeared. Then X is a random variable with values $X(HH) = 2$, $X(TH) = X(HT) = 1$, $X(TT) = 0$. Therefore the values of X are 0, 1, 2.

Note: A random variable X is said to be 'Discrete' if it takes a finite number of values or countably infinite number of values. The above example is discrete.

Probability Mass Function or Probability Function: Let X be a discrete random variable which takes the values $x_1, x_2, x_3, ...x_n$. Let $P[X = x_1] = p_1$ be the probability of x_1. Then the function p is called the probability mass function if $p(x_i) \geq 0$ *for all i and* $\sum_{i=1}^{n} p(x_i) = 1$.

Note: The probability distribution (i.e the values of X and its probability) is usually displayed in the form of a table.

X	x_1	x_2	x_3	.....	x_n
$P(X = x)$	$P(X = x_1)$ $p(x_1)$ p_1	$P(X = x_2)$ $p(x_2)$ p_2	$P(X = x_2)$ $p(x_3)$ p_3	.....	$P(X = x_n)$ $p(x_n)$ p_n

Note: $P[X \leq x_3] = x_1 + x_2 + x_3$, $\quad P[X < x_3] = x_1 + x_2.$

$P[X \geq x_3] = x_3 + x_4 + x_5 + ... + x_n = 1 - P[X < x_3]$ $\qquad P[X > x_3] = x_4 + x_5 + ... + x_n = 1 - P[X \leq x_3]$

Necessary formulae:

Mean = First Moment about origin $\mu_1^{'} = E[X] = \sum_i x_i \, p(x_i)$

NOTE :

1. Expectation $E(X)$ is a theoretical concept and represents the average value of the random variable.
2. Expectation value need not be one of the possible outcomes of the random variable.

3. Expectation of the random variable can be finite or infinite. If the random variable X takes finite number of values , then $E(X)$ is finite. If $E(X)$ is infinite , then the random variable X takes infinite number of values.

Properties of Expectation :

1. $E(cX) = cE(X)$ where c is a constant.
2. $E(c) = c$ where c is a constant.
3. $E(aX+b) = aE(X)+b$ where a and b are constants.

4. $E\left[f(x)+g(x)\right] = E\left[f(x)\right]+E\left[g(x)\right]$ where f and g are two functions such that $E\left[f(x)\right]$ and $E\left[g(x)\right]$ are finite.

Second Moment about origin $\mu_2' = E\left[X^2\right] = \sum_i x_i^2\, p(x_i)$

Variance of X = $E\left[X^2\right] - \left[E(X)\right]^2$ and $V(aX+b) = a^2V(X)$

Moment Generating Function $M_X(t) = E\left[e^{tx}\right] = \sum e^{tx} p(x)$

$\mu_1' = E(X) = M_X'(0)$, $\mu_2' = E(X^2) = M_X''(0)$

$$\mu_1' = E(X) = coeficient\ of\ \frac{t^1}{1!}\ in\ the\ expansion\ of\ M(t)$$

$$\mu_2' = E(X^2) = coeficient\ of\ \frac{t^2}{2!}\ in\ the\ expansion\ of\ M(t)$$

Cumulative Distribution Function $F(x_i) = P(X \le x_i)$ and $P(X = x_i) = F(x_i) - F(x_{i-1})$ with $F(-\infty) = 0,\ F(\infty) = 1$

Example 1: Let X be the random variable which denotes the number of heads in three tosses of a fair coin. Determine the probability mass function of X .

Sample space when tossing coins three times is
$\{HHH, THH, HTH, HT, HHT, TTH, HTT, TTT\}$

Let X denotes the random, variable of getting heads

X	0	1	2	3
$P(X=x)$	1/8	3/8	3/8	1/8

Example 2: Let the random variable X denotes the sum obtained 'm' when rolling a pair of fair dice. Determine the probability mass function of X .

Let the random variable X represents the sum of numbers on them when two dice are thrown.

Possible (x, y)	Sum $X = x + y$	$P(X = x)$
(1,1)	2	1/36
(1,2),(2,1)	3	2/36
(1,3),(2,2),(3,1)	4	3/36
(1,4),(2,3), (3,2),(4,1)	5	4/36
(1,5),(2,4)(3,3),(4,2) (5,1)	6	5/36
(1,6),(2,5)(3,4),(4,3)(5,2),(6,1)	7	6/36
(2,6),(3,5)(4,4),(5,3) (6,2)	8	5/36
(3,6),(4,5)(5,4),(6,3)	9	4/36
(4,6),(5,5) (6,4)	10	3/36
(5,6),(6,5)	11	2/36
(6,6)	12	1/36

Example 3: Evaluate the Mean of a random variable X if its probability distribution is as follows:

X	**−2**	**−1**	**0**	**1**	**2**
$P(X = x)$	a	a	$2a$	a	a

Mean $E(X) = \sum x\, p(x) = -2a - a + 0(2a) + a + 2a = 0$

Example 4: If a random variable X takes the values 1, 2, 3, 4, such that $2P(X=1) = 3P(X=2) = P(X=3) = 5P(X=4)$. Find the probability distribution of X.

Let $P(X=3) = k$.

Then from the given data, we get $P(X=1) = \frac{k}{2}$, $P(X=2) = \frac{k}{3}$, $P(X=4) = \frac{k}{5}$.

We Know that $P(X=1) + P(X=2) + P(X=3) + P(X=4) = 1$

$$\frac{k}{2} + \frac{k}{3} + k + \frac{k}{5} = 1$$

$$\frac{61k}{30} = 1$$

$$k = \frac{30}{61}$$

Required probability distribution is

X = x	1	2	3	4
P(X = x)	$\frac{15}{61}$	$\frac{10}{61}$	$\frac{30}{61}$	$\frac{6}{61}$

Example 5: If the range of X is the set {0,1,2,3,4} and $P(X=x)=0.2$, determine the mean and variance of the random variable.

We tabulate the values of X and its probabilities.

X	0	1	2	3	4
$P(x)$	0.2	0.2	0.2	0.2	0.2
$x\cdot P(x)$	0	0.2	0.4	0.6	0.8
$x^2\cdot P(x)$	0	0.2	0.8	1.8	3.2

We know that Mean $E(X)=\sum_{0}^{4} x\,P(\mathrm{x}) = 0+0.2+0.4+0.6+0.8=2$

$$E\left(X^2\right)=\sum_{0}^{4} x^2 P(x) = 0+0.2+0.8+1.8+3.2=6$$

$$Var(X)=E\left(X^2\right)-\left[E(X)\right]^2 = 6-(2)^2=2$$

Example 6: A random variable X has the following probability function:

$X=x$	0	1	2	3	4	5	6	7
$P(x)$	0	k	$2k$	$2k$	$3k$	k^2	$2k^2$	$7k^2+k$

i. Find k **ii. Evaluate $P(X<6)$, $P(X\ge 6)$, $P(0<X<5)$**

iii. Find the distribution function of X

iv. Find the least value of 'a' such that $P(X\le a)>0.5$.

i. We know that $\sum P(x)=1$

$$10k^2+9k=1$$

$$10k^2+9k-1=0$$

$$k=-1 \text{ or } k=\frac{1}{10}$$

But $P(x)$ cannot be negative. Hence $k=-1$ is neglected. Hence $k=\frac{1}{10}$

ii. $P(X<6)=P(0)+P(1)+.......+P(5)=8k+k^2=\frac{8}{10}+\frac{1}{100}=\frac{81}{100}$

$$P(X\ge 6)=P(6)+P(7)=9k^2+k=\frac{9}{100}+\frac{1}{10}=\frac{19}{100}$$

or $P(X\ge 6)=1-P(X<6)=1-\frac{81}{100}=\frac{19}{100}$

$$P(0 < X < 5) = P(1) + ... + P(4) = 8k = \frac{8}{10}$$

iii.
$$F(0) = P(X \le 0) = P(0) = 0$$
$$F(1) = P(X \le 1) = P(0) + P(1) = k = \frac{1}{10}$$
$$F(2) = P(X \le 2) = P(0) + P(1) + P(2) = 3k = \frac{3}{10}$$
$$F(3) = P(X \le 3) = P(0) + ... + P(3) = 5k = \frac{5}{10}$$
$$F(4) = P(X \le 4) = P(0) + ... + P(4) = 8k = \frac{8}{10}$$
$$F(5) = P(X \le 5) = P(0) + ... + P(5) = 8k + k^2 = \frac{8}{10} + \frac{1}{100} = \frac{81}{100}$$
$$F(6) = P(X \le 6) = P(0) + ... + P(6) = 8k + 3k^2 = \frac{8}{10} + \frac{3}{100} = \frac{83}{100}$$
$$F(7) = P(X \le 7) = P(0) + ... + P(7) = 9k + 10k^2 = \frac{9}{10} + \frac{10}{100} = \frac{100}{100} = 1$$

iv.
$$P(X \le 2) = P(0) + P(1) + P(2) = 3k = \frac{3}{10}$$
$$P(X \le 3) = P(0) + ... + P(3) = 5k = \frac{5}{10} = \frac{1}{2}$$
$$P(X \le 4) = P(0) + ... + P(4) = 8k = \frac{8}{10} > \frac{1}{2} \quad and \quad hence\ a = 4$$

Example 7: Find the Moment Generating Function, Mean, Variance of the distribution $P(X = x) = \frac{1}{2^x}, \ x = 1, 2, 3, ...$ **Also find** $P(X = even)$, $P(X = odd)$, $P(X = divisible\ by\ 3)$**.**

$$M_X(t) = E\left[e^{tx}\right] = \sum_1^{\infty} e^{tx} \quad p(x)$$
$$= e^t . \frac{1}{2^1} + e^{2t} . \frac{1}{2^2} + e^{3t} . \frac{1}{2^3} + ...$$
$$= \left(\frac{e^t}{2}\right) + \left(\frac{e^t}{2}\right)^2 + \left(\frac{e^t}{2}\right)^3 + ...$$
$$= \left(\frac{e^t}{2}\right) \left[1 + \left(\frac{e^t}{2}\right) + \left(\frac{e^t}{2}\right)^2 + \left(\frac{e^t}{2}\right)^3 + ...\right]$$
$$= \left(\frac{e^t}{2}\right) \left[1 - \frac{e^t}{2}\right]^{-1}$$
$$= \left(\frac{e^t}{2}\right) \left[\frac{2 - e^t}{2}\right]^{-1}$$
$$= \left(\frac{e^t}{2}\right) \left[\frac{2}{2 - e^t}\right]$$

Mean

$$E[X] = \sum_{1}^{\infty} x \quad p(x)$$

$$= 1.\frac{1}{2^1} + 2.\frac{1}{2^2} + 3.\frac{1}{2^3} + \dots\dots\dots$$

$$= \frac{1}{2}\left[1. + 2.\left(\frac{1}{2}\right) + 3.\left(\frac{1}{2}\right)^2 + \dots\dots\dots\right]$$

$$= \frac{1}{2}\left[1 - \frac{1}{2}\right]^{-2}$$

$$= 2$$

$$E\left[X^2\right] = \sum_{1}^{\infty} x^2 \quad p(x)$$

$$= 1^2.\frac{1}{2^1} + 2^2.\frac{1}{2^2} + 3^2.\frac{1}{2^3} + \dots\dots\dots$$

$$= \frac{1}{2}\left[1. + 4.\left(\frac{1}{2}\right) + 9.\left(\frac{1}{2}\right)^2 + \dots\dots\dots\right]$$

$$= \frac{1}{2}\left[1 + 4x + 9x^2 + \dots\dots\dots.\right]$$

$$= \frac{1}{2}(1+x)[\ 1-x\ ]^{-1}$$

$$= \frac{1}{2}\left(1+\frac{1}{2}\right)\left[1 - \frac{1}{2}\right]^{-3}$$

$$= 6$$

$$Var(X) = E\left[X^2\right] - \left[E(X)\right]^2$$

$$= 6-4$$

$$= 2$$

$$P[X = even] = P[X = 2] + P[X = 4] + P[X = 6] + \dots\dots\dots.$$

$$= \frac{1}{2^2} + \frac{1}{2^4} + \frac{1}{2^6} + \dots\dots\dots$$

$$= \frac{1}{2^2}\left[1 + \left(\frac{1}{2^2}\right) + \left(\frac{1}{2^2}\right)^2 + \left(\frac{1}{2^2}\right)^3 + \dots\dots\dots\right]$$

$$= \frac{1}{2^2}\left[1 - \frac{1}{2^2}\right]^{-1}$$

$$= \frac{1}{3}$$

$$P[X = odd] = 1 - P[X = even] = 1 - \frac{1}{3} = \frac{2}{3}$$

$$P[X = dicisible \ \ by \ \ 3] = P[X = 3] + \ \ P[X = 6] + P[X = 9] + \ \ ....$$

$$= \frac{1}{2^3} + \frac{1}{2^6} + \frac{1}{2^9} + ...$$

$$= \frac{1}{2^3} + \frac{1}{\left(2^3\right)^2} + \frac{1}{\left(2^3\right)^3} + ...$$

$$= \frac{1}{2^3}\left[1 + \frac{1}{\left(2^3\right)} + \frac{1}{\left(2^3\right)^2} + ...\right]$$

$$= \frac{1}{2^3}\left[1 - \frac{1}{2^3}\right]^{-1}$$

$$= \frac{1}{7}$$

Example 8: Suppose a player plays the following game. A fair die is tossed. If 1 or 2 occurs, he losses Rs. 30; if 3 or 4 or 5 occurs, he gains Rs. 50; If 6 occurs, he gains Rs. 90. If there is an entry charge, what is the amount he will be willing to bet if the game is to be to his advantage in the long run?

In the experiment of tossing the fair die, $P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = \frac{1}{6}$.

Let X represents the amount gains by the player i.e. $X = -30, 50, 90$.

If 1 or 2 occurs, the player losses Rs.30. Therefore	If 3 or 4 or 5 occurs, the player gains Rs. 50. Therefore	If 6 occurs , he gains Rs. 90. Therefore
$P(X = -30) = P(1) + P(2)$ $= \frac{1}{6} + \frac{1}{6}$ $= \frac{1}{3}$	$P(X = 30) = P(1) + P(2) + P(5)$ $= \frac{1}{6} + \frac{1}{6} + \frac{1}{6}$ $= \frac{1}{2}$	$P(X = 90) = P(6)$ $= \frac{1}{6}$

Hence the probability distribution is

X	- 30	50	90
$P(X = x)$	$\frac{2}{6}$	$\frac{3}{6}$	$\frac{1}{6}$

$$\text{Expected gain of the player } = E(X)$$

$$= \sum_x xP(X = x)$$

$$= (\text{-}30)\left(\frac{2}{6}\right) + 50\left(\frac{3}{6}\right) + 90\left(\frac{1}{6}\right)$$

$$= 30$$

He should be willing to bet at most Rs. 30 per play.

Example 8: A box contains 5 red balls and 5 green balls. Balls are drawn one by one without replacement until a green ball is drawn. Let X denote the number of the draw in which the first green ball is drawn. Find the probability distribution of X and its expectation.

Given that the box contains 5 red balls and 5 green balls.
Now the balls are drawn one by one without replacement until a green ball is drawn.
It may be done in any of the following way.
(1) The green ball is drawn in first time.
(2) The red ball is drawn in first time and then the green ball is drawn.
(3) The red balls are drawn in first two times and then the green ball is drawn.
(4) The red balls are drawn in first three times and then the green ball is drawn.
(5) The red balls are drawn in first four times and then the green ball is drawn.
(6) The red balls are drawn in first five times and then the green ball is drawn.

Let X denote the number of the draw in which the first green ball is drawn.
Hence $X = 1, 2, 3, 4, 5, 6$.

Now $P(X=1)=\frac{5C_1}{10C_1}=\frac{5}{10}=\frac{1}{2}$ $\qquad P(X=3)=\frac{5C_1}{10C_1}\times\frac{4C_1}{9C_1}\times\frac{5C_1}{8C_1}=\frac{5}{10}\times\frac{4}{9}\times\frac{5}{5}=\frac{5}{36}$

$P(X=2)=\frac{5C_1}{10C_1}\times\frac{5C_1}{9C_1}=\frac{5}{10}\times\frac{5}{9}=\frac{5}{18}$ $\qquad P(X=4)=\frac{5C_1}{10C_1}\times\frac{4C_1}{9C_1}\times\frac{3C_1}{8C_1}\times\frac{5C_1}{7C_1}=\frac{5}{10}\times\frac{4}{9}\times\frac{3}{8}\times\frac{5}{7}=\frac{5}{84}$

Similarly, $P(X=5)=\frac{5}{252}$, $\quad P(X=6)=\frac{1}{252}$

Hence the probability distribution of X is

X	1	2	3	4	5	6
$P(X=x)$	$\frac{1}{2}$	$\frac{5}{18}$	$\frac{5}{36}$	$\frac{5}{84}$	$\frac{5}{252}$	$\frac{1}{252}$

Expected number of drawn $= E(X)$

$$= \sum_x xP(X=x)$$

$$=1\times\frac{1}{2}+2\times\frac{5}{18}+3\times\frac{5}{36}+4\times\frac{5}{84}+5\times\frac{5}{252}+6\times\frac{1}{252}$$

$$=\frac{2610}{1260}$$

□ 2

Continuous Random Variable

A random variable X is said to be 'continuous' if it takes all values in an interval.

Probability Density Function or Probability Function: Let X be a random variable which takes all values in an interval $(a \le X \le b)$. Then the function $f(x)$ is called the probability density function of X if $f(x) \ge 0$ *for all x and* $\int_{-\infty}^{\infty} f(x)dx = 1$.

Necessary Formulae:

Mean = First Moment about origin = $\mu_1^{'} = E[X] = \int_{-\infty}^{\infty} x\, f(x)dx$

Second Moment about origin = $\mu_2^{'} = E[X^2] = \int_{-\infty}^{\infty} x^2\, f(x)dx$

Variance of X = $E[X^2] - [E(X)]^2$ and $V(aX+b) = a^2V(X)$

Moment Generating Function $M_X(t) = E[e^{tx}] = \int_{-\infty}^{\infty} e^{tx}\, f(x)\, dx$

$\mu_1^{'} = E(X) = M_X^{'}(0)$, $\mu_2^{'} = E(X^2) = M_X^{''}(0)$

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

Cumulative Distribution Function $F(x) = P(X \le x) = \int_{-\infty}^{x} f(x)dx$ with $F(-\infty)=0$, $F(\infty)=1$

Note: CDF must be evaluated in each interval from $-\infty$ to ∞. Also $\frac{d}{dx}F(x) = f(x)$

Properties of CDF $F(x)$.

If $F(x)$ is a distribution function of a random variable X and if $a<b$, then, $P(a < X \le b) = F(b) - F(a)$.

If $F(x)$ is the distribution function of a random variable X, then $0 \le F(x) \le 1$ and $F(x) \le F(x)$ if $x < y$.

Example 1: A continuous random variable X has the probability density function $f(x) = k(1+x),\quad 2 \le x \le 5$. Find $P(X<4)$.

Since $f(x)$ is a pdf,

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

$$\int_2^5 k(1+x)dx = 1$$

$$k\left[\frac{(1+x)^2}{2}\right]_2^5 = 1$$

$$\frac{k}{2}[36-9] = 1$$

$$k = \frac{2}{27}$$

$$P(X<4) = P(2<X<4)$$

$$= \int_2^4 f(x)dx$$

$$= \int_2^4 k(1+x)dx$$

$$= k\left[\frac{(1+x)^2}{2}\right]_2^4$$

$$= \frac{k}{2}[25-9]$$

$$= \frac{16}{27}$$

Example 2: If a random variable X has the probability density function $f(x) = 3x^2, 0 < x < 1$
$= 0, \quad otherwise$ **, find 'k' such that $P(X > k) = 0.05$.**

Given that $P(X > k) = 0.05$

$$P(k < X < 1) = 0.05$$

$$\int_k^1 f(x)dx = 0.05$$

$$\int_k^1 3x^2 \ dx = 0.05$$

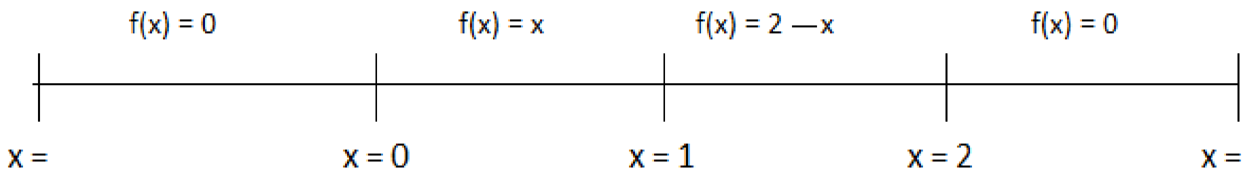
$$3\left[\frac{x^3}{3}\right]_k^1 = 0.05$$

$$\left[1 - k^3\right] = \frac{5}{100}$$

$$k^3 = 1 - \frac{5}{100} = \frac{95}{100} \qquad i.e. \ \ k = \sqrt[3]{\frac{95}{100}} = 0.983$$

Example 3: Find the cumulative distribution function of the random variable with probability density function $f(x) = x \ \ if \ \ 0 < x < 1$
$= 2 - x \ \ if \ \ 1 < x < 2$
$= 0 \ \ if \ \ 2 < x < \infty$

Also find the value of $P(0.5 < X < 1.5)$, $P(1 < X < 2)$, $P(2 < X < 3)$.



CDF in the interval $(-\infty, 0)$

$$F(x) = P(X < x) = P(-\infty < X < x) = \int_{-\infty}^{x} f(x)dx = \int_{-\infty}^{x} 0 \ dx = 0$$

CDF in the interval $(0, 1)$

$$F(x) = P(X < x) = P(-\infty < X < x) = \int_{-\infty}^{x} f(x)dx = \int_{-\infty}^{0} 0 \ dx + \int_0^x x \ dx = \left[\frac{x^2}{2}\right]_0^x = \frac{x^2}{2}$$

CDF in the interval $(1, 2)$

$$F(x)=P(-\infty<X<x)=\int_{-\infty}^{x} f(x)dx=\int_{-\infty}^{0} 0\ dx+\int_{0}^{1} x\ dx+\int_{1}^{x} 2-x\ dx$$

$$=\left[\frac{x^2}{2}\right]_0^1+\left[\frac{(2-x)^2}{-2}\right]_1^x=\frac{1}{2}+\frac{(2-x)^2}{-2}+\frac{1}{2}=1-\frac{(2-x)^2}{2}$$

CDF in the interval $(2, \infty)$

$$F(x)=P(-\infty<X<x)=\int_{-\infty}^{x} f(x)dx=\int_{-\infty}^{0} 0\ dx+\int_{0}^{1} x\ dx+\int_{1}^{2} 2-x\ dx+\int_{2}^{\infty} 0\ dx$$

$$=\left[\frac{x^2}{2}\right]_0^1+\left[\frac{(2-x)^2}{-2}\right]_1^2=\frac{1}{2}+\frac{1}{2}=1$$

Therefore $F(x)=0$, in $-\infty \le x<0$

$F(x)=\frac{x^2}{2}$, in $0 \le x<1$

$F(x)=1-\frac{(2-x)^2}{2}$, in $1 \le x<2$

$F(x)=1$, in $2 \le x<\infty$

$P(0.5<X<1.5)=F(1.5)-F(0.5)$ $=1-\frac{(0.5)^2}{2}-\frac{(0.5)^2}{2}$ $=0.75$	$P(1<X<2)=F(2)-F(1)$ $=1-\left(1-\frac{1}{2}\right)$ $=\frac{1}{2}$	$P(2<X<3)=F(3)-F(2)$ $=1-1$ $=0$

Example 4: The mileage X (in thousand of miles) which car owners get with a certain kind of tyre is a random variable having a probability density function $f(x)=\frac{1}{20}e^{-\frac{x}{20}}, x>0.$ **Find the probabilities that one of these tyres will last (i) at most 10,000 miles (ii) at least 30,000 miles (iii) anywhere from 16,000 to 24,000 miles.**

Given $f(x)=\frac{1}{20}e^{-\frac{x}{20}},\ 0<x\ \infty$

(i) Probability that the tyre will last at most 10,000 miles

$$P[X\le 10]=P[0\le X\le 10]=\int_0^{10}\frac{1}{20}e^{-\frac{1}{20}x}dx=\frac{1}{20}\left[\frac{e^{-\frac{1}{20}x}}{-\frac{1}{20}}\right]_0^{10}=-\left[e^{-\frac{1}{2}}-1\right]=1-e^{-\frac{1}{2}}$$

(ii) Probability that the tyre will last at least 30,000 miles

$$P[X \geq 30] = P[30 \leq X \leq \infty] = \int_{30}^{\infty} \frac{1}{20} e^{-\frac{1}{20}x} dx = \frac{1}{20}\left[\frac{e^{-\frac{1}{20}x}}{-\frac{1}{20}}\right]_{30}^{\infty} = -\left[0 - e^{-\frac{3}{2}}\right] = e^{-\frac{3}{2}}$$

(iii) Probability that the tyre will last between 16,000 to 24,000 miles

$$P[16 \leq X \leq 24] = \int_{16}^{24} \frac{1}{20} e^{-\frac{1}{20}x} dx = \frac{1}{20}\left[\frac{e^{-\frac{1}{20}x}}{-\frac{1}{20}}\right]_{16}^{24} = -\left[e^{-\frac{24}{20}} - e^{-\frac{16}{20}}\right] = e^{-\frac{16}{20}} - e^{-\frac{24}{20}}$$

Example 5: Find the M.G.F of the R.V X having the probability density function $f(x) = \frac{x}{4} e^{-\frac{x}{2}}, x > 0$. Also deduce the first four moments about the origin.

$= 0$, *otherwise*

The moment generating function of given $f(x)$ is given by

$$\begin{aligned}
M_X(t) &= E\left[e^{tx}\right] \\
&= \int_{-\infty}^{\infty} e^{tx} f(x)\, dx \\
&= \frac{1}{4}\int_0^{\infty} xe^{tx} e^{-\frac{1}{2}x} dx \\
&= \frac{1}{4}\int_0^{\infty} xe^{-\left(\frac{1}{2}-t\right)x} dx \\
&= \frac{1}{4}\left[(x)\left(\frac{e^{-\left(\frac{1}{2}-t\right)x}}{-\left(\frac{1}{2}-t\right)}\right) - (1)\left(\frac{e^{-\left(\frac{1}{2}-t\right)x}}{\left(\frac{1}{2}-t\right)^2}\right)\right]_0^{\infty} \\
&= \frac{1}{4}\left[-\left(\frac{2}{1-2t}\right)(x)e^{-\left(\frac{1}{2}-t\right)x} - \left(\frac{2}{1-2t}\right)^2 e^{-\left(\frac{1}{2}-t\right)x}\right]_0^{\infty} \\
&= \frac{1}{4}\left[0 + \left(\frac{2}{1-2t}\right)^2\right] \\
&= (1-2t)^{-2} \\
&= \left[1 + 2(2t) + 3(2t)^2 + 4(2t)^3 + 5(2t)^4 + \dots\dots\dots\right]
\end{aligned}$$

To find the first 4 moments about origin:

$$E[X] = coefficient\ of\ \frac{t^1}{1!}\ in\ the\ expansion\ of\ M_X(t) = 4$$

$$E\left[X^2\right] = coefficient\ of\ \frac{t^2}{2!}\ in\ the\ expansion\ of\ M_X(t) = 3$$

$$E\left[X^3\right] = coefficient\ of\ \frac{t^3}{3!}\ in\ the\ expansion\ of\ M_X(t) = \frac{4}{3}$$

$$E\left[X^4\right] = coefficient\ of\ \frac{t^4}{4!}\ in\ the\ expansion\ of\ M_X(t) = \frac{5}{12}$$

Example 6: Find the probability density function of the random variable X if its cumulative distribution function is $F(x) = 1-(1+x)e^{-x}, \quad x \geq 0.$

The pdf is $f(x) = \frac{d}{dx}F(x)$

$$= \frac{d}{dx}\left[1 - e^{-x} - x\ e^{-x}\right]$$

$$= e^{-x} - \left(-x\ e^{-x} + e^{-x}\right)$$

$$= x\ e^{-x}$$

Example 7: Evaluate $E\left(X^2\right)$ if the probability density function of a random variable is $f(x) = xe^{-x}, x \geq 0$.

$$E\left[X^2\right] = \int_0^{\infty} x^2 f(x)\, dx = \int_0^{\infty} x^3 e^{-x}\, dx = \Gamma 4 = 3! = 6$$

Example 8: Find the constant K if $f(x) = Kx^2, 0 < x < 1$ is the probability density function of a continuous random variable X.

We know that

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

$$\int_0^1 K\,x^2\,dx = 1$$

$$K\left[\frac{x^3}{3}\right]_0^1 = 1$$

$$K\left[\frac{1}{3}\right] = 1$$

$$\therefore K = 3$$

Example 9: Evaluate $P(0 \le x \le 1)$ if the cumulative distribution function is $F(x) = 1-(1+x)e^{-x}$, $x \ge 0$.

Since we are given cumulative distribution function, $P(0 \le x \le 1) = F(1) - F(0) = 1-2e^{-1}$

Binomial Distribution

A random variable X is said to follow binomial distribution if its probability mass function is $P[X = x] = nc_x\, p^x q^{n-x}$, $x = 0,1,2,\ldots n$ $\quad where \quad p+q=1$. It is denoted as $X \approx B(n,p)$ i.e., n, p are the parameters.

- It gives probability of x success in n trials.
- If the trial is repeated for N times, then the required probability is $N \cdot P(x)$.
- If $X \approx B(n_1, p)$ and $Y \approx B(n_2, p)$ then $X+Y \approx B(n_1+n_2, p)$.

Example 1. A random variable X follows Binomial distribution with mean 2, variance 4. Give your comment on this.

Since X follows Binomial distribution, mean $= np$ and variance $= npq$.

Given $np = 2$ and $npq = 4$

Therefore $\dfrac{npq}{np} = \dfrac{4}{2}$ $\textit{this implies}$ $q = 2 > 1$. This is not possible. Hence given data are wrong.

Example 2. For a Binomial distribution of mean 4 and variance 2, find the probability of getting i. at least 2 successes ii. utmost 2 successes iii. $P(5 < X < 8)$

Since X follows Binomial distribution, mean $= np$ and variance $= npq$.

Given $np = 4$ and $npq = 2$

Therefore $\dfrac{npq}{np} = \dfrac{2}{4} = 0.5$ $\textit{this implies}$ $q = 0.5$ $\textit{and hence}$ $p = 1-q = 0.5$

But $np = 2$ $\textit{gives}$ $n(0.5) = 2$ $\textit{i.e.}$ $n = 4$

The $p.m.f$ of binomial distribution is $P[X = x] = nC_x p^x q^{n-x}$, $x = 0, 1, 2, 3, 4$

$$P[X = x] = 4\ c_x\ (0.5)^x (0.5)^{4-x}, \quad x = 0,1,2,3,4$$

i. Probability of getting at least 2 success

$$P[X \geq 2] = P[X = 2,3,4,\ldots\ldots\ldots]$$

$$= 1 - P[X < 2] = 1 - P[X = 0,1]$$

$$= 1 - \left[4c_0 (0.5)^0 (0.5)^{4-0} + 4c_1 (0.5)^1 (0.5)^{4-1}\right]$$

$$= 1 - \left[(0.5)^4 + 4(0.5)(0.5)^3\right]$$

ii. Probability of getting utmost 2 successes

$$P[X \leq 2] = P[X = 0,1,2]$$

$$= \left[4c_0 (0.5)^0 (0.5)^{4-0} + 4c_1 (0.5)^1 (0.5)^{4-1} + + 4c_2 (0.5)^2 (0.5)^{4-2}\right]$$

$$= \left[(0.5)^4 + 4(0.5)(0.5)^3 + 6(0.5)^2 (0.5)^2\right]$$

$$= (0.5)^4 [1 + 4 + 6]$$

iii. Probability of getting success lies between 5 to 8

$$P[5 < X < 8] = P[X = 6,7] = 0$$

Example 3. Find the moment generating function of Binomial distribution and hence find the mean, variance.

The $p.m.f$ of Binomial Distribution is $P[X = x] = nc_x\, p^x q^{n-x}, x = 0,1,2,\ldots\ldots$

Moment Generating Function of Binomial Distribution is given by	Mean
$M_X(t) = E\left[e^{tx}\right]$ $= \sum_{x=0}^{n} e^{tx} p(x)$ $= \sum_{x=0}^{n} e^{tx} nC_x p^x q^{n-x}$ $= \sum_{x=0}^{n} nC_x \left(e^t p\right)^x q^{n-x}$ $= \left(e^t p + q\right)^n$	$E[X] = M'_X(0)$ $= \left[n\left(p e^t + q\right)^{n-1} p e^t\right]_{t=0}$ $= np$ $E\left[X^2\right] = M''_X(0)$ $= np\left[\left(pe^t q\right)^{n-1} e^t + (n-1)\left(pe^t q\right)^{n-2} pe^t\right]_{t=0}$ $= np\left[1 + (n-1)p\right] = np[1 + np - p]$ $= np + n^2p^2 - np^2$

$$\begin{aligned} Var(X) &= E\left[X^2\right]-\left(E[X]\right)^2 \\ &= np+n^2p^2-np^2-n^2p^2 \\ &= np-np^2 \\ &= np(1-p) \\ &= npq \end{aligned}$$

Poisson Distribution

A random variable X is said to follow binomial distribution if its probability mass function is

$P[X=x]=\dfrac{e^{-\lambda}\lambda^x}{x!}$, $x=0,1,2,\dots\dots\dots$ It is denoted as $X \approx P(\lambda)$ i.e., λ is the parameter.

* It gives probability of x success. * It is useful if n is large and p is small.

* If $X \approx P(\lambda_1)$ and $X \approx P(\lambda_2)$ then $X+Y \approx P(\lambda_1+\lambda_2)$.

Example 1. Find the parameter λ of the Poisson distribution if $P(X=1)=2P(X=2)$.

Given $P(X=1)=2P(X=2)$ and X follows Poisson distribution with $P[X=x]=\dfrac{e^{-\lambda}\lambda^x}{x!}$

$$\frac{e^{-\lambda}\lambda^1}{1!}=2\frac{e^{-\lambda}\lambda^2}{2!}$$

$$\frac{1}{1}=\frac{\lambda}{1}$$

$$\lambda=1$$

Example 2. Find the moment generating function of Poisson distribution and hence find mean, variance.

The $p.m.f$ of Poisson distribution is $P[X=x]=\dfrac{e^{-\lambda}\lambda^x}{x!}$, $x=0,1,2,\dots\dots\dots$

$$\begin{aligned} M_X(t) &= E\left(e^{tx}\right) \\ &= \sum_{x=0}^{\infty} e^{tx}p(x) \\ &= \sum_{x=0}^{\infty} e^{tx}\frac{e^{-\lambda}\lambda^x}{x!} \\ &= e^{-\lambda}\sum_{x=0}^{\infty}\frac{\left(\lambda e^t\right)^x}{x!} \\ &= e^{-\lambda}e^{\lambda e^t} \\ &= e^{\lambda\left(e^t-1\right)} \end{aligned}$$

Mean

$$E(X) = M'_X(0)$$

$$= \left(e^{\lambda(e^t - 1)} \lambda e^t \right)_{t=0}$$

$$= \lambda$$

$$E[X^2] = M''_X(0)$$

$$= \lambda \left[e^{\lambda(e^t - 1)}.e^t + e^t e^{\lambda(e^t - 1)} \lambda e^t \right]_{t=0}$$

$$= \lambda [1 + \lambda]$$

$$= \lambda^2 + \lambda$$

$$Var(X) = E(X^2) - (E(X))^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

Example 3. If X be a random variable following Poisson distribution such that $P(X=2) = 9P(X=4) + 90P(X=6)$. Find the mean, variance of X .

The $p.m.f$ of Poisson distribution is $P[X = x] = \dfrac{e^{-\lambda} \lambda^x}{x!}$, $x = 0,1,2,3,......$

Given $P(X=2) = 9P(X=4) + 90P(X=6)$

Therefore $\dfrac{e^{-\lambda} \lambda^2}{2!} = 9 \dfrac{e^{-\lambda} \lambda^4}{4!} + 90 \dfrac{e^{-\lambda} \lambda^6}{6!}$

Cancelling $\dfrac{e^{-\lambda} \lambda^2}{2}$ on both sides, we get,

$$\frac{1}{1} = \frac{9}{12} \lambda^2 + \frac{90}{360} \lambda^4$$

$$\frac{1}{1} = \frac{3}{4} \lambda^2 + \frac{1}{4} \lambda^4$$

$$\lambda^4 + 3\lambda^2 - 4 = 0$$

$$(\lambda^2 - 1)\ (\lambda^2 + 4) = 0$$

$$\lambda^2 = 1 = 0$$

$$\lambda = 1$$

For a Poisson distribution, mean $\lambda = 1$ and variance $\lambda = 1$

Example 4. It is known that the probability of an item produced by a machine will be defective is 0.05 If the products are sold in packets of 20, find the number of packets containing at least, exactly and at most 2 defective items in a consignment of 1000 packets using Poisson distribution.

Let X represents the number of defective items produced and it follows Poisson distribution. Therefore $P[X = x] = \dfrac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0,1,2,\ldots\ldots.$

Probability for producing one defective item is $p = 0.05$

Products are sold in packets of 20 and hence $n = 20$

Therefore $\lambda = np = 20 \times 0.05 = 1$

Probability of a packet containing exactly 2 defective is $P[X = 2] = \dfrac{e^{-1}1^2}{2!} = \dfrac{1}{2e}$

Therefore number of packets containing exactly 2 defectives in a consignment of 1000 packets is $1000 \times \dfrac{1}{2e} = \dfrac{500}{e}$. .

Probability of a packet containing at most 2 defective

$$P[X \le 2] = P[X = 0,1,2] = \frac{e^{-1}1^0}{0!} + \frac{e^{-1}1^1}{1!} + \frac{e^{-1}1^2}{2!} = \frac{1}{e}\left[1 + 1 + \frac{1}{2}\right] = \frac{5}{2e}$$

Therfore number of packets containing at most 2 defectives in a consignment of 1000 packets is $1000 \times \dfrac{5}{2e} = \dfrac{2500}{e}$. .

Probability of a packet containing at least 2 defective

$$P[X \ge 2] = P[X = 2,3,4,\ldots\ldots..]$$

$$= 1 - P[X < 2]$$

$$= 1 - P[X = 0,\ 1]$$

$$= 1 - \left[\frac{e^{-1}1^0}{0!} + \frac{e^{-1}1^1}{1!}\right]$$

$$= 1 - \frac{2}{e}$$

Therfore number of packets containing at most 2 defectives in a consignment of 1000 packets is $1000\left(1 - \dfrac{2}{e}\right)$.

Example 5. A manufacturer of cotton pins knows that 5% of his product is defective. If he sells cotton pins in boxes of 100 and guarantees that not more than 10 pins will be defective. What is the probability that a box will fail to meet the guaranteed quality.

Let X represents the number of defective pins produced.

Probability for producing one defective item is $p = \frac{5}{100}$

Products are sold in packets of 100 and hence $n = 100$

Therefore $\lambda = np = 100 \times 0.05 = 5$

It follows Poisson distribution and hence $P[X = x] = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x = 0,1,2,.......$

The manufacturer gives a guarantee that a packet may contain maximum 10defectives.

If a box contain more than 10 defective items, then the box will fail to meet the guarantee.

Probability of a packet containing more than 10 defective

$$P[X > 10] = P[\ X = 11,\ 12,\ 13,\ ....\ ]$$

$$= 1 - P[\ X \le 10\ ] = 1 - P[\ X = 0,\ 1,\ 2,\ ....,\ 10]$$

$$= 1 - \left[\frac{e^{-5}\ 5^0}{0\ !} + \frac{e^{-5}\ 5^1}{1\ !} + ...+ \frac{e^{-5}\ 5^{10}}{10\ !} \right]$$

$$= 1 - e^{-5} \left[1 + \frac{5^1}{1\ !} + \frac{5^2}{2\ !} + ...+ \frac{5^{10}}{10\ !} \right]$$

Geometric Distribution

A random variable X is said to follow geometric distribution if its probability mass function is

$P[X = x] = p\,q^x,\ x = 0,1,2,.........\quad (or) \quad P[X = x] = p\,q^{x-1},\ x = 1,2,3,.........$

- It gives probability of first success after x failures.

Example 1. Find the moment generating function and hence find the mean, variance of geometric distribution.

The $p.m.f$ of geometric distribution is $P[X = x] = q^{x-1}\ p,\ x = 1,2,3,........$

$$M_X(t) = E\left[e^{tx}\right]$$

$$= \sum_{x=0}^{n} e^{tx}\ p(x)$$

$$= \sum_{x=1}^{\infty} e^{tx} \; q^{x-1} \; p$$

$$= \frac{p}{q} \sum_{x=0}^{n} \left(q \; e^{t} \right)^{x}$$

$$= \frac{p}{q} \left[1 + \left(q \; e^{t} \right) + \left(q \; e^{t} \right)^{1} + \left(q \; e^{t} \right)^{2} + \dots\dots\dots \right]$$

$$= \frac{p}{q} \left[1 - q \; e^{t} \right]^{-1}$$

$$= \frac{p}{q} \; \frac{1}{1 - q \; e^{t}}$$

Mean $E[X] = M'_X(0) = \dfrac{p}{q} \dfrac{-(-qe^t)}{(1-qe^t)^2} = \left[\dfrac{pe^t}{(1-qe^t)^2} \right]_{t=0} = \dfrac{p}{p^2} = \dfrac{1}{p}$

$$E[X^2] = M'_X(0) = \left[\frac{(1-qe^t)^2 \, pe^t - pe^t \, 2(1-qe^t)(-qe^t)}{(1-qe^t)^4} \right]_{t=0} = \frac{p^3 + 2p^2 q}{p^4} = \frac{p+2q}{p^2}$$

$$Var(X) = E\left[X^2\right] - \left(E[X]\right)^2 = \frac{p+2q}{p^2} - \frac{1}{p^2} = \frac{p+2q-1}{p^2} = \frac{q}{p^2}$$

Example 2. State and prove memoryless property of Geometric Distribution

Statement: $P(M > m+n / X > m) = P(X > n)$

Here X follows Geometric Distribution and hence

$$P[X = x] = q^{x-1} p, \; x = 1,2,3, \dots\dots$$

$P[X > n] = \sum_{x=n+1}^{\infty} P(x)$ $= \sum_{x=n+1}^{\infty} q^{x-1} p$ $= p\left[q^n + q^{n+1} + q^{n+2} + \dots\dots\right]$ $= pq^n\left[1 + q + q^2 + q^3 + \dots\dots\right]$ $= p \; q^n \; [1-q]^{-1}$ $= p \, q^n \, p^{-1}$ $= q^n$	$P[X > m+n / X > m] = \dfrac{P[X > m+n \cap X > m]}{P[X > m]}$ $= \dfrac{P[X > m+n]}{P[X > m]}$ $= \dfrac{q^{m+n}}{q^m}$ $= q^n$ $= P[X > n]$

Example 3. A die is cast until 6 appears. What is the probability that it must cast more than five times.

Let X represents the number of tosses required to get the first 6.

Probability of getting 6 is $p = \frac{1}{6}$ and hence $q = 1 - p = \frac{5}{6}$

Also $P[X = x] = q^{x-1}\, p \;=\; \left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^{x-1}$, $x = 1, 2, \ldots\ldots$

$$P[X > 5] = 1 - P[X \le 5] = 1 - \{P[X = 1] + P[X = 2] + P[X = 3] + P[X = 4] + P[X = 5]\}$$

$$= 1 - \left\{\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^0 + \left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^1 + \left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^2 + \left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^3 + \left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^4\right\}$$

$$= 1 - \left(\frac{1}{6}\right)\left\{1 + \left(\frac{5}{6}\right)^1 + \left(\frac{5}{6}\right)^2 + \left(\frac{5}{6}\right)^3 + \left(\frac{5}{6}\right)^4\right\} = 1 - 0.5981 = 0.4019$$

Example 4. A trainee soldier shoots a target in an independent fashion. If the probability that the target is shot on any one shot is 0.8. What is the probability that the target would be first hit at the 6th attempt?. What is the probability that it takes less than 5 shots?

Let X represents the number of shots required to hit the target first.

Probability of hitting the target is $p = 0.8$ and hence $q = 1 - p = 0.2$

Also $P[X = x] = q^{x-1}\, p \;=\; (0.8)\,(0.2)^{x-1}$, $x = 1, 2, \ldots\ldots$

(i) Probability of hitting the target on 6th attempt

$$P[X = 6] = (0.8)\,(0.2)^{6-1} = (0.8)\,(0.2)^5 = 0.00026$$

(ii) Probability of hitting the target in less than 5 attempt

$$P[X < 5] = P[X = 1] + P[X = 2] + P[X = 3] + P[X = 4]$$

$$= (0.8)(0.2)^0 + (0.8)(0.2)^1 + (0.8)(0.2)^2 + (0.8)(0.2)^3$$

$$= (0.8)\left\{1 + (0.2)^1 + (0.2)^2 + (0.2)^3\right\} = 0.9984$$

Example 5. A candidate is applying for driving license has the probability of 0.8 in passing the road test in a given trial. What is the probability that he will pass the test (i) on the fourth trial (ii) in less than four trials.

Let X represents the number of trails required to get the first success.

Probability of getting the license is p = 0.8 and hence q = 1 – p = 0.2

Also $P[X = x] = q^{x-1}\, p = (0.8)\,(0.2)^{x-1}$, $x = 1, 2, \ldots\ldots$

(i) Probability of getting the license in the 4th trial

$$P[X = 4] = (0.8)\,(0.2)^{4-1} = (0.8)\,(0.2)^3 = 0.0064$$

(ii) Probability of hitting the target in less than 4 attempt

$$P[X < 4] = P[X = 1] + P[X = 2] + P[X = 3]$$

$$= (0.8)(0.2)^0 + (0.8)(0.2)^1 + (0.8)(0.2)^2$$

$$= (0.8)\left\{1 + (0.2)^1 + (0.2)^2\right\} = 0.992$$

Uniform Distribution

Let X be a uniform distribution defined in the interval (a, b) then its probability density function is of the form $f(x) = \frac{1}{b-a}$, $a < x < b$.

Example 1. Find the value of 'a' if X follows uniform distribution in the interval $(a, 9)$ and $P(3 < X < 5) = \frac{2}{7}$.

Given X follows uniform distribution in the interval $(a, 9)$ and hence

$f(x) = \frac{1}{9-a}$, $a < x < 9$

Also given that

$$P[\ 3 < X < 5\ ] = \frac{2}{7}$$

$$\int_3^5 f(x)\ dx = \frac{2}{7}$$

$$\int_3^5 \frac{1}{9 - a}\ dx = \frac{2}{7}$$

$$\frac{1}{9 - a}\ [\ x\ ]_3^5 = \frac{2}{7}$$

$$\frac{2}{9 - a} = \frac{2}{7}$$

$$\therefore\ a = 2$$

Example 2. Find the moment generating function and mean, variance of uniform distribution where $X \approx U(a,b)$.

Here X follows uniform distribution in (a, b) and hence $f(x) = \frac{1}{b-a}, a<x<b$

The moment generating function is given by

$$M_X(t) = E\left[e^{tx}\right]$$

$$= \int_a^b e^{tx} \; f(x)\, dx$$

$$= \frac{1}{b-a}\int_a^b e^{tx} \; dx$$

$$= \frac{1}{b-a}\left[\frac{e^{tx}}{t}\right]_a^b$$

$$= \frac{1}{b-a}\left[\frac{e^{tb}}{t} - \frac{e^{ta}}{t}\right]$$

$$E[\ X\ ] = \int_0^1 x \;\; f(x) \;\; dx$$

$$= \frac{1}{b-a} \int_a^b x \;\; dx$$

$$= \frac{1}{b-a}\left[\frac{x^2}{2}\right]_a^b$$

$$= \frac{1}{b-a}\left[\frac{b^2 \; - \; a^2}{2}\right]$$

$$= \frac{a+b}{2}$$

$$E[\ X^2\ ] = \int_0^1 x \;\; f(x) \;\; dx$$

$$= \frac{1}{b-a} \int_a^b x^2 \;\; dx$$

$$= \frac{1}{b-a}\left[\frac{x^3}{3}\right]_a^b$$

$$= \frac{1}{b-a}\left[\frac{b^3 \; - \; a^3}{2}\right]$$

$$= \frac{1}{b-a}\left[\frac{(b-a)\left(b^2 \; +ab+ \; a^2\right)}{3}\right]$$

$$Var(X) = E\left[X^2\right] - \left[E(X)\right]^2 = \frac{a^2+b^2+ab}{3} - \frac{(a+b)^2}{4} = \frac{(a-b)^2}{12}$$

Example 3. Find the $m.g.f$ of uniform distribution in the interval $(0, 1)$. Also find the mean, variance of it.

Here X follows uniform distribution in $(0, 1)$ and hence $f(x) = \frac{1}{1-0}, 0<x<1$

The moment generating function is given by

$$M_X(t) = E\left[e^{tx}\right]$$

$$= \int_0^1 e^{tx}\ f(x)\,dx$$

$$= \frac{1}{1}\int_0^1 e^{tx}\ dx$$

$$= \left[\frac{e^{tx}}{t}\right]_0^1$$

$$= \left[\frac{e^{t}}{t} - \frac{1}{t}\right]$$

$$E[\ X\ ] = \int_0^1 x\ \ f(x)\ \ dx \qquad E\left[X^2\right] = \int_0^1 x\ \ f(x)\ \ dx$$

$$= \int_0^1 x\ \ dx \qquad\qquad = \int_0^1 x^2\ \ dx$$

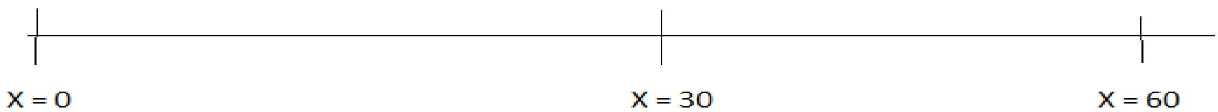
$$= \left[\frac{x^2}{2}\right]_0^1 \qquad\qquad = \left[\frac{x^3}{3}\right]_0^1$$

$$= \frac{1}{2} \qquad\qquad = \frac{1}{3}$$

$$Var(X) = E\left[X^2\right] - \left[\ E(X)\ \right]^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$$

Example 4. Electric trains in a route run every half an hour. Find the probability that a passenger entering the station will have to wait (i) at least 20 minutes (ii) less than 10 minutes.

Electric trains in a route run every half an hour. Therefore it follows uniform distribution with pdf $f(x) = \frac{1}{30}$, $0 < x < 30$



Probability for a passenger to wait at least 20 minutes. It happens if he arrives in between 0 to 10 minutes.

$$P[0 < X < 10] = \int_0^{10} f(x)dx = \frac{1}{30}\int_0^{10} dx = \frac{1}{30}\left[x\right]_0^{10} = \frac{10}{30} = \frac{1}{3}$$

Probability for a passenger to wait less than 10 minutes. It happens if he arrives in between 20 to 30 minutes.

$$P[20<X<30]=\int_{20}^{30} f(x)dx=\frac{1}{30}\int_{20}^{30} dx=\frac{1}{30}[x]_{20}^{30}=\frac{10}{30}=\frac{1}{3}$$

Example 5. A random variable X has an uniform over the interval $(-3,3)$. Compute *i.* $P[X<2]$ *ii.* $P[|X|<2]$ **iii. Find k such that $P[X>k]=\frac{1}{3}$**

Here X follows uniform in the interval $(-3,3)$ and hence $f(x)=\frac{1}{3+3}, -3<x<3$

$$P[X<2]=P[-3<X<2]=\int_{-3}^{2} f(x)dx=\frac{1}{6}\int_{-3}^{2} dx=\frac{1}{6}[x]_{-3}^{2}=\frac{5}{6}$$

$$P[|X|<2]=P[-2<X<2]=\int_{-2}^{2} f(x)dx=\frac{1}{6}\int_{-2}^{2} dx=\frac{1}{6}[x]_{-2}^{2}=\frac{4}{6}=\frac{2}{3}$$

Given $P[X>k]=\frac{1}{3}$

$$P[X>k]=\frac{1}{3}$$

$$P[k<x<3]=\frac{1}{3}$$

$$\int_{k}^{3} f(x)\,dx=\frac{1}{3}$$

$$\frac{1}{6}[x]_{k}^{3}=\frac{1}{3}$$

$$\frac{3-k}{6}=\frac{1}{3}$$

$$k=1$$

Example 6. Let X be a uniform random variable with mean 1 and variance $\frac{4}{3}$. Find $P[X<0]$.

We know that, if X follows uniform distribution in (a,b), then

$$Mean=\frac{a+b}{2},\quad Var=\frac{(b-a)^2}{12}$$

Given mean = 1 and variance $=\frac{4}{3}$.

Therefore $1=\frac{a+b}{2}$ *and* $\frac{4}{3}=\frac{(b-a)^2}{12}$

i.e. $a+b=2$ .....(i) *and* $(b-a)^2=16$ *i.e.* $b-a=4$ ....(ii)

Solving (i) and (ii), we get $a = -1$ and $b = 3$

Therefore $f(x) = \dfrac{1}{b - a} = \dfrac{1}{4}, \quad -1 < x < 3$

$$P[\ X < 0\ ] = P[\ -1 < X < 0\ ] = \int_{-1}^{0} f(x)dx = \frac{1}{4}\int_{-1}^{0} dx = \frac{1}{4}[\ x\ ]_{-1}^{0} = \frac{1}{4}$$

Exponential Distribution

A continuous random variable X is said to have exponential distribution with parameter $\lambda > 0$ if Its probability density function is of the form $f(x) = \lambda e^{-\lambda x}$, $0 < x < \infty$

Example 1. Find the moment generating function, mean, variance of exponential distribution.

The moment generating function of exponential distribution is given by

$$M_X(t) = E\left[e^{tx}\right]$$

$$= \int_{-\infty}^{\infty} e^{tx}\ f(x)\, dx$$

$$= \lambda \int_{0}^{\infty} e^{tx}\ e^{-\lambda x}\, dx$$

$$= \lambda \int_{0}^{\infty} e^{-(\lambda - t)x}\, dx$$

$$= \lambda \left[\frac{e^{-(\lambda - t)x}}{-(\lambda - t)} \right]_0^{\infty}$$

$$= \lambda \left[0 - \frac{1}{-(\lambda - t)} \right]$$

$$= \frac{\lambda}{(\lambda - t)}$$

$$M_X'(t) = \frac{(\lambda - t)(0) - \lambda(-1)}{(\lambda - t)^2} = \frac{\lambda}{(\lambda - t)^2}$$

$$M_X''(t) = \frac{(\lambda - t)^2(0) - \lambda 2(\lambda - t)(-1)}{(\lambda - t)^4} = \frac{2\lambda}{(\lambda - t)^3}$$

$$Mean \quad E(X) = M'_X(0) = \frac{1}{\lambda}$$

$$E(X^2) = M''_X(0) = \frac{2}{\lambda^2}$$

$$Var(X) = E(X^2) - [E(X)]^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

Example 2. State and prove memoryless property of exponential distribution.

If X is exponentially distributed with parameter λ, then for any two positive integers m and n,

$$P[X > m+n \,/\, X > m] = P[X > n]$$

$$P[X > n] = P[n < X < \infty]$$

$$= \int_n^\infty f(x)\, dx$$

$$= \lambda \int_n^\infty e^{-\lambda x}\, dx$$

$$= \lambda \left[\frac{e^{-\lambda x}}{-\lambda} \right]_n^\infty$$

$$= e^{-n\lambda} \qquad ------(1)$$

$$P[X > m+n \,/\, X > m] = \frac{P[X > m+n \cap X > m]}{P[X > m]}$$

$$= \frac{P[X > m+n]}{P[X > m]}$$

$$= \frac{e^{-(m+n)}}{e^{-m}}, \qquad u\sin g\,(1)$$

$$= e^{-n}$$

$$= P[X > n]$$

Example 3. Find $P[X > 10]$ if the probability density function of X is $f(x) = e^{-x}, x > 0$.

Given $f(x) = e^{-x}, x > 0$.

$$P[X > 10] = P[10 < X < \infty]$$

$$= \int_{10}^\infty f(x)\, dx$$

$$= \int_{10}^{\infty} e^{-x} \, dx$$

$$= \left[\frac{e^{-x}}{-1} \right]_{10}^{\infty}$$

$$= e^{-10}$$

Example 4. Suppose that during the rainy season in an island, the length of the shower has an exponential distribution with average 2 minutes. Find the probability that the shower will be there for more than 3 minutes. If the shower has already lasted for 2 minutes, what is the probability that it will last for al least 1 more minute.

Let X represents the length of the shower in minutes and it follows exponential distribution with mean 2. But exponential distribution has mean $\frac{1}{\lambda}$.

Therefore $\frac{1}{\lambda} = 2 \Rightarrow \lambda = \frac{1}{2}$.

The pdf of exponential distribution is $f(x) = \lambda e^{-\lambda x} = \frac{1}{2} e^{-\frac{1}{2}x}, 0 < x < \infty$

(i) Probability that the shower lasts more than 3 minutes

$$P[X > 3] = P[3 < X < \infty]$$

$$= \int_{3}^{\infty} f(x) \, dx$$

$$= \frac{1}{2} \int_{3}^{\infty} e^{-\frac{1}{2}x} \, dx$$

$$= \frac{1}{2} \left[\frac{e^{-\frac{1}{2}x}}{-\frac{1}{2}} \right]_{3}^{\infty}$$

$$= e^{-\frac{3}{2}}$$

(iii) Probability that the shower will last at least one minute, given that it had lasted 2 Minutes

$$P[\, X > 3 \ / \ X > 2 \,] = P[\, X > 1 \,]$$

$$= P[\, 1 < X < \infty]$$

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

$$= \frac{1}{2} \int_{1}^{\infty} e^{-\frac{1}{2}x} \, dx$$

$$= \frac{1}{2} \left[\frac{e^{-\frac{1}{2}x}}{-\frac{1}{2}} \right]_{1}^{\infty}$$

$$= e^{-\frac{1}{2}}$$

Example 5. The life of a lamp (in 1000's of hours) is exponentially distributed with parameter $\lambda = \frac{1}{3}$. Find (i) the probability that the lamp will last longer than its mean life of 3000 hours.
(ii) the probability that the lamp will last for another 1000 hours given that it is operating after 2500 hours.

Let X represents the life of the bulb (in 1000 hours) and it follows exponential distribution with parameter $\lambda = \frac{1}{3}$. Therefore its average life is 3000 hours.

The pdf of exponential distribution is $f(x) = \lambda e^{-\lambda x} = \frac{1}{3} e^{-\frac{1}{3}x}, 0 < x < \infty$

(i) Probability that the life of the bulb lasts more than its average life3000 hours.

$$P[X > 3] = P[3 < X < \infty]$$

$$= \int_{3}^{\infty} f(x) \, dx$$

$$= \frac{1}{3} \int_{3}^{\infty} e^{-\frac{1}{3}x} \, dx$$

$$= \frac{1}{3} \left[\frac{e^{-\frac{1}{3}x}}{-\frac{1}{3}} \right]_{3}^{\infty}$$

$$= e^{-1}$$

(ii) Probability that the shower will last more than 1000 hours, given that it had already lasted 2500 hours

$$P[X > 3.5 \;/\; X > 2.5] = P[X > 1], \quad \textit{by memoryless property}$$

$$= P[1 < X < \infty]$$

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

$$= \frac{1}{3} \int_{1}^{\infty} e^{-\frac{1}{3}x} \, dx$$

$$= \frac{1}{3}\left[\frac{e^{-\frac{1}{3}x}}{-\frac{1}{3}}\right]_1^{\infty}$$

$$= e^{-\frac{1}{3}}$$

Example 6. The daily consumption of milk in a city in excess of 20,000 litres is exponentially distributed. The average excess in consumption of milk is 3,000 litres. The city has a daily stock of 35,000 litres. What is the probability that a day is selected at random and the stock is insufficient for that day.

Let X be the random variable of daily consumption of milk in excess of 20,000 litres.

Now X follows exponential distribution with mean 3000 litres

But exponential distribution has mean $\frac{1}{\lambda}$. Therefore $\frac{1}{\lambda} = 3000 \Rightarrow \lambda = \frac{1}{3000}$.

The pdf of exponential distribution is $f(x) = \lambda e^{-\lambda x} = \frac{1}{3000} e^{-\frac{1}{3000}x}, 0 < x < \infty$

Let Y be the daily consumption of milk. Then $X = Y - 20000$

Probability for the stock is insufficient happens if the daily consumption is more than the stock 35000 litres.

$$P[Y > 35000] = P[X + 20000 > 35000]$$

$$= P[X > 15000]$$

$$= P[15000 < X < \infty]$$

$$= \int_{15000}^{\infty} f(x)\, dx$$

$$= \frac{1}{3000}\int_{15000}^{\infty} e^{-\frac{1}{3000}x}\, dx$$

$$= \frac{1}{3000}\left[\frac{e^{-\frac{1}{3000}x}}{-\frac{1}{3000}}\right]_{15000}^{\infty}$$

$$= \frac{1}{3000}\left[0 - \frac{e^{-5}}{-\frac{1}{3000}}\right]$$

$$= e^{-5}$$

Normal Distribution

Example 1. **If X is a normal variate with $\mu = 30$ and $\sigma = 5$, find**

(i) $P(26 \le X \le 40)$

(ii) $P(X \ge 45)$ (iii) $P(|X - 30| \ge 5)$

Given X is a normal variate with $\mu = 30$ and $\sigma = 5$. Then $Z = \dfrac{X-\mu}{\sigma} = \dfrac{X-30}{5}$

$$(i)\ P(26 \le X \le 40) = P\left[\frac{X-30}{5} < Z < \frac{X-30}{5}\right]$$

$$= P\left[\frac{26-30}{5} < Z < \frac{40-30}{5}\right]$$

$$= P[-0.8 < Z < 2]$$

$$= P[0 < Z < 0.8] + P[0 < Z < 2]$$

$$= 0.2881 + 0.4772$$

$$= 0.7653$$

$$(ii)\ P(X \ge 45) = P\left[Z \ge \frac{X-30}{5}\right]$$

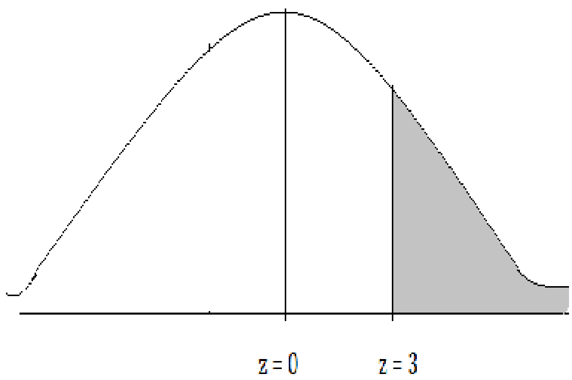
$$= P\left[Z \ge \frac{45-30}{5}\right]$$

$$= P[Z \ge 3]$$

$$= 0.5 - P[0 < Z < 3]$$

$$= 0.5 - 0.4987$$

$$= 0.0013$$



$(iii)\ P(|X-30|>5) = 1 - P(|X-30|<5)$

$$= 1 - P[-5 < X - 30 < 5]$$

$$= 1 - P[25 < X < 35]$$

$$= 1 - P\left[\frac{X-30}{5} < Z < \frac{X-30}{5}\right]$$

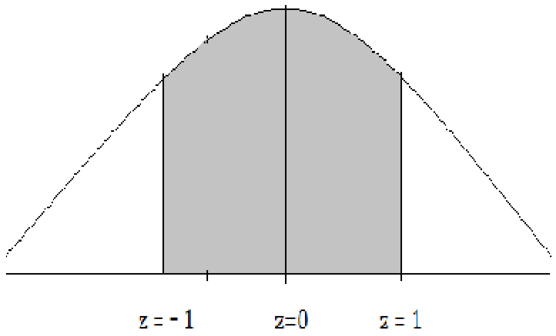
$$= 1 - P\left[\frac{25-30}{5} < Z < \frac{35-30}{5}\right]$$

$$= 1 - P[-1 < Z < 1]$$

$$= 1 - 2P[0 < Z < 1]$$

$$= 1 - 2\,(0.3413)$$

$$= 0.3174$$



Example 2. A certain type of storage battery lasts on the average 3 years with standard deviation 0.5 year. Assuming that the battery lives are normally distributed, find the probability that the given battery will last less than 2.3 years.

Given X is a normal variate with $\mu = 3$ and $\sigma = 0.5$ Then $Z = \frac{X-\mu}{\sigma} = \frac{X-3}{0.5}$

To find the probability that the battery will last less than 2.3 years

$$P[X < 2.3] = P\left[Z < \frac{X-3}{0.5}\right]$$

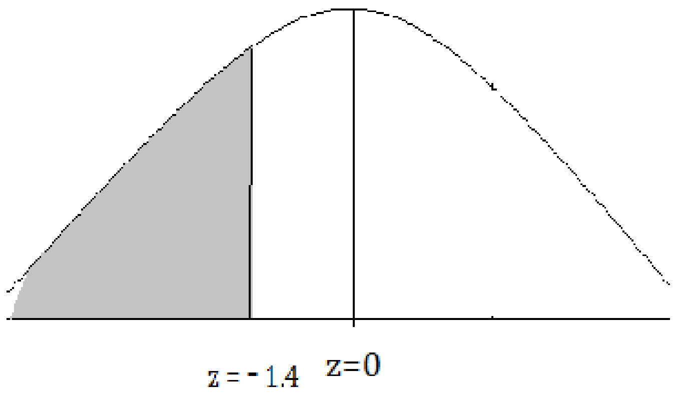
$$= P\left[Z < \frac{2.3-3}{0.5}\right]$$

$$= P[Z < -1.4]$$

$$= 0.5 - P[0 < Z < 1.4]$$

$$= 0.5 - 0.4192$$

$$= 0.0808$$



Example 3. The average percentage of marks of candidates in an examination is 42 with a standard deviation of 10 .If the minimum mark for pass is 50% and 1000 candidates appear for the examination, how many candidates can be expected to get the pass mark if the marks follow normal distribution? If it is required, that double the number of the candidates should pass, What should be the minimum mark for pass?

Let X denote the marks of the candidates. Given $\mu = 42,\ \sigma = 10$.

Let $Z=\frac{X-\mu}{\sigma}=\frac{X-42}{10}$

Probability to get pass marks $=P(X\geq 50)$

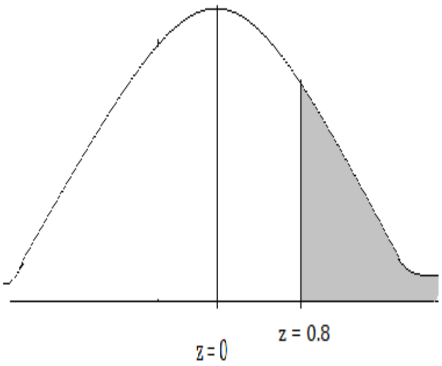
$$=P\left(Z\geq \frac{50-42}{10}\right)$$

$$=P(Z\geq 0.8)$$

$$=\left[0.5-P(0\leq Z\leq 0.8)\right]$$

$$=\left[0.5-0.2881\right]$$

$$=0.2119$$



Therefore expected number of students to get pass marks $=1000\times P(X\geq 50)$

$$=1000\times 0.2119$$
$$=212$$

To double the number of passed students i.e. 424. Then $P(Z>z)=0.424$

$$P(0<Z<z)=0.5-0.424=0.076$$

From the normal table, $z=0.19$

$$0.19=\frac{50-x}{10}$$

$$x\approx 48$$

Therefore the pass mark should be nearly 48

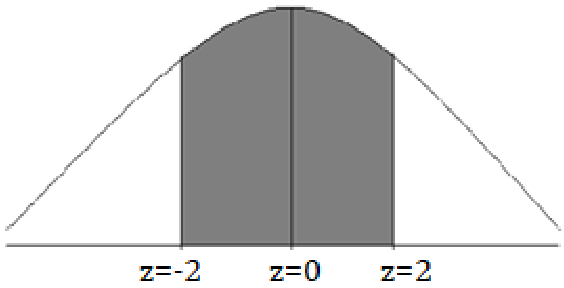
Example 4. A production line manufactures 1000 ohm resistors that have 10% tolerance. Let X denotes the resistance of resistor. Assuming that X is a normal random variable with mean 1000 and variance 2500, find the probability that a resistor picked at random will be rejected.

Given that the resistance of resistor X is a normal random variable with mean $\mu=1000$ and variance $\sigma^2=2500$. Therefore $Z=\frac{X-1000}{\sqrt{2500}}$.

1000ohm resistor is produced. Since the tolerance limit is 10%, the resistor is accepted if the resistance capacity is 900ohm to 1100ohm.

Therefore probability for a resistor to be accepted is $P(900<X<1100)$.

Therefore probability for a resistor to be rejected is $=1-P(900<X<1100)$



$$=1-P\left(\frac{900-1000}{\sqrt{2500}}<Z<\frac{1100-1000}{\sqrt{2500}}\right)$$

$$=1-P(-2<Z<2)$$

$$=1-2P(0<Z<2)$$

$$=1-2\times(0.4772)$$

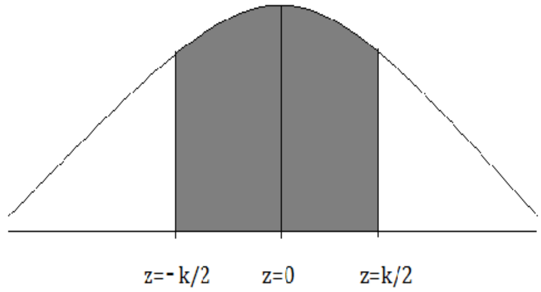
$$=0.0456$$

Example 5. **If** X **is** $N(3,4)$**, find** k **so that** $P(|X-3|>k)=0.05$

Given X is a normal variate with

$\mu=3,\ \sigma=2$.

Then $Z=\dfrac{X-\mu}{\sigma}=\dfrac{X-3}{2}$



Given that

$$P\left[\,|X-3|>k\,\right]=0.05$$

$$1-P\left[\,|X-3|<k\,\right]=0.05$$

$$P\left[\,|X-3|<k\,\right]=1-0.05$$

$$P\left[-k<X-3<k\right]=0.95$$

$$P\left[-k<X-3<k\right]=0.95$$

$$P\left[3-k<X<k+3\right]=0.95$$

$$P\left[\frac{X-3}{2}<Z<\frac{X-3}{2}\right]=0.95$$

$$P\left[\frac{3-k-3}{2}<Z<\frac{k+3-3}{2}\right]=0.95$$

$$P\left[\frac{-k}{2}<Z<\frac{k}{2}\right]=0.95$$

$$2\ P\left[0<Z<\frac{k}{2}\right]=0.95$$

$$P\left[0<Z<\frac{k}{2}\right]=0.475$$

$$\frac{k}{2}=1.96\ \text{\{from the normal table\}}$$

i.e. $k=3.92$

Example 6. Given that X is distributed normally with $P(X<45)=0.31$ and $P(X<64)=0.08$.

Find the mean and standard deviation of the distribution.

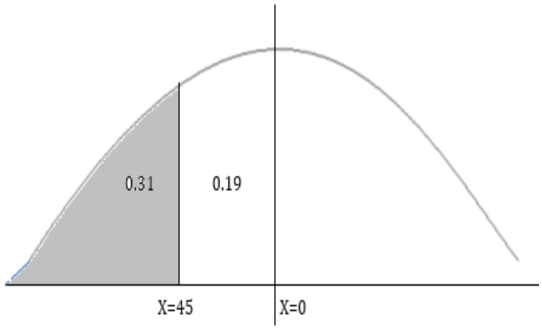
Given $P(X<45)=0.31$.

Therefore $P(45<X<0)=0.19$

The value of z corresponds to the area 0.19 = 0.5

$$\therefore \frac{45-\mu}{\sigma}=-0.5$$

$$45-\mu=-0.5\sigma \quad ----(1)$$



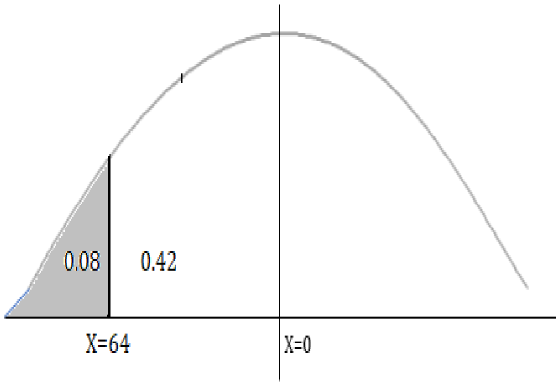
Given $P(X<64)=0.08$.

Therefore $P(64<X<0)=0.42$

The value of z corresponds to the area 0.42 = 1.4

$$\therefore \frac{64-\mu}{\sigma}=-1.4$$

$$64-\mu=-1.4\sigma \quad ----(2)$$



Solving (1) and (2) mean=50, SD=10.

Exercise

1. If $P(A)=0.35$, $P(B)=0.45$, find $P\left(A\cap\overline{B}\right)$ where A and B are mutually exclusive events.

2. If $P(A)=0.3$, $P(B)=0.7$ and $P(A\cap B)=0.1$ find $P\left(\overline{A}\cup\overline{B}\right)$.

3. If A and B are mutually exclusive events with $P(A)=0.31$, $P(B)=0.17$, find $P(A\cup B)$.

4. What is the probability of drawing one white ball from a bag containing 6 red 8 black and 10 yellow balls?

5. If two dice are tossed simultaneously what is the probability of getting 7 as the sum of the resultant faces?

6. If four coins are tossed simultaneously what is the probability of getting exactly 3 heads?

7. When a pair of balanced dice are rolled what is the probability of getting 2 or 12.

8. A fair coin is tossed 3 times. What is the probability of getting exactly two tails?

9. Let an urn contains 4 tickets numbered 1 , 2 , 3 , 4 and another urn contains 6 tickets numbered 2 , 4 , 6 , 7 , 8 , 9 . If one of the two urns is chosen at random and a ticket is drawn at random from the chosen urn, find the probabilities that the ticket drawn bears the number 1 or 9

10. In a bolt factory, machines A, B and C manufacture 25% , 35% and 40% respectively of the total output. The probability of a defective product are respectively 5 , 4, 2 percent from the three plants. A bolt is chosen at random from a days production and found to be defective. What is the probability that it was manufactured by C ?

11. A certain firm has plants A, B and C producing 35 % , 15 % and 50 % respectively of the total output. The probability of a non defective product are respectively 0.75 , 0.95 and 0.85 from three plants. A customer receives a defective product. What is the probability that it came from plant C ?.

12. Suppose that coloured balls are distributed in three indistinguishable boxes as follows:

	Box 1	Box 2	Box 3
Red	2	4	3
White	3	1	4
Blue	5	3	5

A box is selected at random from which a ball is selected at random and it is observed to be red. What is the probability that the box 3 was selected?

13. Three machines A, B and C of equal capacities are producing washers. The probabilities that these machines produce defective washers are 0.15 , 0.24 and 0.18. A washer taken at random from a days production was found to be defective. Find the probability that it was produced by C .

14. If two fair dice are thrown, find the expectation of the sum of outcomes and also find the variance.

15. There are four different choices available to the customer who wants to buy a transistor set. The first type costs Rs. 400 , the second type Rs. 340 , the third type Rs. 440 and the fourth type Rs. 380. The probability that the customer will buy these types is $\frac{1}{3}$, $\frac{1}{6}$, $\frac{1}{4}$ and $\frac{1}{4}$ respectively. The retailer of these transistors gets a commission 20 % , 12 % , 25 % and 15 % for these sets respectively. What is the expected commission of the retailer ?

16. For scores X on a admission aptitude test , the mean $E(X)=60$ and $V(X)=40$. Find the mean and variance of $\frac{X-60}{10}$.

17. Let X be a discrete random variable with $p.m.f$ $P(X=x)=\begin{cases}\frac{x}{10} & ;x=1,2,3,4\\ 0 & otherwise\end{cases}$.

Compute $P(X<3)$ and $E\left(\frac{X}{2}\right)$.

18. A continuous random variable has the pdf given by $f(x)=\begin{cases}\alpha(1+x^2) & ,2\le x\le 5\\ 0 & otherwise\end{cases}$.
Find α and $P(X<4)$.

19. The DF of a continuous random variable X is given by $F(x)=\begin{cases}0, & x<0\\ x^2, & 0\le x<\frac{1}{2}\\ 1-\frac{3}{25}(3-x)^2, & \frac{1}{2}\le x<3\\ 1, & x\ge 3\end{cases}$.

Find the pdf of X and evaluate $P(|X|\le 1)$ and $P\left(\frac{1}{3}<X<4\right)$ using both pdf and cdf.

20. Find the MGF of the random variable X having the pdf $f(x)=\begin{cases}x, & 0<x<1\\ 2-x, & 1<x<2\\ 0, & 2<x<\infty\end{cases}$.

21. 6 dice are thrown 729 times. How many times do you expect at least three dice to show a five (or) a six?

22. The number of monthly breakdown of a computer is a random variable having a Poisson distribution with mean equal to 1.8. Find the probability that this computer will function for a month (i) without a breakdown (ii) with only one breakdown.

23. If the probability that an applicant for a driver's license will pass the road test on any given trail is 0.8, what is the probability that he will finally pass the test (i) on the 4^{th} trial (ii) in fewer than 4 trails?

24. Let X be a uniformly distributed random variable over $[-5,-5]$. Determine $(i)P(X\le 2)$ (ii) $P(|X|>2)$

ANSWERS

1. $P(A)$ **2.** 0.9 **3.** 0.48 **4.** 0 **5.** 0.1666 **6.** 0.25 **7.** 0.0555 **8.** 0.1111 **9.** 0.2083

10. 0.2318 **11.** 0.4411 **12.** 0.1132 **13.** 0.3157 **14.** M=7, V=5.8 **15.** 75.22

16. M=0, V=0.5 **17.** $\frac{3}{10}, \frac{3}{2}$ **18.** $\frac{1}{42}, \frac{31}{63}$ **19.** $f(x)=\frac{d}{dx}F(x)$, $\frac{13}{25}, \frac{8}{9}$ **20.** $\frac{1}{t^2}\left(e^t-1\right)^2$

21. $729\times P\left(X\geq 3\right)=233$ **22.** 0.16, 0.29, 0.83 **23.** 0.0064, 0.992 **24.** $\frac{7}{10}, \frac{6}{10}$

Unit-II Two Dimensional Random Variables

Two Dimensional Discrete Random Variable

The probability mass function of a two dimensional discrete random variable taking the values $x_0,\ x_1,\ x_2$ and $y_0,\ y_1,\ y_2$ is tabulated as follows:

$X \backslash Y$	y_0	y_1	y_2
x_0	p_{00}	p_{01}	p_{02}
x_1	p_{10}	p_{11}	p_{12}
x_2	p_{20}	p_{21}	p_{22}

Here all $p_{ij} \geq 0$ and $\sum\sum p_{ij} = 1$. Notation: $P[X = x_0, Y = y_0] = p_{00}$

Necessary Formulae:

Marginal distribution of X

$$P[X = x_0] = p_{00} + p_{01} + p_{02}$$

$$P[X = x_1] = p_{10} + p_{11} + p_{12}$$

$$P[X = x_2] = p_{20} + p_{21} + p_{22}$$

Marginal distribution of Y

$$P[Y = y_0] = p_{00} + p_{10} + p_{20}$$

$$P[Y = y_1] = p_{01} + p_{11} + p_{21}$$

$$P[Y = y_2] = p_{02} + p_{12} + p_{22}$$

Conditional distribution of X given $Y = y_0$	Conditional distribution of Y given $X = x_1$
$P[X = x_0 / Y = y_0] = \dfrac{P[X = x_0, Y = y_0]}{P[Y = y_0]} = \dfrac{p_{00}}{P[Y = y_0]}$	$P[Y = y_0 / X = x_1] = \dfrac{P[Y = y_0, X = x_1]}{P[X = x_1]} = \dfrac{p_{10}}{P[X = x_1]}$
$P[X = x_1 / Y = y_0] = \dfrac{P[X = x_1, Y = y_0]}{P[Y = y_0]} = \dfrac{p_{10}}{P[Y = y_0]}$	$P[Y = y_1 / X = x_1] = \dfrac{P[Y = y_1, X = x_1]}{P[X = x_1]} = \dfrac{p_{11}}{P[X = x_1]}$

https://doi.org/10.5281/zenodo.15090192

$P[X = x_2 / Y = y_0] = \dfrac{P[X = x_2, Y = y_0]}{P[Y = y_0]}$ $= \dfrac{p_{20}}{P[Y = y_0]}$	$P[Y = y_2 / X = x_1] = \dfrac{P[Y = y_2, X = x_1]}{P[X = x_1]}$ $= \dfrac{p_{12}}{P[X = x_1]}$

If X and Y are independent, then $P[X = x_i, Y = y_j] = P[X = x_i] \times P[Y = y_j]$ $for\,all\ i, j$

Mean of X : $E[X] = \sum x P[X = x_i] \quad \forall i$ Mean of Y : $E[Y] = \sum y P[Y = y_j] \quad \forall j$

$E[XY] = \sum \sum xy P(x,y) \quad \forall x, y$

Conditional expectation of X given $Y = y_j$: $E[X / Y = y_j] = \sum x_i P[X = x_i / Y = y_j] \forall i$

Conditional expectation of Y given $X = x_j$: $E[Y / X = x_i] = \sum y_j P[Y = y_j / X = x_i] \forall j$

$$P[X \le x_1 \ , \ Y < y_2] = P[X = x_0 \ , \ x_1 \ , \ Y = y_0 \ , \ y_1]$$
$$= P[X = x_0, \ Y = y_0] + P[X = x_0 \ , \ Y = y_1] +$$
$$P[X = x_1 \ , \ Y = y_0] + P[X = x_1 \ , \ Y = y_1]$$

Example 1. The joint probability function of the random variables (X, Y) is given by
$P(X,Y) = k(2x + 3y), \ x = 0,1,2; \ y = 1,2,3$

i. **Find the marginal distributions**
ii. **Find the probability distribution of $X + Y$**
iii. **Find all the conditional distribution of X given Y**

The joint probability distribution function of (X, Y) is given below:

X \ Y	1	2	3
0	$3k$	$6k$	$9k$
1	$5k$	$8k$	$11k$
2	$7k$	$10k$	$13k$

To find k: we know that $\sum \sum P(x,y) = 1$

$$72k = 1$$

$$k = \frac{1}{72}$$

Marginal distribution of X .

$$P[X = 0] = 3k + 6k + 9k = 18k = \frac{18}{72}$$

$$P[X = 1] = 5k + 8k + 11k = 24k = \frac{24}{72}$$

$$P[X = 2] = 7k + 10k + 13k = 30k = \frac{30}{72}$$

Marginal distribution of Y

$$P[Y = 1] = 3k + 5k + 7k = 15k = \frac{15}{72}$$

$$P[Y = 2] = 6k + 8k + 10k = 24k = \frac{24}{72}$$

$$P[Y = 3] = 9k + 11k + 13k = 33k = \frac{33}{72}$$

Probability distribution of $X + Y$: Here $X + Y$ takes the values 1, 2, 3, 4, 5

Possible combination	(0,1)	(1,1) or (0,2)	(0,3) or (1,2) or (2,1)	(2,2) or (1,3)	(2,3)
$X + Y$	1	2	3	4	5
$P(X + Y)$	$3k$ $= \frac{3}{72}$	$5k + 6kk$ $= 11k$ $= \frac{11}{72}$	$9k + 8k + 7k$ $=24k$ $= \frac{24}{72}$	$10k + 11k$ $=21k$ $= \frac{21}{72}$	$13k$ $= \frac{13}{72}$

.

Conditional distribution of X given Y = 1:

$$P[X = x / Y = 1]$$

$$P[X = 0 / Y = 1] = \frac{P[X = 0, Y = 1]}{P[Y = 1]}$$

$$= \frac{3k}{15k} = \frac{3}{15}$$

$$P[X = 1 / Y = 1] = \frac{P[X = 1, Y = 1]}{P[Y = 1]}$$

$$= \frac{5k}{15k} = \frac{5}{15}$$

$$P[X = 2 / Y = 1] = \frac{P[X = 2, Y = 1]}{P[Y = 1]}$$

$$= \frac{7k}{15k} = \frac{7}{15}$$

Conditional distribution of X given Y = 2:

$$P[X = x / Y = 2]$$

$$P[X = 0 / Y = 2] = \frac{P[X = 0, Y = 2]}{P[Y = 2]}$$

$$= \frac{6k}{24k} = \frac{6}{24}$$

$$P[X = 1 / Y = 2] = \frac{P[X = 1, Y = 2]}{P[Y = 2]}$$

$$= \frac{8k}{24k} = \frac{8}{24}$$

$$P[X = 2 / Y = 2] = \frac{P[X = 2, Y = 2]}{P[Y = 2]}$$

$$= \frac{10k}{24k} = \frac{10}{24}$$

Conditional distribution of X given Y = 3: $P\left[X = x / Y = 3\right]$

$$P\left[\ X = 0/Y = 3\ \right] = \frac{P\left[\ X = 0\ , Y = 3\ \right]}{P\left[Y = 3\ \right]} = \frac{9k}{33k} = \frac{9}{33}$$

$$P\left[\ X = 1/Y = 3\ \right] = \frac{P\left[\ X = 1\ , Y = 3\ \right]}{P\left[Y = 3\ \right]} = \frac{11k}{33k} = \frac{11}{33}$$

$$P\left[\ X = 2/Y = 3\ \right] = \frac{P\left[\ X = 2\ , Y = 3\ \right]}{P\left[Y = 3\ \right]} = \frac{13k}{33k} = \frac{13}{33}$$

2. **The two dimensional random variable (X, Y) has the joint density function**

$$f(x,y) = \frac{x+2y}{27},\ x = 0,1,2;\ \ y = 0,1,2.$$

Find the marginal distributions and conditional distributions.

To find marginal distributions:

X \ Y	0	1	2	$P(X = x)$
0	0	1/27	2/27	3/27
1	2/27	3/27	4/27	9/27
2	4/27	5/27	6/27	15/27
$P(Y = y)$	6/27	9/27	12/27	

The conditional distribution of Y given $X = x$ is given by $f(y/x) = \dfrac{f(x,y)}{f(x)}$

X \ Y	0	1	2
0	0	1/3	2/3
1	2/9	3/9	4/9
2	4/15	5/15	6/15

The conditional distribution of X given $Y=y$ is given by $f(x/y)=\dfrac{f(x,y)}{f(y)}$

Y / X	0	1	2
0	0	1/9	2/12
1	2/6	3/9	4/12
2	4/6	5/9	6/12

Example . Three balls are drawn at random without replacement from a box containing 2 white, 3 red and 4 black balls. If X denotes the number of white balls drawn and Y denotes the number of red balls drawn, find the joint probability distribution of (X,Y).

Here X takes the values 0, 1, 2 and values of Y are 0, 1, 2, 3. The joint $p.m.f$ of (X,Y) is required. There are $2W$, $3R$, $4B$ balls and in total 9 balls.

The probabilities when three balls are drawn at a time without replacement are given below:

$P(X=0,Y=0)=\dfrac{4C_3}{9C_3}=\dfrac{1}{21}$ $(W0,\ R0,\ B3)$	$P(X=1,Y=0)=\dfrac{2C_1\cdot 4C_2}{9C_3}=\dfrac{1}{7}$ $(W1,\ R0,\ B2)$	$P(X=2,Y=0)=\dfrac{2C_2\cdot 4C_1}{9C_3}=\dfrac{1}{21}$ $(W2,\ R0,\ B1)$
$P(X=0,Y=1)=\dfrac{3C_1\cdot 4C_2}{9C_3}=\dfrac{3}{14}$ $(W0,\ R1,\ B2)$	$P(X=1,Y=1)=\dfrac{2C_1\cdot 3C_1\cdot 4C_1}{9C_3}=\dfrac{2}{7}$ $(W1,\ R1,\ B1)$	$P(X=2,Y=1)=\dfrac{2C_2\cdot 3C_1}{9C_3}=\dfrac{1}{28}$ $(W2,\ R1,\ B0)$
$P(X=0,Y=2)=\dfrac{3C_2\cdot 4C_1}{9C_3}=\dfrac{1}{7}$ $(W0,\ R2,\ B1)$	$P(X=1,Y=2)=\dfrac{2C_1\cdot 3C_2}{9C_3}=\dfrac{1}{14}$ $(W1,\ R2,\ B0)$	$P(X=2,Y=2)=0$ $(W2,\ R2,\ B0)$
$P(X=0,Y=3)=\dfrac{3C_3}{9C_3}=\dfrac{1}{84}$ $(W0,\ R3,\ B0)$	$P(X=1,Y=3)=0$ $(W1,\ R3,\ B0)$	$P(X=2,Y=3)=0$ $(W2,\ R3,\ B0)$

Therefore the joint $p.m.f$ is

X / Y	0	1	2
0	$\frac{1}{21}$	$\frac{1}{7}$	$\frac{1}{21}$
1	$\frac{3}{14}$	$\frac{2}{7}$	$\frac{1}{28}$
2	$\frac{1}{7}$	$\frac{1}{14}$	0
3	$\frac{1}{84}$	0	0

3. **X and Y are independent random variables with variance 2 and 3. Find the variance of $3X+4Y$.**

Given $\text{Var}(X)=2$ and $\text{Var}(Y)=3$

We know $\text{Var}(X+Y) = \text{Var}(X)+\text{Var}(Y)+2\text{Cov}(X,Y)$ and $\text{Var}(aX)=a^2\ \text{Var}(X)$

Since X and Y are independent then $\text{Cov}(X,Y) = 0$

$\text{Var}(3X+4Y) = 9\text{Var}(X) + 16\text{Var}(Y) = 9(2) + 16(3) = 18 + 48 = 66$

Two Dimensional Continuous Random Variable

A function $f(x,y)$ is said to be joint pdf if $(i)\ f(x,y)\geq 0$ $(ii)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dxdy = 1$

Marginal distribution of X

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

Marginal distribution of Y

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

Conditional density function of X given Y

$$f(x/y)=\frac{f(x,y)}{f(y)}$$

Conditional density function of Y given X

$$f(y/x)=\frac{f(x,y)}{f(x)}$$

$$P[a<X<b]=\int_a^b f(x)dx$$

$$P[a<Y<b]=\int_a^b f(y)dy$$

$$P[a<X<b, c<Y<d]=\int_a^b\int_c^d f(x,y)dy\ dx$$

Conditional mean of X give $Y=a$

$$E[X/Y=a]=\int_{-\infty}^{\infty} x\ f(x/y)_{y=a}\ dx$$

Conditional mean of Y given $X=a$

$$E[Y/X=a]=\int_{-\infty}^{\infty} y\ f(y/x)_{x=a}\ dy$$

Properties of joint distribution function

(i) $0 \le F(x,y) \le 1$ (ii) $F(x,y)$ is non decreasing function

(iii) $P(a<x<b, Y \le y) = F(b,y) - F(a,y)$ (iv) $P(X \le x, a<y<b) = F(x,b) - F(x,a)$

(v) $P(a<X<b, c<Y<d) = F(b,d) - F(a,d) - F(b,c) + F(a,c)$

(vi) $F(x,-\infty) = F(-\infty,y) = 0,\ \ F(\infty,\infty) = 1$

(vii) At the point of continuity of $f(x,y)$, $f(x,y) = \dfrac{\partial^2}{\partial x \partial y} F(x,y)$

Example 1. If X and Y are two random variables having joint density function

$$f(x,y) = \begin{cases} \frac{1}{8}(6-x-y); & 0<x<2,\ 2<y<4 \\ 0 & ; \ otherwise \end{cases}.$$

Find $(i)\ P(X<1, Y<3)$ $(ii)\ P(X<1/Y<3)$ $(iii)\ P(X+Y<3)$

$$P[X<1\ ,\ Y<3] = P[\ 0<X<1\ ,\ 2<Y<3\ ]$$

$$= \int_0^1 \int_2^3 f(x,y)\ dy\ dx$$

$$= \frac{1}{8} \int_0^1 \int_2^3 (6-x-y)\ dy\ dx = \frac{1}{8} \int_0^1 \left(6y - xy - \frac{y^2}{2}\right)_2^3 dx$$

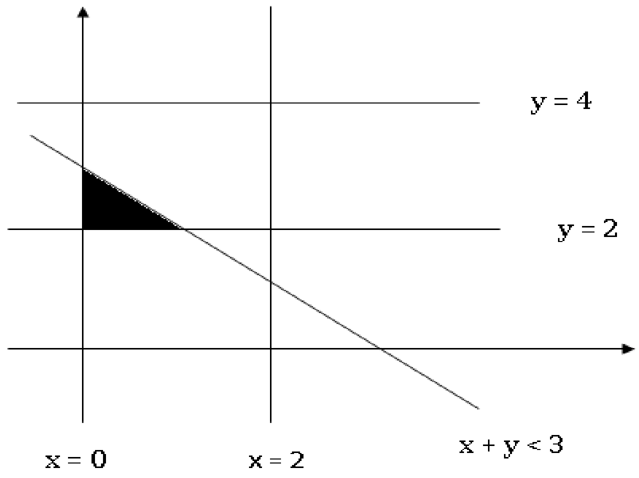
$$= \frac{1}{8} \int_0^1 \left(18 - 3x - \frac{9}{2}\right) - (12 - 2x - 2)\ dx$$

$$= \frac{1}{8} \int_0^1 \left(\frac{7}{2} - x\right) dx = \frac{1}{8} \left(\frac{7}{2}x - \frac{x^2}{2}\right)_0^1$$

$$= \frac{1}{8}\left(\frac{7}{2} - \frac{1}{2}\right) = \frac{3}{8}$$

Marginal density function of Y :	$P[\ Y<3\ ] = P[\ 2<Y<3\ ]$
$f(y) = \int_{-\infty}^{\infty} f(x,y)\ dx$	$= \int_2^3 f(y)\ dy$
$= \frac{1}{8} \int_0^1 (\ 6 - x - y)\ dx$	$= \frac{1}{8} \int_2^3 \left(\frac{11}{2} - y\right) dy$
$= \frac{1}{8} \left(6x - \frac{x^2}{2} - yx\right)_0^1$	$= \frac{1}{8}\left(\frac{11}{2}y - \frac{y^2}{2}\right)_2^3$
$= \frac{1}{8}\left(6 - \frac{1}{2} - y\right)$	$= \frac{1}{8}\left[\left(\frac{33}{2} - \frac{9}{2}\right) - (11-2)\right]$
$= \frac{1}{8}\left(\frac{11}{2} - y\right)$	$= \frac{3}{8}$

$$P[X<1/Y<3]=P[0<X<1/\ 2<Y<3]=\frac{P[0<X<1\ ,\ 2<Y<3]}{P[\ 2<Y<3]} \ =\ \frac{3}{8}\times\frac{8}{3}=1$$



$$P[\ X+\ Y<3\ ]\ =\ \int_0^1 \int_2^{3-x} f(x,y)\,dy\,dx$$

$$=\ \frac{1}{8}\int_0^1 \int_2^{3-x} (\ 6-x-y)dy\ dx\ \ =\ \frac{1}{8}\int_0^1 \left(6y-xy-\frac{y^2}{2}\right)_2^{3-x} dx$$

$$=\ \frac{1}{8}\int_0^1 \left(6(3-x)-x(3-x)-\frac{(3-x)^2}{2}\right)\ -(12-2x-2)\ dx$$

$$=\ \frac{1}{8}\int_0^1 (3-x)\ \left(6-x-\frac{1}{2}(3-x)\right)-10+2x\ \ dx$$

$$=\ \frac{1}{8}\int_0^1 18-3x-\frac{9}{2}+\frac{3x}{2}-6x+x^2+\frac{3x}{2}-\frac{x^2}{2}-10+2x\ dx$$

$$=\ \frac{1}{8}\int_0^1 \frac{7}{2}-4x+\frac{x^2}{2}\ dx$$

$$=\ \frac{1}{8}\left(\frac{7}{2}x-2x^2+\frac{x^3}{6}\ \right)_0^1$$

$$=\ \frac{5}{24}$$

Example . The joint $p.d.f$ of a random variable (X,Y) is $f(x,y)=xy^2+\frac{x^2}{8},\ 0\le x\le 2,\ 0\le y\le 1$

Find (i) $P(X<Y)$ (ii) $P(X<1)$ (iii) $P\left(Y<\frac{1}{2}\right)$ (iv) $P\left(X>1/Y<\frac{1}{2}\right)$.

Marginal $p.d.f$ of X

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

$$= \int_0^1 xy^2 + \frac{x^2}{8}\ dy$$

$$= \left(\frac{xy^3}{3} + \frac{x^2 y}{8}\right)_0^1$$

$$= \frac{x}{3} + \frac{x^2}{8}$$

Marginal $p.d.f$ of Y

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

$$= \int_0^2 xy^2 + \frac{x^2}{8}\ dx$$

$$= \left(\frac{x^2y^2}{2} + \frac{x^3}{24}\right)_0^2$$

$$= 2y^2 + \frac{1}{3}$$

$$P(X<1) = P(0<X<1)$$

$$= \int_0^1 f(x)\ dx$$

$$= \int_0^1 \frac{x}{3} + \frac{x^2}{8}\ dx$$

$$= \left(\frac{x^2}{6} + \frac{x^3}{24}\right)_0^1$$

$$= \frac{1}{6} + \frac{1}{24} = \frac{5}{24}$$

$$P\left(Y < \frac{1}{2}\right) = P\left(0 < Y < \frac{1}{2}\right)$$

$$= \int_0^{\frac{1}{2}} f(y)\ dy$$

$$= \int_0^{\frac{1}{2}} 2y^2 + \frac{1}{3}\ dy$$

$$= \left(\frac{2y^3}{3} + \frac{1}{3}y\right)_0^{\frac{1}{2}}$$

$$= \frac{1}{12} + \frac{1}{6} = \frac{1}{4}$$

.

$$P\left(1 < X < 2\ ,\ 0 < Y < \frac{1}{2}\right) = \int_1^2 \int_0^{\frac{1}{2}} f(x,y)\ dy\ dx$$

$$= \int_1^2 \int_0^{\frac{1}{2}} xy^2 + \frac{x^2}{8}\ dy\ dx$$

$$= \int_1^2 \left(x\frac{y^3}{3} + \frac{x^2 y}{8}\right)_0^{\frac{1}{2}}\ dx$$

$$= \int_1^2 \left(\frac{x}{24} + \frac{x^2}{16}\right)\ dx$$

$$= \left(\frac{x^2}{48} + \frac{x^3}{48}\right)_1^2$$

$$= \frac{4}{48} + \frac{8}{48} - \frac{1}{48} - \frac{1}{48}$$

$$= \frac{10}{48}$$

$$P(X<Y) = P(0<X<Y,\ 0<Y<1)$$

$$= \int_0^1 \int_0^y xy^2 + \frac{x^2}{8}\ dx\ dy$$

$$= \int_0^1 \left[\frac{x^2y^2}{2} + \frac{x^3}{24}\right]_0^y\ dy$$

$$= \int_0^1 \frac{y^4}{2} + \frac{y^3}{24}\ dy$$

$$= \left(\frac{y^5}{10} + \frac{y^4}{96}\right)_0^1$$

$$= \frac{1}{10} + \frac{1}{96}$$

$$= \frac{106}{960}$$

$$P\left(X > 1/Y < \frac{1}{2}\right) = P\left(1 < X < 2/0 < Y < \frac{1}{2}\right)$$

$$= \frac{P\left(1 < X < 2\ ,\ 0 < Y < \frac{1}{2}\right)}{P\left(0 < Y < \frac{1}{2}\right)}$$

$$= \frac{10}{48} \times \frac{4}{1} = \frac{5}{6}$$

Example 2. If the joint probability density function of (X, Y) is given by $f(x,y) = 2,\ 0 < x < y < 1$. Find i. Marginal density function of X and Y ii. Conditional density $f(x,y)$ iii. Conditional mean of X given $Y = 2$.

Marginal density function of X

$$\begin{aligned} f(x) &= \int_x^1 f(x,y)\,dy \\ &= \int_x^1 2\ dy \\ &= 2[y]_x^1 \\ &= 2(1-x) \end{aligned}$$

Marginal density function of Y

$$\begin{aligned} f(y) &= \int_0^y f(x,y)\,dx \\ &= \int_0^y 2\ dx \\ &= 2[x]_0^y \\ &= 2y \end{aligned}$$

Conditional density function of X given Y

$$f(x/y) = \frac{f(x,y)}{f(y)} = \frac{2}{2y} = \frac{1}{y}$$

Conditional mean of X given $Y = 2$

$$\begin{aligned} E[X/Y=2] &= \int_0^1 x\ f(x/y)_{y=2}\ dx \\ &= \int_0^1 x\ \frac{1}{2}\ dx \\ &= \frac{1}{2}\left[\frac{x^2}{2}\right]_0^1 \\ &= \frac{1}{4} \end{aligned}$$

Example . If the joint probability distribution function of a two dimensional random variable (X,Y) is given by $F(x,y)=\begin{cases}(1-e^{-x})(1-e^{-y}), & x>0,\ y>0\\ 0 & , \textit{otherwise}\end{cases}$. Find the marginal densities of X and Y. Are X and Y independent? Find $P(1<X<3, 1<Y<2)$.

Given the cumulative distribution function $F(x,y)$. Hence The joint $p.d.f$ is $f(x,y)=\frac{\partial^2}{\partial x\partial y}F(x,y)$

$$f(x,y)=\frac{\partial^2}{\partial x\partial y}(1-e^{-x})(1-e^{-y})$$

$$=\frac{\partial}{\partial x}(1-e^{-x})(e^{-y})$$

$$=(e^{-x})(e^{-y}),\ x>0,\ y>0$$

The marginal density function of X is

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

$$=\int_0^{\infty} e^{-x}e^{-y}\,dy$$

$$=e^{-x}\left(-e^{-y}\right)_0^{\infty}$$

$$=e^{-x}$$

The marginal density function of Y is

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

$$=\int_0^{\infty} e^{-x}e^{-y}\,dx$$

$$=e^{-y}\left(-e^{-x}\right)_0^{\infty}$$

$$=e^{-y}$$

Here $f_X(x)\cdot f_Y(y)=f(x,y)$. Hence X and Y are independent.

Also $P(1<X<3, 1<Y<2)=\int_1^3\int_1^2 f(x,y)\,dy\,dx$

$$=\int_1^3\int_1^2 e^{-x}e^{-y}\,dy\,dx$$

$$=\int_1^3 e^{-x}\,dx\int_1^2 e^{-y}\,dy$$

$$=\left(-e^{-x}\right)_1^3\left(-e^{-y}\right)_1^2$$

$$=\left(-e^{-3}+e^{-1}\right)\left(-e^{-2}+e^{-1}\right)$$

Example . Given the joint $p.d.f$ of X and Y is $f_{X,Y}(x,y)=\begin{vmatrix} Cx(x-y); & 0<x<2, -x<y<x \\ 0 & ; \textit{otherwise} \end{vmatrix}$

(1) Evaluate C (2) Find marginal pdf of X (3) Find the conditional density of Y/X

By definition of $p.d.f$, we have $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\, dy\, dx = 1$

$$C\int_0^2\int_{-x}^{x} x^2 - xy\ dy\ dx = 1$$

$$2C\int_0^2\int_0^{x} x^2\ dy\ dx = 1 \quad \text{\{applying odd/even integral property\}}$$

$$2C\int_0^2 \left(x^2 y\right)_0^x dx = 1$$

$$2C\int_0^2 x^3\ dx = 1$$

$$2C\left(\frac{x^4}{4}\right)_0^2 = 1$$

$$8C = 1$$

$$C = \frac{1}{8}$$

Marginal distribution of X is $f_X(x) = \int_{-\infty}^{\infty} f(x,y)\ dy$

$$f_X(x) = \frac{1}{8}\int_{-x}^{x} x^2 - xy\ dy$$

$$= \frac{2}{8}\int_0^x x^2\ dy \qquad \text{\{applying odd/even integral property\}}$$

$$= \frac{2}{8}\left(x^2 y\right)_0^x$$

$$= \frac{x^3}{4},\ 0<x<2$$

Conditional density function of Y/X is

$$f_{Y/X}(y/x) = \frac{f(x,y)}{f(x)} = \frac{\frac{1}{8}x(x-y)}{\frac{x^3}{4}} = \frac{1}{2x^2}(x-y),\ -x<y<x$$

Example 3. The joint probability density function of the random variables (X,Y) is given by $f(x,y)=kxye^{-(x^2+y^2)},\ x>0,\ y>0$. Find the value of k and also prove that X and Y are independent.

Since $f(x,y)$ is a joint $p.d.f$, we have $\int_0^\infty \int_0^\infty kxye^{-(x^2+y^2)}\,dxdy=1$

$$k\int_0^\infty xe^{-x^2}\,dx\int_0^\infty ye^{-y^2}\,dy=1$$

$$\frac{k}{2\times 2}\int_0^\infty e^{-x^2}\,d(x^2)\int_0^\infty e^{-y^2}\,d(y^2)=1$$

$$\frac{k}{4}\left(e^{-x^2}\right)_0^\infty\left(e^{-y^2}\right)_0^\infty=1$$

$$\frac{k}{4}(0-1)(0-1)=1$$

$$k=4$$

Marginal density function of X	Marginal density function of Y
$f_X(x)=\int_0^\infty 4xye^{-(x^2+y^2)}\,dy$	$f_Y(y)=\int_0^\infty 4xye^{-(x^2+y^2)}\,dx$
$=4xe^{-x^2}\int_0^\infty ye^{-y^2}\,dy$	$=4ye^{-y^2}\int_0^\infty xe^{-x^2}\,dx$
$=\frac{4}{2}xe^{-x^2}\int_0^\infty e^{-y^2}\,d(y^2)$	$=\frac{4}{2}ye^{-y^2}\int_0^\infty e^{-x^2}\,d(x^2)$
$=-2xe^{-x^2}$	$=-2ye^{-y^2}$

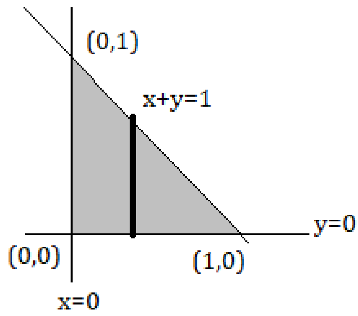
Since $f(x,y)=f_X(x)\times f_Y(y)$, X and Y are independent.

Definition: The correlation between two random variables X and Y is defined as $E(XY)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\,f(x,y)\,dx\,dy$.

- Two random variables X and Y are uncorrelated if $E(XY)=E(X)E(Y)$
- Two random variables X and Y are orthogonal if $E(XY)=0$
- Covariance between X and Y is $Cov(X,Y)=E(XY)-E(X)E(Y)$
- Correlation coefficient between X and Y is $r=\dfrac{Cov(X,Y)}{\sigma_X\cdot\sigma_Y}$
- Correlation coefficient lies between -1 and 1.
- If two variables are independent, then they are uncorrelated and covariance is zero. But the converse need not be true.

Example . The joint probability density function of the random variables (X,Y) is $f(x,y)=3(x+y),\ 0\le x\le 1,\ 0\le y\le 1,\ x+y\le 1$. Find $Cov(X,Y)$.

Marginal density function of X and Y



$$f_X(x)=\int_{-\infty}^{\infty} f(x,y)\,dy$$
$$=3\int_0^{1-x} x+y\,dy$$
$$=3\left(xy+\frac{y^2}{2}\right)_0^{1-x}$$
$$=3\left(x(1-x)+\frac{1}{2}(1-x)^2\right)$$
$$=\frac{3}{2}\left(1-x^2\right),\ 0\le x\le 1$$

Similarly $f_Y(y)=\int_{-\infty}^{\infty} f(x,y)\,dx\ =\frac{3}{2}\left(1-y^2\right),\ 0\le y\le 1$

$$E(X)=\int_{-\infty}^{\infty} xf(x)\,dx$$
$$=\frac{3}{2}\int_0^1 x\left(1-x^2\right)dx$$
$$=\frac{3}{2}\left(\frac{x^2}{2}-\frac{x^4}{4}\right)_0^1$$
$$=\frac{3}{2}\left(\frac{1}{2}-\frac{1}{4}\right)$$
$$=\frac{3}{8}$$

Similarly $E(Y)=\frac{3}{8}$

$$E(XY)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)\,dx$$
$$=3\int_0^1\int_0^{1-x} xy(x+y)\,dy\,dx$$
$$=3\int_0^1\left(\frac{x^2y^2}{2}+\frac{xy^3}{3}\right)_0^{1-x}dx$$
$$=3\int_0^1\left(\frac{x^2(1-x)^2}{2}+\frac{x(1-x)^3}{3}\right)dx$$
$$=3\int_0^1\left(\frac{x^2+x^4-2x^3}{2}+\frac{x+3x^3-3x^2-x^4}{3}\right)dx$$
$$=3\left[\frac{1}{2}\left(\frac{x^3}{3}+\frac{x^5}{5}-\frac{x^4}{2}\right)+\frac{1}{3}\left(\frac{x^2}{2}+\frac{3x^4}{4}-\frac{3x^3}{3}-\frac{x^5}{5}\right)\right]_0^1$$
$$=3\left[\frac{1}{2}\left(\frac{1}{3}+\frac{1}{5}-\frac{1}{2}\right)+\frac{1}{3}\left(\frac{1}{2}+\frac{3}{4}-\frac{3}{3}-\frac{1}{5}\right)\right]$$
$$=\frac{1}{10}$$

$$Cov(X,Y)=E(XY)-E(X)\cdot E(Y)$$
$$=\frac{1}{10}-\frac{3}{8}\cdot\frac{3}{8}$$
$$=-\frac{13}{320}$$

Example: Two random variables X and Y have joint density function $f(x,y)=x^2+\frac{xy}{3},\ 0\le x\le 1,\ 0\le y\le 2$. Are X and Y independent? Find the conditional density functions and check whether they are valid.

Marginal density function of X

$$f_X(x)=\int_{-\infty}^{\infty} f(x,y)\,dy$$
$$=\int_0^2 x^2+\frac{xy}{3}\,dy$$
$$=\left[x^2y+\frac{xy^2}{6}\right]_0^2$$
$$=2x^2+\frac{2}{3}x,\ \ 0\le x\le 1$$

Marginal density function of Y

$$f_X(x)=\int_{-\infty}^{\infty} f(x,y)\,dx$$
$$=\int_0^1 x^2+\frac{xy}{3}\,dx$$
$$=\left[\frac{x^3}{3}+\frac{x^2y}{6}\right]_0^1$$
$$=\frac{1}{3}+\frac{y}{6},\quad 0\le y\le 2$$

Since $f_X(x)\ne f_Y(y)$, X and Y are not independent.

Conditional density of X given Y

$$f_X(X/Y)=\frac{f_{XY}(x,y)}{f_Y(y)}$$
$$=\frac{x^2+\frac{xy}{3}}{\frac{1}{3}+\frac{y}{6}}$$
$$=\frac{6x^2+2xy}{2+y},\ \ 0\le x\le 1,\ \ 0\le y\le 2$$

Conditional density of Y given X

$$f_Y(Y/X)=\frac{f_{XY}(x,y)}{f_X(x)}$$
$$=\frac{x^2+\frac{xy}{3}}{2x^2+\frac{2}{3}x}$$
$$=\frac{3x+y}{6x+2},\ \ 0\le x\le 1,\ \ 0\le y\le 2$$

Validity of conditional densities

$$\int_0^1 f_X(X/Y)\,dx=\int_0^1 \frac{6x^2+2xy}{2+y}\,dx$$
$$=\frac{1}{2+y}\left[\frac{6x^3}{3}+\frac{2x^2y}{2}\right]_0^1$$
$$=\frac{1}{2+y}[2+y]$$
$$=1$$

$$\int_0^2 f_Y(Y/X)\,dy=\int_0^2 \frac{3x+y}{6x+2}\,dy$$
$$=\frac{1}{6x+2}\left[3xy+\frac{y^2}{2}\right]_0^2$$
$$=\frac{1}{6x+2}[6x+2]$$
$$=1$$

Hence the conditional densities are valid.

Example . Two random variables X and Y are defined as $Y = 4X + 9$. Find the correlation coefficient between X and Y.

$$E(Y) = E(4X+9) = 4E(X)+9 \qquad E(X)E(Y) = 4[E(X)]^2 + 9E(X)$$

$$E(XY) = E[X(4X+9)] = E[(4X^2+9X)] = 4E(X^2)+9E(X)$$

$$E(Y^2) = E(4X+9)^2 = E(16X^2+81+72X) = 16E(X^2)+72E(X)+81$$

$$[E(Y)]^2 = [4E(X)+9]^2 = 16[E(X)]^2 + 81 + 72E(X)$$

$$\sigma_Y^2 = E(Y^2) - [E(Y)]^2$$

$$= 16E(X^2)+72E(X)+81-16[E(X)]^2-81-72E(X)$$

$$= 16E(X^2)-16[E(X)]^2$$

$$= 16\sigma_X^2$$

$$E(XY)-E(X)E(Y) = 4E(X^2)+9E(X)-4[E(X)]^2-9E(X) = 4\sigma_X^2$$

The correlation coefficient is $r = \dfrac{Cov(X,Y)}{\sigma_X \cdot \sigma_Y}$

$$= \frac{E(XY)-E(X)E(Y)}{\sigma_X \cdot \sigma_Y}$$

$$= \frac{4\sigma_X^2}{\sigma_X \cdot 4\sigma_X}$$

$$= 1$$

CORRELATION AND REGRESSION

Two or more variables are interrelated in many situations such as business , industry , agriculture and so on. Two variables are correlated if a change in one variable affects the other variable. Otherwise the two variables are called uncorrelated variables. In particular two variables are said to be positively correlated if one variable increases (or decreases) as the other variable increases (or decreases). Two variables are said to be negatively correlated if one variable increases (or decreases) as the other variable decreases (or increases). There are two measures of correlation coefficient between two variables. They are (a) Karl Pearson's coefficient of correlation (b) Spearman's coefficient of rank correlation.

KARL PEARSON'S COEFFICIENT OF CORRELATION

The Karl Pearson's coefficient of correlation between two variables X and Y is usually denoted by $r(x,y)$ and is defined by

$$r(x,y)=\frac{\frac{1}{n}\sum(x-\overline{x})(y-\overline{y})}{\sqrt{\frac{1}{n}\sum(x-\overline{x})^2}\sqrt{\frac{1}{n}\sum(y-\overline{y})^2}}=\frac{\operatorname{cov}(X,Y)}{\sigma_X\sigma_Y}$$ were $x_1, x_2, ..., x_n$ and $y_1, y_2, ..., y_n$ are the values of the random variables X and Y respectively and $\overline{x}$, $\overline{y}$ are the mean values of X and Y respectively.

NOTE :

The above formula is useful if $\overline{x}$ and $\overline{y}$ are integers. The following formulae are useful when $\overline{x}$ and $\overline{y}$ are not integers and the given observations are large.

DIFFERENT FORMULAE FOR COEFFICIENT OF CORRELATION

(A) The coefficient of correlation between two variables X and Y is given by

$$r(x,y)=\frac{\frac{1}{n}\sum xy-\overline{x}\,\overline{y}}{\sqrt{\frac{1}{n}\sum x^2-\overline{x}^2}\sqrt{\frac{1}{n}\sum y^2-\overline{y}^2}}$$

(B) The coefficient of correlation between two variables X and Y is given by

$$r(x,y)=\frac{n\sum xy-\sum x\sum y}{\sqrt{n\sum x^2-\left(\sum x\right)^2}\sqrt{n\sum y^2-\left(\sum y\right)^2}}$$

NOTE :

Two variables X and Y are uncorrelated if $r(x,y)=0$.

Two variables X and Y are uncorrelated if $\operatorname{cov}(X,Y)=0$.

PROPERTIES OF CORRELATION COEFFICIENT

1. The coefficient of correlation $r(x,y)$ lies between – 1 and 1.

NOTE :

(i) If $r=1$, then the coefficient of correlation between two variables is perfect.

(ii) If $r=-1$, then the inverse of coefficient of correlation between two variables is perfect.

(iii) If $r=0$, then the two variables are independent (or uncorrelated). i.e There is no relation between two variables.

2. The coefficient of correlation is independent of change of scale and origin of the variables X and Y.

3. If X and Y are random variables , then $r(x,y)=\frac{\sigma_X^2+\sigma_Y^2-\sigma_{X-Y}^2}{2\sigma_X\,\sigma_Y}$

Example 1. If X and Y are uncorrelated random variables with variances 16 and 9. Find the correlation coefficient between $U = X+Y$ and $V = X-Y$.

Since X and Y are uncorrelated, $Cov(X,Y) = 0$.

Given that $V(X) = \sigma_X^2 = 16$ and $V(Y) = \sigma_Y^2 = 9$

i.e. $E(X)^2 - (E(X))^2 = 16$ and $E(Y)^2 - (E(Y))^2 = 9$

$$E(U) = E(X+Y) = E(X) + E(Y)$$

$$\begin{aligned} E(U^2) &= E(X+Y)^2 \\ &= E(X^2 + Y^2 + 2XY) \\ &= E(X^2) + E(Y^2) + 2E(X)E(Y) \end{aligned}$$

$$\begin{aligned} (E(U))^2 &= [E(X) + E(Y)]^2 \\ &= (E(X))^2 + (E(Y))^2 + 2E(X)E(Y) \end{aligned}$$

$$\begin{aligned} E(U^2) - (E(U))^2 &= E(X^2) + E(Y^2) + 2E(X)E(Y) \\ &\quad - (E(X))^2 - (E(Y))^2 - 2E(X)E(Y) \\ &= \sigma_X{}^2 + \sigma_Y{}^2 \end{aligned}$$

$$E(V) = E(X-Y) = E(X) - E(Y)$$

$$\begin{aligned} E(V^2) &= E(X-Y)^2 \\ &= E(X^2 + Y^2 - 2XY) \\ &= E(X^2) + E(Y^2) - 2E(X)E(Y) \end{aligned}$$

$$\begin{aligned} (E(V))^2 &= [E(X) - E(Y)]^2 \\ &= (E(X))^2 + (E(Y))^2 - 2E(X)E(Y) \end{aligned}$$

$$\begin{aligned} E(V^2) &= E(X^2) + E(Y^2) - 2E(X)E(Y) \\ &\quad - (E(X))^2 - (E(Y))^2 + 2E(X)E(Y) \\ &= \sigma_X{}^2 + \sigma_Y{}^2 \end{aligned}$$

$$\begin{aligned} E(UV) &= E[(X+Y)(X-Y)] \\ &= E(X^2 - Y^2) \\ &= E(X^2) - E(Y^2) \end{aligned}$$

$$\begin{aligned} E(U)E(V) &= (E(X) + E(Y))(E(X) - E(Y)) \\ &= (E(X))^2 - (E(Y))^2 \end{aligned}$$

$$\begin{aligned} E(UV) - E(U)E(V) &= E(X^2) - E(Y^2) - (E(X))^2 + (E(Y))^2 \\ &= \sigma_X{}^2 - \sigma_Y{}^2 \end{aligned}$$

$$r_{UV} = \frac{E(UV) - E(U)E(V)}{\sqrt{E(U)^2 - (E(U))^2}\sqrt{E(V)^2 - (E(V))^2}}$$

$$= \frac{\sigma_X{}^2 - \sigma_Y{}^2}{\sqrt{\sigma_X{}^2 + \sigma_Y{}^2}\sqrt{\sigma_X{}^2 + \sigma_Y{}^2}}$$

$$= \frac{\sigma_X{}^2 - \sigma_Y{}^2}{\sigma_X{}^2 + \sigma_Y{}^2}$$

$$= \frac{16-9}{16+9}$$

$$= \frac{7}{25}$$

Example . Two independent random variables X and Y are defined by $f_X(x)=\begin{vmatrix}4ax, & 0<x<1\\ 0, & otherwise\end{vmatrix}$

and $f_Y(y)=\begin{vmatrix}4by, & 0<y<1\\ 0, & otherwise\end{vmatrix}$. Show that $U=X+Y$ and $V=X-Y$ are uncorrelated.

Since X and Y are independent, $E(XY)=E(X)E(Y)$.

By the definition of $p.d.f$

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

$$\int_0^1 4a\,x\;dx=1$$

$$4a\left(\frac{x^2}{2}\right)_0^1=1$$

$$4a\left(\frac{1}{2}\right)=1$$

$$a=\frac{1}{2}$$

$$E(X)=\int_0^1 x\;f(x)\;dx$$

$$=\int_0^1 2x^2\;dx$$

$$=2\left(\frac{x^3}{3}\right)_0^1$$

$$=\frac{2}{3}$$

Similarly $E(Y)=\frac{2}{3}$

By the definition of $p.d.f$

$$\int_{-\infty}^{\infty} f_Y(y)dy=1$$

$$\int_0^1 4b\,y\;dy=1$$

$$4b\left(\frac{y^2}{2}\right)_0^1=1$$

$$4b\left(\frac{1}{2}\right)=1$$

$$b=\frac{1}{2}$$

$$E(X^2)=\int_0^1 x^2\;f(x)\;dx$$

$$=\int_0^1 2x^3\;dx$$

$$=2\left(\frac{x^4}{4}\right)_0^1$$

$$=\frac{1}{2}$$

Similarly $E(Y^2)=\frac{1}{2}$

$$\sigma_X{}^2=Var(X)=E(X^2)-(E(X))^2$$

$$=\frac{1}{2}-\frac{4}{9}=\frac{1}{18}$$

$$\sigma_Y{}^2=Var(Y)=E(Y^2)-(E(Y))^2$$

$$=\frac{1}{2}-\frac{4}{9}=\frac{1}{18}$$

If U and V are uncorrelated, then correlation coefficient $r=0$

$$i.e.\quad Cov(U,V)=0$$

$$i.e.\quad E(UV)-E(U)E(V)=0$$

$$i.e.\quad E((X+Y)(X-Y))-E(X+Y)E(X-Y)=0$$

$$i.e.\quad E(X^2-Y^2)-(E(X)+E(Y))(E(X)-E(Y))=0$$

$$i.e.\quad E(X^2)-E(Y^2)-(E(X))^2+E(X)E(Y)-E(Y)E(X)+(E(Y))^2=0$$

i.e. $\left[E\left(X^2\right)-\left(E\left(X\right)\right)^2\right]-\left[E\left(Y^2\right)-\left(E\left(Y\right)\right)^2\right]=0$

$$\left[\frac{1}{18}\right]-\left[\frac{1}{18}\right]=0$$

Example . Find the coefficient of correlation between X and Y from the data given below.

X :	65	66	67	67	68	69	70	72
Y :	67	68	65	68	72	72	69	71

X	Y	X^2	Y^2	XY
65	67	4225	4489	4355
66	68	4356	4624	4488
67	65	4489	4225	4355
67	68	4489	4624	4556
68	72	4624	5184	4896
69	72	4761	5184	4968
70	69	4900	4761	4830
72	71	5184	5041	5112
544	**552**	**37028**	**38132**	**37560**

$$\bar{X}=\frac{\sum X}{n}=\frac{544}{8}=68 \quad \text{and} \quad \sigma_X=\sqrt{\left(\frac{\sum X^2}{n}\right)-\left(\frac{\sum X}{n}\right)^2}=\sqrt{\left(\frac{37028}{8}\right)-(68)^2}=2.121$$

$$\bar{Y}=\frac{\sum Y}{n}=\frac{552}{8}=69 \quad \text{and} \quad \sigma_Y=\sqrt{\left(\frac{\sum Y^2}{n}\right)-\left(\frac{\sum Y}{n}\right)^2}=\sqrt{\left(\frac{38132}{8}\right)-(69)^2}=2.345$$

$$Cov\left(X,Y\right)=\frac{\sum XY}{n}-\frac{\sum X}{n}\cdot\frac{\sum Y}{n}=\frac{37560}{8}-68\times 69=3$$

Coefficient of correlation $r=\dfrac{Cov\left(X,Y\right)}{\sigma_X.\sigma_Y}=\dfrac{3}{2.121\times 2.345}=0.603$

EXAMPLE : The following table gives the number of blind per lakh of population in different age groups. Find out the correlation between age and blindness.

Age in years	0 – 10	10–20	20 – 30	30 – 40	40 – 50	50 – 60	60 – 70	70 - 80
Number of Blind	55	67	100	111	150	200	300	500

Since the grouped distribution is given , we use the idea of change of scale and origin. Take $u_i = \frac{x_i - 45}{10}$ and $v_i = y_i - 200$

The coefficient of correlation between X and Y is given by

$$r(x,y) = r(u,v) = \frac{\frac{1}{n}\sum uv - \bar{u}\,\bar{v}}{\sqrt{\frac{1}{n}\sum u^2 - \bar{u}^2}\sqrt{\frac{1}{n}\sum v^2 - \bar{v}^2}} \text{ .... } (1)$$

Age	Mid value x_i	No. of blind y_i	$u_i = \frac{x_i - 45}{10}$	$v_i = y_i - 200$	$u_i v_i$	u_i^2	v_i^2
0 – 10	5	55	– 4	- 145	580	16	21025
10 – 20	15	67	- 3	- 133	399	9	17689
20 – 30	25	100	- 2	- 100	200	4	10000
30 – 40	35	111	- 1	- 89	89	1	7921
40 – 50	45	150	0	- 50	0	0	2500
50 – 60	55	200	1	0	0	1	0
60 – 70	65	300	2	100	200	4	10000
70 – 80	75	500	3	300	900	9	90000
			- 4	- 117	2368	44	159135

Then the mean of U = $\frac{1}{n}\sum_{i=1}^{n} u_i$ i.e. $\bar{u} = \frac{-4}{8} = -0.5$

Also the mean of V = $\frac{1}{n}\sum_{i=1}^{n} v_i$ i.e. $\bar{v} = \frac{-117}{8} = -14.625$

$$\text{Hence } r(x,y) = \frac{\frac{1}{8}(2368) - (-0.5)(-14.625)}{\sqrt{\frac{1}{8}(44) - (-0.5)^2}\sqrt{\frac{1}{8}(159135) - (-14.625)^2}} = 0.89$$

Since the coefficient of correlation $r(x,y)$ is positive, the age and blindness are positively correlated.

RANK CORRELATION

Let $x_1, x_2, \ldots, x_n$ and $y_1, y_2, \ldots, y_n$ are the values of the random variables X and Y respectively . Then the Spearman's rank correlation coefficient between two variables X and Y is usually denoted by $\rho(x,y)$ and is defined by

$$\rho(x,y)=1-\left[\frac{6\sum_{i=1}^{n} d_i^2}{n\left(n^2-1\right)}\right]$$ where n is the number of observations and

d_i = rank of x_i – rank of y_i

Note: In the formula add the correction factor $\frac{m\left(m^2-1\right)}{12}$ to $\sum d_i^2$, where m is the number of times an item is repeated.

EXAMPLE : Distribution of marks in Economics and mathematics for ten students in a certain test is given below.

Student Number	1	2	3	4	5	6	7	8	9	10
Marks in Economics	25	28	32	36	40	38	39	42	41	45
Marks in Mathematics	70	80	85	75	65	59	48	50	54	66

Calculate the value of rank correlation coefficient.

Student number	Marks in Economics x_i	Marks in Mathematics y_i	Rank of x_i (R_1)	Rank of y_i (R_2)	$d_i = R_1 - R_2$	d_i^2
1	25	70	10	4	6	36
2	28	80	9	2	7	49
3	32	85	8	1	7	49
4	36	75	7	3	4	16
5	40	65	4	6	- 2	4
6	38	59	6	7	- 1	1
7	39	48	5	10	- 5	25
8	42	50	2	9	- 7	49
9	41	54	3	8	- 5	25
10	45	66	1	5	- 4	16
						270

Then the Spearman's rank correlation coefficient between two variables X and Y is

$$\rho(x,y)=1-\left[\frac{6\sum_{i=1}^{n} d_i^2}{n\left(n^2-1\right)}\right]=-0.6364$$

Hence the Spearman's rank correlation coefficient is $\rho(x,y) = -0.6364$.

EXAMPLE : From the following data , calculate the coefficient of rank correlation between X and Y.

X	32	55	49	60	43	37	43	49	10	20
Y	40	30	70	20	30	50	72	60	45	25

X	Y	Rank of X (R_1)	Rank of Y (R_2)	$d_i = R_1 - R_2$	d_i^2
32	40	8	6	2	4
55	30	2	7.5	- 5.5	30.25
49	70	3.5	2	1.5	2.25
60	20	1	10	- 9	81
43	30	5.5	7.5	- 2	4
37	50	7	4	3	9
43	72	5.5	1	4.5	20.25
49	60	3.5	3	0.5	0.25
10	45	10	5	5	25
20	25	9	9	0	0
					176

Since some ranks are repeated , we have

The correction factor C.F = $\sum_{m} \dfrac{m\left(m^2 - 1\right)}{12}$ where m is the number of times a rank repeated.

Therefore C.F = $\dfrac{2\left(2^2 - 1\right)}{12} + \dfrac{2\left(2^2 - 1\right)}{12} + \dfrac{2\left(2^2 - 1\right)}{12} = 1.5$

Hence corrected $\sum_{i=1}^{n} d_i^2$ = Correction factor + actual $\sum_{i=1}^{n} d_i^2$

$= 1.5 + 176$

$= 177.5$

Then the Spearman's rank correlation coefficient between two variables X and Y is

$$\rho(x,y)=1-\left[\frac{6\sum_{i=1}^{n} d_i^2}{n\left(n^2-1\right)}\right]=-0.0758$$

Hence the Spearman's rank correlation coefficient is $\rho(x,y) = -0.0758$.

EXAMPLE : Ten competitors in a beauty contest are ranked by three judges in the following order.

First judge	1	6	5	10	3	2	4	9	7	8
Second Judge	3	5	8	4	7	10	2	1	6	9
Third Judge	6	4	9	8	1	2	3	10	5	7

Use the method of rank correlation to determine which pair of judges have the nearest approach to common likings in beauty.

Ranks by the Judges			d_1	d_2	d_3	d_1^2	d_2^2	d_3^2
First (R_1)	Second (R_2)	Third (R_3)	$R_1 - R_2$	$R_2 - R_3$	$R_1 - R_3$			
1	3	6	-2	-3	-5	4	9	25
6	5	4	1	1	2	1	1	4
5	8	9	-3	-1	-4	9	1	16
10	4	8	6	-4	2	36	16	4
3	7	1	-4	6	2	16	36	4
2	10	2	-8	8	0	64	64	0
4	2	3	2	-1	1	4	1	1
9	1	10	8	-9	-1	64	81	1
7	6	5	1	1	2	1	1	4
8	9	7	-1	2	1	1	4	1
						200	214	60

Then the Spearman's rank correlation coefficient between first and second judges is

$$\rho_{1,2}(x,y)=1-\left[\frac{6\sum_{i=1}^{n} d_i^2}{n\left(n^2-1\right)}\right]=-0.2121$$

Then the Spearman's rank correlation coefficient between second and third judges is

$$\rho_{2,3}(x,y)=1-\left[\frac{6\sum_{i=1}^{n} d_i^2}{n\left(n^2-1\right)}\right]=-0.297$$

Then the Spearman's rank correlation coefficient between first and third judges is

$$\rho_{1,3}(x,y)=1-\left[\frac{6\sum_{i=1}^{n} d_i^2}{n\left(n^2-1\right)}\right]=0.6363$$

Since the Spearman's rank correlation coefficient between first and third judges is positive and maximum, we conclude that the pair of judges 1 and 3 has the nearest approach to common likings in beauty.

Regression Analysis

Regression shows a relationship between the average values of two variables. Thus regression is helpful in estimating the average value of one variable for a given value of the other variable.

The best average value of one variable associated with the given value of the other variable may also be estimated by means of an equation, known as regression equation.

Regression line of x on y. It gives the estimate of the value of x for a specified value of y.

$$(x-\bar{x})=r\frac{\sigma_x}{\sigma_y}(y-\bar{y})$$

Regression line of y on x. It gives the estimate of the value of y for a specified value of x.

$$(y-\bar{y})=r\frac{\sigma_y}{\sigma_x}(x-\bar{x}) \text{ where}$$

$$r=\frac{n\sum xy-\sum x\sum y}{\sqrt{n\sum x^2-\left(\sum x\right)^2}\sqrt{n\sum y^2-\left(\sum y\right)^2}},\ \sigma_x=\sqrt{\sum x^2-\frac{\left(\sum x\right)^2}{n}} \text{ and } \sigma_y=\sqrt{\sum y^2-\frac{\left(\sum y\right)^2}{n}}.$$

Note: $r=\dfrac{\operatorname{cov}(x,y)}{\sigma_x\sigma_y}$

Regression coefficient of x on y is $b_{xy}=r\dfrac{\sigma_x}{\sigma_y}=\dfrac{\sum xy-n\bar{x}\,\bar{y}}{\sum y^2-n(\bar{y})^2}=\dfrac{\sum(x-\bar{x})(y-\bar{y})}{\sum(y-\bar{y})^2}$.

Regression coefficient of y on x is $b_{yx}=r\dfrac{\sigma_y}{\sigma_x}=\dfrac{\sum xy-n\bar{x}\,\bar{y}}{\sum x^2-n(\bar{x})^2}=\dfrac{\sum(x-\bar{x})(y-\bar{y})}{\sum(x-\bar{x})^2}$.

Let A and B respectively be the assumed mean for x and y. Let $dx=x-A$ and $dy=y-B$.

Then $b_{xy}=r\dfrac{\sigma_x}{\sigma_y}=\dfrac{\sum dxdy-\dfrac{\sum dx\cdot\sum dy}{n}}{\sum dy^2-\dfrac{\left(\sum dy\right)^2}{n}}$ and $b_{yx}=r\dfrac{\sigma_y}{\sigma_x}=\dfrac{\sum dxdy-\dfrac{\sum dx\cdot\sum dy}{n}}{\sum dx^2-\dfrac{\left(\sum dx\right)^2}{n}}$.

Properties of Regression Coefficient:

1. The correlation coefficient is the geometric mean of the regression coefficients.

$$i.e.\ \ r = \sqrt{b_{xy}\ .\ b_{yx}}$$

2. If one of the regression coefficients is greater than unity, then the other is less than unity.
3. Both the regression coefficients will have the same sign.
4. The sign of correlation coefficient is same as that of regression coefficients.

Properties of linear regressions:

1. Two regression lines always intersect at the means.
2. If $r = 0$, then the regression coefficients are zero and the regression lines are perpendicular.
3. If $r = \pm 1$, the regression lines becomes identical.

If θ is the angle between two regression lines , then $\tan\theta = \left(\frac{1-r^2}{r}\right)\frac{\sigma_x\sigma_y}{\sigma_x{}^2+\sigma_y{}^2}$.

Example . The equations of two regression lines are $3x+12y=19$ and $9x+3y=46$. Find $\overline{x}$, $\overline{y}$ and the correlation coefficient between X and Y .

Regression lines are passing through $(\overline{x}, \overline{y})$. Therefore

$$3\overline{x}+12\overline{y}=19 \text{ ....(1)}$$
$$9\overline{x}+3\overline{y}=46 \text{ ....(2)}$$

$(1)\times 3$ gives, $\quad 9\overline{x}+36\overline{y}=57$ ....(3)

$(3)-(2)$ gives, $\quad 33\overline{y}=11$ and hence $\overline{y}=\frac{1}{3}$

From (1), $3\overline{x}+4=19$ and hence $\overline{x}=5$.

Let the regression line of Y on X be $3x+12y=19$ i.e. $x=-\frac{3}{12}x+\frac{19}{12}$

$$\text{Hence } b_{yx} = -\frac{3}{12}$$

Let the regression line of X on Y be $9x+3y=46$ i.e. $x=\frac{46}{9}-\frac{3}{9}y$

$$\text{Hence } b_{xy} = -\frac{3}{9}$$

But $r^2 = b_{yx}\times b_{xy} = \left(-\frac{3}{9}\right)\times\left(-\frac{3}{12}\right)=\frac{1}{12}$. Therefore $r=-\frac{1}{2\sqrt{3}}$

Example: Given that the variance of $x=9$ **and the regression equations are** $8x-10y+66=0,\ 40x-18y=214.$ **Find (a) mean values of** x **and** y **(b) coefficient of correlation between** x **and** y **(c) standard deviation of** y **.**

We know that the regression equations are passing through $(\overline{x},\ \overline{y})$. Therefore

$$8\overline{x}-10\overline{y}=-66.$$

$$40\bar{x} - 18\bar{y} = 214.$$

Solving the above equations, we get $\bar{x} = 13, \quad \bar{y} = 17.$

Let us assume that the line $8x - 10y + 66 = 0$ be regression equation of y on x.

Therefore

$$y = \frac{8}{10}x + \frac{66}{10}$$

$$i.e.\ b_{yx} = \frac{8}{10}$$

Let us assume that the line $40x - 18y = 214$ be regression equation of x on y.

Therefore

$$x = \frac{18}{40}y + \frac{214}{48}$$

$$i.e.\ b_{xy} = \frac{18}{40}$$

But $r = \sqrt{b_{xy}\ .\ b_{yx}} = \sqrt{\frac{8}{18} \times \frac{18}{40}} = +\frac{6}{10}$

It is given that $\sigma_x^2 = 9 \quad \Rightarrow \quad \sigma_x = 3$

But $b_{yx} = r\frac{\sigma_y}{\sigma_x}$

$$\frac{4}{5} = \frac{6}{10} \cdot \frac{\sigma_y}{3}$$

$$\sigma_y = 4$$

Example . Marks obtained by 10 students in Mathematics (X) and Statistics (Y) are given below:

X	60	34	40	50	45	40	22	43	42	64
Y	75	32	33	40	45	33	12	30	34	51

Find the two regression lines. Also find Y when $X = 55$.

X	Y	X^2	Y^2	XY
60	75	3600	5625	4500
34	32	1156	1024	1088
40	33	1600	1089	1320
50	40	2500	1600	2000

45	45	2025	2025	2025
40	33	1600	1089	1320
22	12	484	144	264
43	30	1849	900	1290
42	34	1764	1156	1428
64	51	4096	2601	3264
440	**385**	**20674**	**17253**	**18499**

$$\bar{X} = \frac{\sum X}{n} = \frac{440}{10} = 44 \qquad \text{and} \qquad \bar{Y} = \frac{\sum Y}{n} = \frac{385}{8} = 38.5$$

$$r\frac{\sigma_X}{\sigma_Y} = b_{XY} = \frac{n\sum XY - \sum X \sum Y}{n\sum Y^2 - \left(\sum Y\right)^2} = \frac{10(18499) - (440)\cdot(385)}{10(17253) - (385)^2} = 0.641$$

$$r\frac{\sigma_Y}{\sigma_X} = b_{YX} = \frac{n\sum XY - \sum X \sum Y}{n\sum X^2 - \left(\sum X\right)^2} = \frac{10(18499) - (440)\cdot(385)}{10(20674) - (440)^2} = 1.186$$

Regression line of Y on X

$$\left(Y - \bar{Y}\right) = r\frac{\sigma_Y}{\sigma_X}\left(X - \bar{X}\right)$$

$$\left(Y - 38.5\right) = 1.186\left(X - 44\right)$$

$$Y = 1.186X - 13.7$$

When $X = 55$, $Y = 1.186(55) - 13.7 = 51.5$

Regression line of X on Y

$$\left(X - \bar{X}\right) = r\frac{\sigma_X}{\sigma_Y}\left(Y - \bar{Y}\right)$$

$$\left(X - 44\right) = 0.641\left(Y - 38.5\right)$$

$$X = 0.641Y - 19.3$$

Transformation of Random Variable

Let the two random variables X and Y have the joint probability density function $f_{XY}(x,y)$. Two new random variables U and V are formed from X and Y as $U = g(X,Y)$ and $V = h(X,Y)$ where g and h are differentiable functions. Now the joint probability function of U and V is $f_{UV}(u,v) = f(x,y)|J|$ where $J = \begin{vmatrix} X_u & X_v \\ Y_u & Y_v \end{vmatrix}$.

Example 1. If the $p.d.f$ of a two dimensional random variable (X,Y) is given by $f(x,y) = x + y,\ 0 < x,\ y < 1$. Find the $p.d.f$ of $U = XY$.

Given $u = xy$ and let $v = y$. Re writing, we have $y = v,\ x = \frac{u}{y} = \frac{u}{v}$

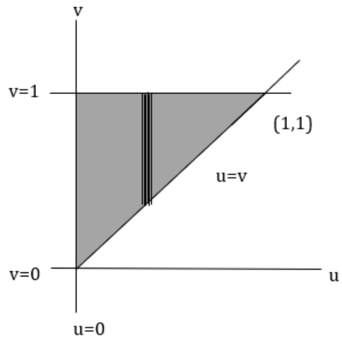
Now $J = \frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} \frac{1}{v} & -\frac{u}{v^2} \\ 0 & 1 \end{vmatrix} = \frac{1}{v}$

Therefore $p.d.f$ of (U,V) is $g(u,v) = f(x,y)|J|$

$$= (x+y)\frac{1}{v}$$

$$= \left(\frac{u}{v}+v\right)\frac{1}{v}$$

$$= \frac{u}{v^2}+1$$



Given limits are $0<x<1$ and $0<y<1$

i.e. $0<\frac{u}{v}<1$ and $0<v<1$

i.e. $0<u<v$

Therefore the $p.d.f$ of U is $g(u) = \int_{-\infty}^{\infty} g(u,v)\, dv$

$$= \int_{u}^{1} \frac{u}{v^2}+1\ dv$$

$$= \left(-\frac{u}{v}+v\right)_{u}^{1}$$

$$= \left(-\frac{u}{1}+1\right) - \left(-\frac{u}{u}+u\right)$$

$$= 2-2u,\ 0 \le u \le 1$$

Example 2. Let (X,Y) be a two dimensional non-negative continuous random variable having the joint density $f(x,y) = \begin{vmatrix} 4xye^{-(x^2+y^2)}, & x>0,\ y>0 \\ 0 & ,\ otherwise \end{vmatrix}$. Find the density function of $U=\sqrt{X^2+Y^2}$.

Given $u=\sqrt{x^2+y^2}$ and let $v=y$. Re writing, we have $y=v,\ \ x^2=u^2-y^2$

$$x=\sqrt{u^2-v^2}$$

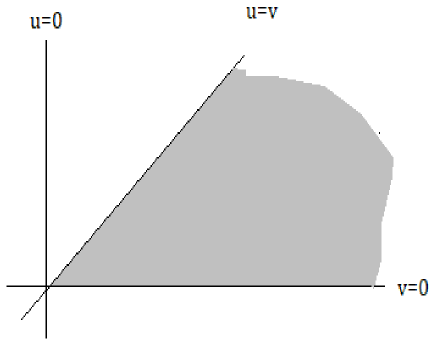
Now $J = \frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} \frac{u}{\sqrt{u^2-v^2}} & \frac{-v}{\sqrt{u^2-v^2}} \\ 0 & 1 \end{vmatrix} = \frac{u}{\sqrt{u^2-v^2}}$

Therefore $p.d.f$ of (U,V) is

$$g(u,v) = f(x,y)|J|$$

$$= \left(4\sqrt{u^2-v^2}\; v\, e^{-u^2}\right) \cdot \frac{u}{\sqrt{u^2-v^2}}$$

$$= 4uve^{-u^2}$$



Given limits are $x > 0$ and $y > 0$

i.e. $\sqrt{u^2-v^2} > 0$ and $v > 0$

i.e. $u^2 - v^2 > 0$

i.e. $u > v$

Therefore the $p.d.f$ of U is $g(u) = \int_{-\infty}^{\infty} g(u,v)\, dv$

$$= \int_0^u 4uve^{-u^2}\, dv$$

$$= 4ue^{-u^2}\left(\frac{v^2}{2}\right)_0^u$$

$$= 2u^3e^{-u^2},\ 0 < u < \infty$$

Example 3. If X and Y are independent with common $p.d.f$ (exponential), find the $p.d.f$ of $X - Y$.

Since X and Y follows exponential distribution, its $p.d.f$ is $f(x) = e^{-x},\ x > 0$ and

$f(y) = e^{-y},\ y > 0$

Also X and Y are independent. Hence $f(x,y) = f(x) \times f(y) = e^{-x}e^{-y},\ x > 0;\ y > 0$

Let $u = x - y$ and let $v = y$. Re writing, we have $y = v,\quad x = u + y$

$$x = u + v$$

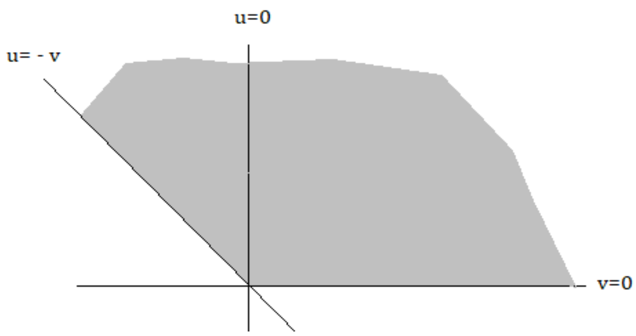
Now $J = \dfrac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} = 1$

Therefore $p.d.f$ of (U,V) is

$$g(u,v) = f(x,y)|J|$$

$$= e^{-(x+y)} \cdot 1$$

$$= e^{-(u+2v)}$$



Given limits are $x > 0$ and $y > 0$

i.e. $u + v > 0$ and $v > 0$

i.e. $u > -v$

Therefore when $u < 0;\ -u < v < \infty$ and when $u > 0;\ 0 < v < \infty$

Therefore the $p.d.f$ of U is $g(u) = \int_{-\infty}^{\infty} g(u,v)\, dv$

$$= \left| \begin{array}{l} \int_{-u}^{\infty} e^{-u-2v}\ dv,\ u < 0 \\ \int_{0}^{\infty} e^{-u-2v}\ dv,\ u > 0 \end{array} \right.$$

$$= \left| \begin{array}{l} \left(-\frac{1}{2} e^{-u-2v}\right)_{-u}^{\infty},\ u < 0 \\ \left(-\frac{1}{2} e^{-u-2v}\right)_{0}^{\infty},\ u > 0 \end{array} \right.$$

$$= \left| \begin{array}{l} \frac{1}{2} e^{u},\ u < 0 \\ \frac{1}{2} e^{-u},\ u > 0 \end{array} \right.$$

$$= \frac{1}{2} e^{-|u|},\ -\infty < u < \infty$$

Example 4. If X and Y are independent random variables with probability density functions $f(x) = 4e^{-4x}$, $x > 0$ and $f(y) = 2e^{-2y}$, $y > 0$. (i) Find the density function of $U = \dfrac{X}{X+Y}$, $V = X + Y$ (ii) Are U and V independent? (iii) What is $P(U > 0.5)$?

Since X and Y are independent, $f(x,y) = f(x) \times f(y) = 8e^{-4x}e^{-2y}$, $x > 0;\ y > 0$

Given $u = \dfrac{x}{x+y}$ and $v = x + y$. Re writing, we have, $u = \dfrac{x}{v}$ and $y = v - x$

$x = uv$ and $y = v - uv$

Now $J = \dfrac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} v & u \\ -v & 1-u \end{vmatrix} = v - uv + uv = v$

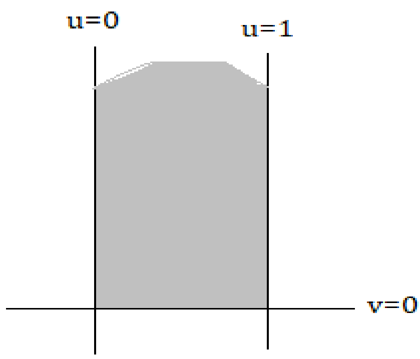
Therefore $p.d.f$ of (U,V) is

$$g(u,v) = f(x,y)|J|$$

$$= 8e^{-(4x+2y)} \cdot v$$

$$= 8v\, e^{-(4uv+2v-2vu)}$$

$$= 8v\, e^{-2v(u+1)}$$



Given limits are $x > 0$ and $y > 0$

i.e. $uv > 0$ and $v - uv > 0$

i.e. $u > 0$ & $v > 0$ and $1 - u > 0$ & $v > 0$

Combining, we get $0 < u < 1$ & $v > 0$

Now the $p.d.f$ of U is

$$g(u) = \int_{-\infty}^{\infty} g(u,v)\, dv$$

$$= \int_{0}^{\infty} 8ve^{-2v(1+u)}\, dv$$

$$= 8\left[v \cdot \frac{e^{-2v(1+u)}}{-2(1+u)} - 1 \cdot \frac{e^{-2v(1+u)}}{4(1+u)^2} \right]_0^{\infty}$$

$$= \frac{2}{(1+u)^2}$$

The $p.d.f$ of V is $g(v) = \int_{-\infty}^{\infty} g(u,v)\, du$

$$= \int_{0}^{1} 8ve^{-2v(1+u)}\, du$$

$$= 8v\left[\frac{e^{-2v(1+u)}}{-2v} \right]_0^1$$

$$= 8v\left[\frac{e^{-4v}}{-2v} + \frac{e^{-2v}}{2v} \right]$$

$$= 4\left[e^{-2v} - e^{-4v} \right]$$

Consider $g(u)\times g(v)=\frac{2}{(1+u)^2}\times 4\left[e^{-2v}-e^{-4v}\right]\neq g(u,v)$.

Therefore U and V are not independent.

$$P(U>0.5)=P(0.5<U<1)$$
$$=\int_{0.5}^{1}\frac{2}{(1+u)^2}\,du$$
$$=\left(-\frac{2}{1+u}\right)_{0.5}^{1}$$
$$=\frac{1}{3}$$

Example 5. If X and Y are independent random variables each normally distributed with mean zero and variance σ^2, find the $p.d.f$ of $R=\sqrt{X^2+Y^2}$ and $\theta=\tan^{1}\left(\frac{Y}{X}\right)$.

Since X and Y follows normal distribution with mean zero and variance σ^2, its $p.d.f$ is

$$f(x)=\frac{1}{\sqrt{2\pi}\,\sigma}e^{-\left(\frac{x^2}{2\sigma^2}\right)},\ -\infty<x<\infty \text{ and } f(y)=\frac{1}{\sqrt{2\pi}\,\sigma}e^{-\left(\frac{y^2}{2\sigma^2}\right)},\ -\infty<y<\infty.$$

Also X and Y are independent. Hence $f(x,y)=f(x)\times f(y)=\frac{1}{2\pi\sigma^2}e^{-\left(\frac{x^2+y^2}{2\sigma^2}\right)},\ -\infty<x,\ y<\infty$

Given transformations are $r=\sqrt{x^2+y^2}$ and $\theta=\tan^{1}\left(\frac{y}{x}\right)$.

Re writing, we have $x=r\cos\theta$ $y=r\sin\theta$

Therefore $p.d.f$ of (R,θ) is

$$g(r,\theta)=f(x,y)|J|$$
$$=\frac{1}{2\pi\sigma^2}e^{-\left(\frac{x^2+y^2}{2\sigma^2}\right)}\cdot r$$
$$=\frac{r}{2\pi\sigma^2}e^{-\left(\frac{r^2}{2\sigma^2}\right)}$$

Here

$$J=\frac{\partial(x,y)}{\partial(r,\theta)}=\begin{vmatrix}\frac{\partial x}{\partial r} & \frac{\partial x}{\partial\theta}\\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial\theta}\end{vmatrix}$$
$$=\begin{vmatrix}\cos\theta & -r\sin\theta\\ \sin\theta & r\cos\theta\end{vmatrix}$$
$$=r\left(\cos^2\theta+\sin^2\theta\right)$$
$$=r$$

Given limits are $-\infty<x,\ y<\infty$

i.e. $0<r<\infty$ and $0<\theta<2\pi$

The $p.d.f$ of R is $g(r)=\int_{-\infty}^{\infty} g(r,\theta)\,d\theta$

$$=\int_0^{2\pi} \frac{r}{2\pi\sigma^2}e^{-\left(\frac{r^2}{2\sigma^2}\right)}d\theta$$

$$=\frac{r}{2\pi\sigma^2}e^{-\left(\frac{r^2}{2\sigma^2}\right)}(2\pi)$$

$$=\frac{r}{\sigma^2}e^{-\left(\frac{r^2}{2\sigma^2}\right)}$$

The $p.d.f$ of Θ is $g(\theta)=\int_{-\infty}^{\infty} g(r,\theta)\,dr$

$$=\int_0^{\infty} \frac{r}{2\pi\sigma^2}e^{-\left(\frac{r^2}{2\sigma^2}\right)}dr$$

$$=\frac{1}{4\pi\sigma^2}\int_0^{\infty} e^{-\left(\frac{r^2}{2\sigma^2}\right)}d\left(r^2\right)$$

$$=\frac{1}{4\pi\sigma^2}\left[\frac{e^{-\left(\frac{r^2}{2\sigma^2}\right)}}{-\left(\frac{1}{2\sigma^2}\right)}\right]_0^{\infty}$$

$$=\frac{1}{2\pi}$$

Central Limit Theorem

We know that the sampling distribution of the mean will equal to the population mean. As the sample size increases, the sampling distribution of the mean will approach normality. The relationship between the shape of the population distribution and the shape of the sampling distribution of the mean is called the central limit Theorem.

Theorem: The probability distribution function of the sum of a large number of independent, identically distributed random variables X_i with mean μ and variance σ^2 approaches a Gaussian distribution with mean 0 and variance 1.

Definition : When sampling is done from a population with mean and standard deviation σ the sampling distribution of the sample mean $\bar{X}$ $\left(or\ \bar{X}=\frac{X_1+X_2+\cdots+X_n}{n}\right)$ will tend to a normal distribution with mean μ and standard deviation $\frac{\sigma}{\sqrt{n}}$ as the sample size n becomes large.

If $\bar{X}=X_1+X_2+\cdots+X_n$, then it will tend to a normal distribution with mean $n\mu$ and standard deviation $\sigma\sqrt{n}$ as the sample size n becomes large.

Note: The significance of the central limit theorem is that it permits us to use sample statistics (mean, variance etc.,) to make inferences about population parameters.

Example: The life time of a certain brand of an electric bulb may be considered as a random variable with mean 1200 hours and standard deviation 250 hours. Find the probability using central limit theorem that the average life time of 60 bulbs exceed 1250 hours.

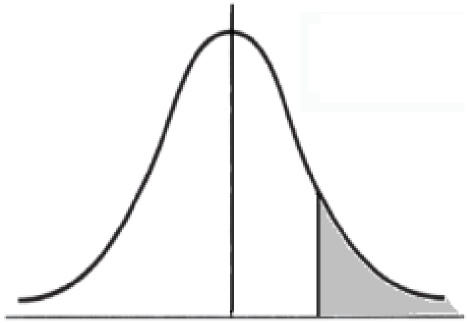
Given $n=60,\quad \mu=1200,\quad \sigma=250$.

Let $\bar{X} = \dfrac{X_1 + X_2 + \cdots + X_{60}}{60}$. By Central Limit Theorem, $\bar{X}$ follows normal distribution with mean μ and standard deviation $\dfrac{\sigma}{\sqrt{n}}$.

Therefore $P\left(\bar{X} \geq 1250\right) = P\left[Z \geq \dfrac{\bar{X} - \text{mean}}{\text{SD}}\right]$

$$= P\left[Z \geq \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{\text{n}}}}\right]$$

$$= P\left[Z \geq \frac{1250 - 1200}{\frac{250}{\sqrt{60}}}\right]$$



$$= P\left[Z \geq 1.55\right]$$

$$= 0.5 - P\left[0 \leq z \leq 1.55\right]$$

$$= 0.5 - 0.4394$$

$$= 0.0606$$

Example: A random sample of size 100 is taken from a population whose mean is 60 and variance 400. Using central limit theorem find what probability that we can assert that the mean of the sample will not differ from μ more than 4.

Given $n = 100, \quad \mu = 60, \quad \sigma^2 = 400, \quad \sigma = 20$.

Let $\bar{X}$ = sample mean. By Central Limit Theorem, $\bar{X}$ follows normal distribution with mean $\mu = 60$ and standard deviation $\dfrac{\sigma}{\sqrt{n}} = \dfrac{20}{\sqrt{100}} = 2$. We have to find the probability for $\left|\bar{X} - \mu\right| \leq 4$.

Therefore $P\left(\left|\bar{X} - \mu\right| \leq 4\right) = P\left(-4 \leq \bar{X} - \mu \leq 4\right)$

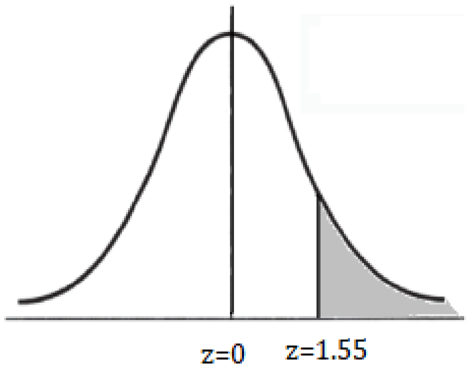
$$= P\left(\mu - 4 \leq \bar{X} \leq \mu + 4\right)$$

$$= P\left(56 \leq \bar{X} \leq 64\right)$$

$$= P\left[\frac{\bar{X} - \text{mean}}{\text{SD}} \leq Z \leq \frac{\bar{X} - \text{mean}}{\text{SD}}\right]$$

$$= P\left[\frac{X-\mu}{\frac{\sigma}{\sqrt{n}}} \le Z \le \frac{X-\mu}{\frac{\sigma}{\sqrt{n}}}\right]$$

$$= P\left[\frac{56-60}{2} \le Z \le \frac{64-60}{2}\right]$$



$$= P[-2 \le Z \le 2]$$

$$= 2 \times P[0 \le z \le 2]$$

$$= 2 \times 0.4772$$

$$= 0.9544$$

Example: Consider 12 uniformly distributed independent random variables in the range $(0, 1)$. Find $P\left[\sum X_i > 8\right]$.

Here each X_i are uniformly distributed in $(0, 1)$. Hence mean $E(X) = \frac{a+b}{2} = \frac{0+1}{2} = \frac{1}{2}$ and variance $Var(X) = \frac{b^2-a^2}{12} = \frac{1-0}{12} = \frac{1}{12}$

Here $\bar{X} = X_1 + X_2 + \cdots + X_{12}$. Therefore mean $n\mu = 12 \times \frac{1}{2} = 6$ and standard deviation $\sigma\sqrt{n} = \sqrt{\frac{1}{12}} \times \sqrt{12} = 1$

Therefore $P\left[\sum X_i > 8\right] = P\left[Z > \frac{8-mean}{SD}\right]$

$$= P\left[Z > \frac{8-6}{\sqrt{1}}\right]$$

$$= P[Z > 2]$$

$$= 0.5 - P[0 \le Z \le 2]$$

$$= 0.5 - 0.4772$$

$$= 0.0228$$

Example : A box contains many 100ohm resistors with a tolerance of ± 10 ohms. If 10 resistors are drawn and connected in series find the probability that the resistance of the circuit is between 900 and 1100.

Let X be a random variable representing resistance of one resistor. It follows uniform distribution in $(100-10, 100+10) = (90, 110)$.

Hence mean $E(X)=\frac{a+b}{2}=\frac{90+110}{2}=100$ and variance $Var(X)=\frac{b^2-a^2}{12}=\frac{110^2-90^2}{12}=\frac{4000}{12}$

Here $\bar{X}=X_1+X_2+\cdots+X_{10}$. Therefore mean $n\mu=10\times100=1000$ and standard deviation

$$\sigma\sqrt{n}=\sqrt{\frac{4000}{12}}\times\sqrt{10}=\sqrt{10\cdot\frac{4000}{12}}$$

Therefore $P\left[900\le\sum X_i\le1100\right]=P\left[\frac{900-mean}{SD}\le Z\le\frac{1100-mean}{SD}\right]$

$$=P\left[\frac{900-1000}{\sqrt{\frac{40000}{12}}}\le Z\le\frac{1100-1000}{\sqrt{\frac{40000}{12}}}\right]$$

$$=P\left[-1.732\le Z\le1.732\right]$$

$$=2\times P\left[0\le Z\le1.732\right]$$

$$=0.9164$$

Example : 20 dice are thrown. Find the probability that the sum obtained is between 65 and 75 using central limit theorem.

Let X_i be a random variable representing the number shown on the i^{th} die. $\therefore$ $P(X_i)=\frac{1}{6}$.

Then $\mu=E(X)=\sum xP(x)=1.\frac{1}{6}+2.\frac{1}{6}+3.\frac{1}{6}+4.\frac{1}{6}+5.\frac{1}{6}+6.\frac{1}{6}=\frac{7}{2}$

$$E(X^2)=\sum x^2P(x)=1^2.\frac{1}{6}+2^2.\frac{1}{6}+3^2.\frac{1}{6}+4^2.\frac{1}{6}+5^2.\frac{1}{6}+6^2.\frac{1}{6}=\frac{91}{6}$$

$$\sigma^2=Var(X)=E(X^2)-\left[E(X)\right]^2=\frac{91}{6}-\frac{49}{4}=\frac{35}{12}$$

Let $\bar{X}=X_1+X_2+\cdots+X_{20}$ be the sum of total faces.

Therefore $\bar{X}$ follows normal distribution with mean $n\mu=20\times\frac{7}{2}=70$ and standard deviation

$$\sigma\sqrt{n}=\sqrt{\frac{35}{12}}\times\sqrt{20}=\sqrt{\frac{175}{3}}$$

Therefore $P\left[65\le\sum X_i>70\right]=P\left[\frac{65-70}{\sqrt{\frac{175}{3}}}\le Z\le\frac{75-70}{\sqrt{\frac{175}{3}}}\right]$

$$=P\left[-0.65\le Z\le0.65\right]$$

$$=2\times P\left[0\le Z\le0.65\right]$$

$$= 2 \times 0.2422$$
$$= 0.4844$$

Example : An economist wishes to estimate the average family income in a certain population. The population standard deviation is known to be 4,500, and the economist uses a random sample of size $n = 225$. What is the probability that the sample mean will fall within 800 of the population mean?

Given $n = 225, \quad \mu = 800, \qquad \sigma = 4500$.

Let $\bar{X}$ = sample mean. By Central Limit Theorem, $\bar{X}$ follows normal distribution with mean $\mu = 800$ and standard deviation $\dfrac{\sigma}{\sqrt{n}} = \dfrac{4500}{\sqrt{225}} = 300$. We have to find the probability for $\left|\bar{X} - \mu\right| \leq 800$.

Therefore $P\left(\left|\bar{X} - \mu\right| \leq 800\right) = P\left(-800 \leq \bar{X} - \mu \leq 800\right)$

$$= P\left(\mu - 800 \leq \bar{X} \leq \mu + 800\right)$$

$$= P\left(0 \leq \bar{X} \leq 1600\right)$$

$$= P\left[0 \leq Z \leq \frac{\bar{X} - \text{mean}}{\text{SD}}\right]$$

$$= P\left[0 \leq Z \leq \frac{X - \mu}{\frac{\sigma}{\sqrt{\text{n}}}}\right]$$

$$= P\left[0 \leq Z \leq \frac{1600 - 800}{300}\right]$$

$$= P\left[0 \leq Z \leq 2.67\right]$$
$$= 0.4962$$

Exercise

1. Let X and Y be two discrete random variables with joint probability mass function

$$P[X=x,Y=y]=\begin{cases}\frac{1}{18}(2x+y), & x=1,2 \quad \& \quad y=1,2\\ =0 & otherwise\end{cases}$$

Find the marginal probability mass functions of X and Y.

2. The two lines of regression are $8X-10Y+66=0$, $40X-18Y-214=0$. The variance of X is 9. Find the mean values of X and Y. Variance of Y. Also find the coefficient of correlation between the variables X and Y.

3. If the joint $p.d.f$ of (X,Y) is $f(x,y)=\frac{1}{4}$, $0<x,\ y<2$. Find $P(X+Y\le 1)$.

4. If the joint $p.d.f$ of two dimensional random variable (X,Y) is given by $f(x,y)=x^2+\frac{xy}{43}$, $0<x<1, 0<y<2$. Find (i) $P\left(X>\frac{1}{2}\right)$ (ii) $P(Y<X)$.

5. The joint $p.d.f$ of the two dimensional random variable (X,Y) is given by $f(x,y)=\frac{8}{9}xy$, $1\le x\le y\le 2$. Find (i) Marginal densities of X and Y (ii) Conditional density functions $f(y/x)$ and $f(x/y)$.

6. The joint probability density function of the two dimensional random variable (X,Y) is $f(x,y)=2-x-y$, $0\le x,y\le 1$. Find the correlation coefficient between X and Y.

7. If the independent random variables X and Y have variances 36 and 16 respectively, find the correlation coefficient between $X+Y$ and $X-Y$.

8. If X and Y are independent random variables with probability density functions $f(x)=e^{-x}$, $x>0$ and $f(y)=e^{-y}$, $y>0$. (i) Find the $p.d.f$ of $U=\frac{X}{X+Y}$, $V=X+Y$ (ii) Are U and V independent?

9. A small college has 90 male and 30female professors. An ad-hoc committee of 5 is selected at random to write the vision and mission of the college. If X and Y are the number of men and women in the committee, respectively what is the joint probability mass function of X and Y ?

ANSWERS

1. $P_X(x)=\frac{1}{18}(4x+3),\ \ x=1,2$ & $P_Y(y)=\frac{1}{18}(2y+6),\ \ y=1,2$ **2.** $\overline{X}=13, \overline{Y}=17$, $\sigma_Y=4,\ r=0.6$

3. $\frac{1}{8}$ **4.** $\frac{5}{6}, \frac{7}{24}$ **5.** $\frac{4}{9}\left(4x-x^3\right);\ \frac{4}{9}\left(y^3-y\right)$ and $\frac{2x}{4-x^2}, \frac{2y}{y^2-1}$ **6.** $-\frac{1}{11}$ **7.** $\sqrt{\frac{5}{13}}$

8. $g(u,v)=ve^{-v},\ \le u\le 1,\ v\ge 0$, Independent **9.** $p(x,y)=\frac{90C_x\ 30C_y}{120C_5},\ \ \begin{array}{l} x=0,1,\ldots,90 \\ y=0,1,\ldots,30 \end{array}$

Unit-III Random Processes

A stochastic process is a collection of random variables evolving over time, which is used to model uncertainty in dynamic systems. Stochastic processes are also used to represent signals with random behavior, such as thermal noise in receivers or fading in wireless channels. A Poisson process is often used to model the arrival of packets in communication networks, particularly in queueing theory.

Random variable is a function $X(s)$ that assigns a real number to each outcome of an experiment. But random process is a function of time $X(s,t)$ that is assigned to each sample point based on some rule.

Classification of Random Processes

- If both s and t are discrete, the random process is called a discrete random sequence.
- If s is discrete and t is continuous, the random process is called a discrete random process.
- If both s is continuous and t is discrete, the random process is called a continuous random sequence.
- If both s and t are continuous, the random process is called a continuous random process.

Let $\{X(t)\}$ be a random process. The mean of $\{X(t)\}$ is defined by

$E\{X(t)\} = \mu_X(t) = \int_{-\infty}^{\infty} x\, f_X(x,t)\, dx$ where $X(t)$ is treated as a random variable for a fixed value of t

The auto correlation of the process $\{X(t)\}$ is defined as

$$R_{XX}(t_1\ ,\ t_2) = E\left[\{X(t_1)\quad X(t_2)\ \}\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2\ f_X(x_1, x_2 : t_1, t_2)\, dx_1\ dx_2$$

The auto covariance of the process $\{X(t)\}$ is defined as $C_{XX}(t_1\ ,\ t_2) = R_{XX}(t_1,t_2) - \mu_X(t_1)\mu_X(t_2)$

Special Classes of Random Processes

If certain probability distribution or averages do not depend on t, then the random process is called stationary.

A random process $\{X(t)\}$ is said to be first order stationary process if $\mu = E[\{X(t)\}]$ is a constant

A random process $\{X(t)\}$ is said to be strongly stationary or strict sense stationary if all its finite dimensional distributions are invariant under translation of time parameter.

A random process $\{X(t)\}$ is said to be second order stationary if the second order density must be invariant under translation of time parameter. In particular, the auto correlation function is a function of time difference.

https://doi.org/10.5281/zenodo.15090289

A random processes $\{X(t)\}$ is called Wide Sense Stationary(Weak Stationary/Covariance Stationary), if its mean is constant and autocorrelation function depends only on the time difference. i.e.

$E\{X(t)\}$ is always a constant and

$E\{X(t)\cdot X(t+\tau)\} = R_{XX}(\tau)$ (or) $E\{X(t_1)\cdot X(t_2)\} = R_{XX}(t_2 - t_1)$.

Two random processes $\{X(t)\}$ and $\{Y(t)\}$ is called jointly wide sense stationary, if each is WSS and their cross correlation function depends only on the time difference. i.e. $E\{X(t)\cdot Y(t+\tau)\} = R_{XY}(\tau)$

A random process $\{X(t)\}$ is said to be mean ergodic if ensembled mean is equal to time average mean. i.e. $E\{X(t)\} = \underset{T\to\infty}{Lt} \dfrac{1}{2T}\int_{-T}^{T} X(t)\, dt$

A random process $\{x(t)\}$ is said to be correlation ergodic if $R_{XX}(t_1, t_2) = \underset{T\to\infty}{Lt} \dfrac{1}{2T}\int_{-T}^{T} X(t_1)\cdot X(t_2) dt$

A random process $\{X(t)\}$ which is not stationary in any sense is called evolutionary random process. Poisson process is an evolutionary process.

Properties of Auto Correlation Function

- $R_{XX}(\tau)$ is an even function
- $R_{XX}(\tau)$ is maximum at $\tau = 0$ i.e. $|R(\tau)| \le R(0)$
- Mean of $[X(t)] = \sqrt{\underset{\tau\to\infty}{Lt} R_{XX}(\tau)}$
- $E[X^2(t)] = R(0)$
- $R_{YX}(\tau) = R_{XY}(-\tau)$
- $|R_{XY}(\tau)| = \sqrt{R_{XX}(0)\cdot R_{YY}(0)}$
- Two random processes $\{X(t)\}$ and $\{Y(t)\}$ are orthogonal if $R_{XY}(\tau) = 0$

Example: The probability distribution of the process $\{X(t)\}$ is given by

$$P[X(t)=n] = \begin{cases} \dfrac{(at)^{n-1}}{(1+at)^{n+1}}, & n=1,2,3,\ldots\ldots \\ \dfrac{at}{1+at} & , n=0 \end{cases}$$

Show that $\{X(t)\}$ is not stationary.

Given $\{X(t)\}$ and $P\{X(t)\}$ are tabulated as follows:

$\{X(t)\}$	n	0	1	2	.....
$P\{X(t)\}$	$P\{X_n\}$	$\frac{at}{1+at}$	$\frac{at}{(1+at)^2}$	$\frac{at}{(1+at)^3}$	.....

$$E[X(t)] = \sum_{n=0}^{\infty} n\, p(x_n)$$

$$= 0.\frac{at}{1+at} + 1.\frac{1}{(1+at)^2} + 2.\frac{at}{(1+at)^3} + 3.\frac{(at)^2}{(1+at)^4} + \ldots\ldots\ldots$$

$$= \frac{1}{(1+at)^2}\left[1 + 2.\frac{at}{(1+at)} + 3.\frac{(at)^2}{(1+at)^2} + \ldots\ldots\ldots\right]$$

$$= \frac{1}{(1+at)^2}\left[1 - \frac{at}{(1+at)}\right]^{-2}$$

$$= \frac{1}{(1+at)^2}\left[\frac{1}{(1+at)}\right]^{-2}$$

$$= \frac{1}{(1+at)^2}(1+at)^2$$

$$= 1,\ a\ \text{constant}$$

$$E\left[X^2(t)\right] = \sum_{n=0}^{\infty} n^2\ p(x_n)$$

$$= \sum_{n=0}^{\infty}\left[\, n\ (n+1) - n\ \right]\ p(x_n)$$

$$= \sum_{n=0}^{\infty}\left[\, n\ (n+1)\ \right]\ p(x_n) - \sum_{n=0}^{\infty} n\ p(x_n)$$

$$= \left[1\ .2\ \frac{1}{(1+at)^2} + 2\ .\ 3\ \frac{at}{(1+at)^3} + 3\ .\ 4\ \frac{(at)^2}{(1+at)^4} + \ldots\ldots\ldots\right] - 1$$

$$= \frac{1\ .\ 2}{(1+at)^2}\left[1 + \frac{2\ .\ 3}{1\ .\ 2}\ \frac{at}{(1+at)} + \frac{3\ .\ 4}{1\ .\ 2}\ \frac{(at)^2}{(1+at)^2} + \ldots\ldots\ldots\right] - 1$$

$$= \frac{2}{(1+at)^2}\left[1 - \frac{at}{(1+at)}\right]^{-3} - 1$$

$$= \frac{2}{(1+at)^2} \left[\frac{1}{(1+at)} \right]^{-3} - 1$$

$$= 2 \left(1 + at \right) - 1$$

$$= 1 + 2at, \ \textit{which is dependent on } t.$$

Therefore {X(t)} is not stationary

Example : The auto correlation function of a stationary random process is $R(\tau) = 16 + \frac{9}{1+6\tau^2}$. Find the mean and variance of the process.

$E\left[X^2(t)\right] = R(0) = 16 + 9 = 25$ and $\left(E\left[X(t)\right]\right)^2 = \underset{\tau \to \infty}{Lt} R(\tau) = \underset{\tau \to \infty}{Lt} 16 + \frac{9}{1+6\tau^2} = 16$

Hence Mean $E\left[X(t)\right] = 4$. Therefore $Var\left[X(t)\right] = E\left[X^2(t)\right] - \left(E\left[X(t)\right]\right)^2 = 25 - 16 = 9$

Example : A random processes $\{X(t)\}$ is defined by $X(t) = K\cos\omega t, t \geq 0$ where ω is a constant and K is uniformly distributed over $(0,2)$. Find the autocorrelation of $X(t)$.

Since K is uniformly distributed over $(0,2)$, $f_K(k) = \frac{1}{2}, 0 < k < 2$

Autocorrelation $R_{XX}(t_2 - t_1) = E\left\{X(t_1) \cdot X(t_2)\right\}$

$$= E\left(K\cos\omega t_1 \cdot K\cos\omega t_2\right)$$

$$= \left(\cos(\omega t_1) \cdot \cos(\omega t_2)\right) E\left(K^2\right)$$

$$= \frac{1}{2}\left(\cos(\omega t_1 + \omega t_2) + \cos(\omega t_1 - \omega t_2)\right) \int_0^2 k^2 . \frac{1}{2} dk$$

$$= \frac{1}{4}\left(\cos(\omega t_1 + \omega t_2) + \cos(\omega t_1 - \omega t_2)\right) \left[\frac{k^3}{3}\right]_0^2$$

$$= \frac{2}{3}\left(\cos(\omega t_1 + \omega t_2) + \cos(\omega t_1 - \omega t_2)\right)$$

Example : Show that the process $X(t) = A\cos(\omega t + \theta)$ is not stationary if θ is uniformly distributed in $(0,\pi)$.

Given $X(t) = A\cos(\omega t + \theta)$ and θ is uniformly distributed in $(0,\pi)$.

Therefore $f(\theta) = \frac{1}{\pi}, \ 0 < \theta < \pi$

$$E[X(t)] = E[A\,cos(\,wt+\theta\,)]$$

$$= A\int_{0}^{\pi} cos(\,wt+\theta\,)\; f(\theta)\, d\theta$$

$$= \frac{A}{\pi}\int_{0}^{\pi} cos(\,wt+\theta\,)\; d\theta$$

$$= \frac{A}{\pi}\,[sin(\,wt+\theta\,)]_{0}^{\pi}$$

$$= \frac{A}{\pi}\,[sin(\,\pi+wt\,) - sin\,\omega t\,]$$

$$= \frac{A}{\pi}\,[-\,sin\;wt\,-\,sin\,\omega t\,]$$

Since $E\{X(t)\}$ depends on t, it is not stationary.

Example: Show that the random process $X(t) = A\sin(\omega t+\theta)$ is wide sense stationary where A and ω are constants and θ is uniformly distributed in $(0, 2\pi)$.

Since θ is uniformly distributed in $(0, 2\pi)$, The $p.d.f$ is $f(\theta) = \frac{1}{2\pi},\; 0<\theta<2\pi$

$$E[X(t)] = E[A\,sin(\,wt+\theta\,)]$$

$$= A\int_{0}^{2\pi} sin(\,wt+\theta\,)\; f(\theta)\, d\theta$$

$$= \frac{A}{2\pi}\int_{0}^{2\pi} sin(\,wt+\theta\,)\; d\theta$$

$$= \frac{A}{2\pi}\,[-cos(\,wt+\theta\,)]_{0}^{2\pi}$$

$$= \frac{A}{2\pi}\,[-\,cos(\,2\pi+wt\,) + cos\,\omega t\,]$$

$$= \frac{A}{2\pi}\,[-\,cos\;wt\,+\,cos\,\omega t\,]$$

$$= 0,\; a\; constant$$

$$E[X(t_1)X(t_2)] = E[A\sin(wt_1+\theta)A\sin(wt_2+\theta)]$$

$$= A^2\int_0^{2\pi} \sin(wt_1+\theta)A\sin(wt_2+\theta)\; f(\theta)\,d\theta$$

$$= \frac{A^2}{2\pi}\int_0^{2\pi} \sin(wt_1+\theta)\sin(wt_2+\theta)\; d\theta$$

$$= \frac{A^2}{2\pi}\int_0^{2\pi} \cos(wt_1-wt_2) - \cos(wt_1+wt_2+2\theta)\; d\theta$$

$$= \frac{A^2}{2\pi}\cos(wt_1-wt_2)[\,\theta\,]_0^{2\pi} - \frac{A^2}{2\pi}[\sin(wt_1+wt_2+2\theta)]_0^{2\pi}$$

$$= A^2\cos(wt_1-wt_2) - \frac{A^2}{2\pi}[\sin(4\pi+wt_1+wt_2)-\sin(wt_1+wt_2)]$$

$$= A^2\cos(wt_1-wt_2) - \frac{A^2}{2\pi}[\sin(wt_1+wt_2)-\sin(wt_1+wt_2)]$$

$$= A^2\cos(wt_1-wt_2),\; a\;\; function\, of\; (t_1-t_2)$$

Therefore $\{X(t)\}$ is $W.S.S$.

Example: Show that the random process $X(t)=\cos(\lambda t+Y)$ is stationary in the wide sense If Y is the random variable with characteristic function $\phi(\omega)$ and if $\phi(1)=0$ and $\phi(2)=0$.

The characteristic function $\phi(\omega)=E\left[e^{i\omega Y}\right]$

$$=E[\cos\omega Y+i\sin\omega Y]$$

$$=E[\cos\omega Y]+i\;E[\sin\omega Y]$$

But $\phi(1)=E[\cos Y]+i\;E[\sin Y]$

$0=E[\cos Y]+i\;E[\sin Y]$

Therefore $E[\cos Y]=0, \quad E[\sin Y]=0$.

Similarly $\phi(2)=0$ gives $E[\cos 2Y]=0, \quad E[\sin 2Y]=0$

Now $E(X(t))=E(\cos(\lambda t+Y))$

$$=E(\cos\lambda t\,\cos Y+\sin\lambda t\,\sin Y)$$

$$=\cos\lambda t\cdot E(\cos Y)+\sin\lambda t\cdot E(\sin Y)$$
$$=0$$

Example : A random processes $\{X(t)\}$ defined by $X(t) = A\cos t + B\sin t$ where A and B are independent random variables each of which has a value -2 with probability 1/3 and a value 1 with probability 2/3. Show that $X(t)$ is a wide sense stationary process.

The probability distribution of the discrete random variables A and B are as follows:

A	-2	1
$P(A)$	$\frac{1}{3}$	$\frac{2}{3}$
B	-2	1
$P(B)$	$\frac{1}{3}$	$\frac{2}{3}$

Therefore $E(A) = \sum A \cdot P(A) = -2\left(\frac{1}{3}\right) + 1\left(\frac{2}{3}\right) = 0$

Also $E\left(A^2\right) = \sum A^2 \cdot P(A) = 4\left(\frac{1}{3}\right) + 1\left(\frac{2}{3}\right) = 2$

Hence $E(B) = 0$, $E\left(B^2\right) = 2$

Since A and B are independent $E(AB) = E(A) \cdot E(B) = 0$.

$$\begin{aligned} E(X(t)) &= E(A\cos t + B\sin t) \\ &= E(A)\cos t + E(B)\sin t \\ &= 0 \end{aligned}$$

$$\begin{aligned} R_{XX}(t_1, t_2) &= E\{X(t_1) \cdot X(t_2)\} \\ &= E\left[(A\cos t_1 + B\sin t_1) \cdot (A\cos t_2 + B\sin t_2)\right] \\ &= E\left[A^2 \cos t_1 \cos t_2 + AB\cos t_1 \sin t_2 + BA\sin t_1 \cos t_2 + B^2 \sin t_1 \sin t_2\right] \\ &= E\left(A^2\right)\cos t_1 \cos t_2 + E(AB)\cos t_1 \sin t_2 + E(BA)\sin t_1 \cos t_2 + E\left(B^2\right)\sin t_1 \sin t_2 \\ &= 2\cos t_1 \cos t_2 + 2\sin t_1 \sin t_2 \\ &= 2\cos(t_2 - t_1) \end{aligned}$$

Hence $X(t)$ is a wide sense stationary process.

Example : Show that the random processes $\{X(t)\}$ defined by $X(t) = A\cos \omega t + B\sin \omega t$ where A and B are independent random variables with zero means and equal variance is wide sense stationary.

Given $E(A) = E(B) = 0$. Also $\left[E(A)\right]^2 - E(A) = \left[E(B)\right]^2 - E(B) = \sigma^2$.

Therefore $\left[E(A)\right]^2 = \left[E(B)\right]^2 = \sigma^2$

Since A and B are independent $E(AB) = E(A) \cdot E(B) = 0$.

$$E\left(X(t)\right)=E\left(A\cos\omega t+B\sin\omega t\right)$$
$$=E(A)\cos\omega t+E(B)\sin\omega t$$
$$=0$$

$$R_{XX}(t_1,t_2)=E\left\{X(t_1)\cdot X(t_2)\right\}$$
$$=E\left[\left(A\cos\omega t_1+B\sin\omega t_1\right)\cdot\left(A\cos\omega t_2+B\sin\omega t_2\right)\right]$$
$$=E\left[A^2\cos\omega t_1\cos\omega t_2+AB\cos\omega t_1\sin\omega t_2+BA\sin\omega t_1\cos\omega t_2+B^2\sin\omega t_1\sin\omega t_2\right]$$
$$=E\left(A^2\right)\cos\omega t_1\cos\omega t_2+E(AB)\cos\omega t_1\sin\omega t_2+E(BA)\sin\omega t_1\cos\omega t_2+E\left(B^2\right)\sin\omega t_1\sin\omega t_2$$
$$=\sigma^2\cos\omega t_1\cos\omega t_2+\sigma^2\sin\omega t_1\sin\omega t_2$$
$$=\sigma^2\cos\left(t_2-t_1\right)$$

Hence $X(t)$ is a wide sense stationary process.

Example : Let $X(t)=A\cos\omega t+B\sin\omega t$ be a random process where A and B are independent normally distributed random variables with $N\left(0,\sigma^2\right)$. Is $X(t)$ covariance stationary?

Given $E(A)=E(B)=0$. Also $\left[E(A)\right]^2-E(A)=\left[E(B)\right]^2-E(B)=\sigma^2$.

Therefore $\left[E(A)\right]^2=\left[E(B)\right]^2=\sigma^2$

Since A and B are independent $E(AB)=E(A)\cdot E(B)=0$.

$$E\left(X(t)\right)=E\left(A\cos\omega t+B\sin\omega t\right)$$
$$=E(A)\cos\omega t+E(B)\sin\omega t$$
$$=0$$

$$R_{XX}(t_1,t_2)=E\left\{X(t_1)\cdot X(t_2)\right\}$$
$$=E\left[\left(A\cos\omega t_1+B\sin\omega t_1\right)\cdot\left(A\cos\omega t_2+B\sin\omega t_2\right)\right]$$
$$=E\left[A^2\cos\omega t_1\cos\omega t_2+AB\cos\omega t_1\sin\omega t_2+BA\sin\omega t_1\cos\omega t_2+B^2\sin\omega t_1\sin\omega t_2\right]$$
$$=E\left(A^2\right)\cos\omega t_1\cos\omega t_2+E(AB)\cos\omega t_1\sin\omega t_2+E(BA)\sin\omega t_1\cos\omega t_2+E\left(B^2\right)\sin\omega t_1\sin\omega t_2$$
$$=\sigma^2\cos\omega t_1\cos\omega t_2+\sigma^2\sin\omega t_1\sin\omega t_2$$
$$=\sigma^2\cos\left(t_2-t_1\right)$$

Therefore covariance $C(t_1,t_2)=R(t_1,t_2)-E(t_1)E(t_2)$

$$=\sigma^2\cos(t_2-t_1)$$

Since covariance is a function of (t_2-t_1) and $E[X(t)]$ is constant, $X(t)$ is covariance stationary.

Example: If $X(t)=A\cos(100t+\theta)$ where A and θ are independent random variables with $E(A)=2$, $V(A)=6$ and θ is uniformly distributed in $(-\pi,\pi)$. Prove that $\{X(t)\}$ is WSS process.

Since θ is uniformly distributed over $(-\pi,\pi)$, the p.d.f of θ is $f(\theta)=\dfrac{1}{2\pi}, -\pi<\theta<\pi$

$$E(X(t))=E(A\cos(100t+\theta))$$
$$=E(A)E(\cos(100t+\theta))$$
$$=2\int_{-\pi}^{\pi}\frac{1}{2\pi}\cos(100t+\theta)\,d\theta$$
$$=2\times2\times\frac{1}{2\pi}\int_{0}^{\pi}\cos(100t+\theta)\,d\theta$$
$$=\frac{2}{\pi}\left[\sin(100t+\theta)\right]_0^{\pi}$$
$$=\frac{2}{\pi}\left[\sin(100t+\pi)-\sin(100t)\right]$$
$$=\frac{2}{\pi}\left[\sin(100t)-\sin(100t)\right]$$
$$=0$$

We know that $V(A)=E(A^2)-[E(A)]^2$

$$6=E(A^2)-[2]^2$$
$$E(A^2)=10$$

We know that $R_{XX}(t_1,t_2)=E\{X(t_1)\cdot X(t_2)\}$

$$=E\left[\left(A\cos(100t_1+\theta)\right)\left(A\cos(100t_2+\theta)\right)\right]$$
$$=E(A^2)\cdot E\left[\left(\cos(100t_1+\theta)\right)\left(\cos(100t_2+\theta)\right)\right]$$
$$=\frac{E(A^2)}{2}\cdot E\left[\cos(100t_1+\theta+100t_2+\theta)+\cos(100t_1+\theta-100t_2-\theta)\right]$$

$$= \frac{E(A^2)}{2} \cdot E\left[\cos\left(100(t_1+t_2)+2\theta\right)+\cos\left(100(t_1-t_2)\right)\right]$$

$$= \frac{E(A^2)}{2} \cdot E\left[\cos\left(100(t_1+t_2)+2\theta\right)\right]+\frac{E(A^2)}{2} \cdot E\left[\cos\left(100(t_1-t_2)\right)\right]$$

$$= \frac{E(A^2)}{2} \cdot E\left[\cos\left(100(t_1+t_2)+2\theta\right)\right]+\frac{E(A^2)}{2} \cdot \cos\left(100(t_1-t_2)\right)$$

Consider

$$E\left[\cos\left(100(t_1+t_2)+2\theta\right)\right] = \int_{-\pi}^{\pi} \cos\left(100(t_1+t_2)+2\theta\right) \cdot f(\theta)\, d\theta$$

$$= \frac{2}{2\pi}\int_{0}^{\pi} \cos\left(100(t_1+t_2)+2\theta\right) d\theta$$

$$= \frac{2}{2\pi}\left[\frac{\sin\left(100(t_1+t_2)+2\theta\right)}{2}\right]_0^{\pi}$$

$$= \frac{1}{2\pi}\left[\sin 100(t_1+t_2) - \sin 100(t_1+t_2)\right]$$

$$= 0$$

Therefore $R_{XX}(t_1, t_2) = \frac{E(A^2)}{2} \cdot \cos\left(100(t_1-t_2)\right)$

Since $E(X(t))$ is a constant and auto correlation is a function of time difference (t_1-t_2), $\{X(t)\}$ is a wide sense stationary.

Example : If $X(t) = A\cos(\Omega t+\theta)$ where Ω is a random variable with density function $f_\Omega(\omega)$ and θ is uniformly distributed in $(-\pi, \pi)$ and is independent with Ω is WSS process.

Since θ is uniformly distributed over $(-\pi, \pi)$, the p.d.f of θ is $f(\theta) = \frac{1}{2\pi}, -\pi < \theta < \pi$

$$E(X(t)) = E\left(A\cos(\Omega t+\theta)\right)$$

$$= A \cdot E\left(\cos(\Omega t+\theta)\right)$$

$$= A \cdot E\left(\cos\Omega t \cos\theta - \sin\Omega t \sin\theta\right)$$

$$= A \cdot E(\cos\Omega t)E(\cos\theta) - A \cdot E(\sin\Omega t)E(\sin\theta)$$

$$= A \cdot E(\cos\Omega t)\frac{1}{2\pi}\int_{-\pi}^{\pi} \cos\theta\, d\theta - A \cdot E(\sin\Omega t)\frac{1}{2\pi}\int_{-\pi}^{\pi} \sin\theta\, d\theta$$

$$= A \cdot E(\cos \Omega t) \frac{2}{2\pi} \int_0^{\pi} \cos\theta \, d\theta - A \cdot E(\sin \Omega t)(0)$$

$$= A \cdot E(\cos \Omega t) \frac{2}{2\pi} [\sin\theta]_0^{\pi}$$

$$= A \cdot E(\cos \Omega t) \frac{2}{2\pi} [\sin\pi - \sin 0]$$

$$= 0$$

Also $R_{XX}(t_1, t_2) = E\{X(t_1) \cdot X(t_2)\}$

$$= E\left[\left(A\cos(\Omega t_1 + \theta)\right)\left(A\cos(\Omega t_2 + \theta)\right)\right]$$

$$= E(A^2) E\left[\left(\cos(\Omega t_1 + \theta)\right)\left(\cos(\Omega t_2 + \theta)\right)\right]$$

$$= E(A^2) E\left[(\cos\Omega t_1 \cos\theta - \sin\Omega t_1 \sin\theta)(\cos\Omega t_2 \cos\theta - \sin\Omega t_2 \sin\theta)\right]$$

$$= E(A^2) E\begin{bmatrix} \cos\Omega t_1 \cos\Omega t_2 \cos^2\theta - \cos\Omega t_1 \cos\theta \sin\Omega t_2 \sin\theta \\ -\sin\Omega t_1 \sin\theta \cos\Omega t_2 \cos\theta + \sin\Omega t_1 \sin\Omega t_2 \sin^2\theta \end{bmatrix}$$

$$= E(A^2) E[\cos\Omega t_1 \cos\Omega t_2] E\left[\cos^2\theta\right] + E(A^2) E[\sin\Omega t_1 \sin\Omega t_2] E\left[\sin^2\theta\right]$$

$$-E(A^2) E[\cos\Omega t_1 \sin\Omega t_2 + \sin\Omega t_1 \cos\Omega t_2] E[\sin\theta\cos\theta]$$

$$= E(A^2) E[\cos\Omega t_1 \cos\Omega t_2] E\left[\cos^2\theta\right] + E(A^2) E[\sin\Omega t_1 \sin\Omega t_2] E\left[\sin^2\theta\right]$$

$$-\frac{1}{2} E(A^2) E\left[\sin(\Omega t_1 + \Omega t_2)\right] E[\sin 2\theta]$$

$$= E(A^2) E[\cos\Omega t_1 \cos\Omega t_2] \int_{-\pi}^{\pi} \cos^2\theta \; f(\theta) \, d\theta$$

$$+ E(A^2) E[\sin\Omega t_1 \sin\Omega t_2] \int_{-\pi}^{\pi} \sin^2\theta \; f(\theta) \, d\theta$$

$$-\frac{1}{2} E(A^2) E\left[\sin(\Omega t_1 + \Omega t_2)\right] \int_{-\pi}^{\pi} \sin 2\theta \; f(\theta) \, d\theta$$

$$= E(A^2) E[\cos\Omega t_1 \cos\Omega t_2] \frac{1}{2\pi} \int_{-\pi}^{\pi} \cos^2\theta \; d\theta$$

$$+ E(A^2) E[\sin\Omega t_1 \sin\Omega t_2] \frac{1}{2\pi} \int_{-\pi}^{\pi} \sin^2\theta \; d\theta$$

$$-\frac{1}{2} E(A^2) E\left[\sin(\Omega t_1 + \Omega t_2)\right] \frac{1}{2\pi} \int_{-\pi}^{\pi} \sin 2\theta \; d\theta$$

But $\int_{-\pi}^{\pi} \cos^2\theta \; d\theta = \int_{-\pi}^{\pi} \sin^2\theta \; d\theta = 1$ and $\int_{-\pi}^{\pi} \sin 2\theta \;\; d\theta = 0$

$$= E\left(A^2\right)\frac{1}{2\pi}\; E\left[\cos\Omega t_1 \cos\Omega t_2 + \sin\Omega t_1 \sin\Omega t_2\right]$$

$$= E\left(A^2\right)\frac{1}{2\pi}\; E\left[\cos\Omega\left(t_1 - t_2\right)\right]$$

Since $E\left(X\left(t\right)\right)$ is a constant and auto correlation is a function of time difference $\left(t_1 - t_2\right)$ whatever be the value of $f_{\Omega}\left(\omega\right)$, $\left\{X\left(t\right)\right\}$ is a wide sense stationary.

Example : Show that the random process $X\left(t\right) = A\cos\left(\omega t + \theta\right)$ is wide sense stationary if A and ω are constants and θ is uniformly distributed random variable in $\left(0,\, 2\pi\right)$.

Since θ is uniformly distributed in $\left(0, 2\pi\right)$, The $p.d.f$ is $f\left(\theta\right) = \frac{1}{2\pi}$, $0 < \theta < 2\pi$

$$\begin{aligned}
E\left[X\left(t\right)\right] &= E\left[A\cos\left(wt + \theta\right)\right] \\
&= A\int_0^{2\pi} \cos\left(wt + \theta\right) \; f\left(\theta\right) \; d\theta \\
&= \frac{A}{2\pi}\int_0^{2\pi} \cos\left(wt + \theta\right) \; d\theta \\
&= \frac{A}{2\pi}\left[\; \sin\left(wt + \theta\right) \;\right]_0^{2\pi} \\
&= \frac{A}{2\pi}\left[\sin\left(2\pi + wt\right) - \sin\omega t\right] \\
&= \frac{A}{2\pi}\left[\sin wt \;-\; s\text{in}\,\omega t\;\right] \\
&= 0, \quad a \quad cons\tan t
\end{aligned}$$

$$\begin{aligned}
E\left[X\left(t_1\right)X\left(t_2\right)\;\right] &= E\left[A\cos\left(wt_1 + \theta\;\right)A\cos\left(wt_2 + \theta\right)\right] \\
&= A^2\int_0^{2\pi} \cos\left(wt_1 + \theta\right)A\cos\left(wt_2 + \theta\right) f\left(\theta\right) \; d\theta \\
&= \frac{A^2}{2\pi}\int_0^{2\pi} \cos\left(wt_1 + \theta\right)\cos\left(wt_2 + \theta\right) d\theta \\
&= \frac{A^2}{4\pi}\int_0^{2\pi} \cos\;\left(\; wt_1 - \; wt_2\;\right) \;+\; \cos\;\left(wt_1 \;+ wt_2 + 2\theta\;\right) \;\; d\theta \\
&= \frac{A^2}{4\pi}\;\cos\left(wt_1 - wt_2\right)\;\left[\;\theta\;\right]_0^{2\pi} \;+\; \frac{A^2}{4\pi}\;\left[\sin\left(wt_1 + wt_2 + 2\theta\;\right)\right]_0^{2\pi} \\
&= \frac{A^2}{4\pi}\left(2\pi\right)\cos\left(wt_1 - wt_2\right) + \frac{A^2}{4\pi}\left[\sin\left(4\pi + wt_1 + wt_2\right) - \sin\left(wt_1 + wt_2\right)\right]
\end{aligned}$$

$$= \frac{A^2}{2}\cos(wt_1 - wt_2) - \frac{A^2}{4\pi}\left[\sin(wt_1 + wt_2) - \sin(wt_1 + wt_2)\right]$$

$$= \frac{A^2}{2} \cos(wt_1 - wt_2), \ a \ function \ of \ (t_1 - t_2)$$

Therefore $\{X(t)\}$ is $w.s.s$.

Example: If $\{X(t)\}$ is a WSS process with auto correlation $R_{XX}(\tau) = A \cdot e^{-\alpha|\tau|}$, determine the second order moment of random variable $X(8) - X(5)$.

$$E\left[X(8) - X(5)\right]^2 = E\left[X^2(8) + X^2(5) - 2X(8)X(5)\right]$$

$$= E\left[X^2(8)\right] + E\left[X^2(5)\right] - 2E\left[X(8)X(5)\right]$$

We know that $E\left[X^2(t)\right] = R_{XX}(0)$. Therefore $E\left[X^2(8)\right] = E\left[X^2(5)\right] = Ae^{-\alpha|0|} = A$

$$E\left[X(8)X(5)\right] = R_{XX}(8-5) = Ae^{-3\alpha}$$

Therefore $E\left[X(8) - X(5)\right]^2 = A + A - 2Ae^{-3\alpha}$

Example: If $\{X(t)\}$ is a WSS process with auto correlation $R_{XX}(\tau)$ and if $Y(t) = X(t+a) - X(t-a)$ show that $R_{YY}(\tau) = 2R_{XX}(\tau) - R_{XX}(\tau + 2a) - R_{XX}(\tau - 2a)$.

$$R_{YY}(\tau) = E\left[Y(t) \cdot Y(t+\tau)\right]$$

$$= E\left[\left(X(t+a) - X(t-a)\right) \cdot \left(X(t+\tau+a) - X(t+\tau-a)\right)\right]$$

$$= E\left[X(t+a)X(t+\tau+a)\right] - E\left[X(t+a)X(t+\tau-a)\right]$$
$$- E\left[X(t-a)X(t+\tau+a)\right] + E\left[X(t-a)X(t+\tau-a)\right]$$

$$= R_{XX}(\tau) - E\left[X(t+a)X(t+\tau-a)\right] - E\left[X(t-a)X(t+\tau+a)\right] + R_{XX}(\tau)$$

$$= R_{XX}(\tau) - R_{XX}(\tau - 2a) - R_{XX}(\tau + 2a) + R_{XX}(\tau)$$

Example : Find the average power of the process $X(t) = 10\cos(100t + \theta)$, where θ is uniformly distributed over $\left(0, \frac{\pi}{2}\right)$.

Since θ is uniformly distributed over $\left(0, \frac{\pi}{2}\right)$, $f(\theta) = \frac{2}{\pi}, \ 0 < \theta < \frac{\pi}{2}$

$$E\left[X^2(t)\right] = E\left[100\cos^2(100t + \theta)\right]$$

$$= 100 \times \frac{1}{2} E\left[1 + \cos 2(100t + \theta)\right]$$

$$= 50 + 50E\left[\cos 2(100t + \theta)\right]$$

$$= 50 + 50 \int_0^{\pi/2} \cos 2(100t + \theta)\ f(\theta)\ d\theta$$

$$= 50 + 50 \times \frac{2}{\pi} \int_0^{\pi/2} \cos 2(100t + \theta)\ \ d\theta$$

$$= 50 + 50 \times \frac{2}{\pi} \left[\frac{\sin 2(100t + \theta)}{2}\right]_0^{\pi/2}$$

$$= 50 + \frac{50}{\pi}\left[\sin 2\left(100t + \frac{\pi}{2}\right) - \sin 2(100t)\right]$$

$$= 50 + \frac{50}{\pi}\left[-\sin 2(100t) - \sin 2(100t)\right]$$

$$= 50 - \frac{100}{\pi} \sin 2(100t)$$

Average Power $P_{XX} = \underset{T \to \infty}{Lt} \frac{1}{2T} \int_{-T}^{T} E\left[X^2(t)\right] dt$

$$= \underset{T \to \infty}{Lt} \frac{1}{2T} \int_{-T}^{T} 50 - \frac{100}{\pi} \sin 2(100t)\ dt$$

$$= \underset{T \to \infty}{Lt} \frac{1}{2T} \int_{-T}^{T} 50\ dt$$

$$= \underset{T \to \infty}{Lt} \frac{1}{2T} \left[502T\right]$$

$$= 50$$

Example : Let $X(t) = A\cos\omega t + B\sin\omega t$ be a random process where A and B are random variables with $E(A) = E(B) = 0$, $\left[E(A)\right]^2 = \left[E(B)\right]^2$ and $E(AB) = 0$. Prove that $\{X(t)\}$ is mean ergodic.

Given $E(A) = E(B) = 0$, $\left[E(A)\right]^2 = \left[E(B)\right]^2 = \sigma^2$, $E(AB) = 0$

$$E(X(t)) = E(A\cos\omega t + B\sin\omega t)$$

$$= E(A)\cos\omega t + E(B)\sin\omega t$$

$$= 0$$

$$\overline{X_T} = \underset{T \to \infty}{Lt} \frac{1}{2T} \int_{-T}^{T} X(t)\ dt$$

$$= \underset{T \to \infty}{Lt} \frac{1}{2T} \int_{-T}^{T} (A\cos\omega t + B\sin\omega t)\ dt$$

$$= \underset{T \to \infty}{Lt} \frac{A}{T} \int_{0}^{T} \cos\omega t\ \ dt$$

$$= \underset{T \to \infty}{Lt} \frac{A}{T} \left[\frac{\sin\omega T}{w}\right]$$

$$= 0$$

Since ensembled mean and time average mean are equal, $\{X(t)\}$ is mean ergodic.

Example : If the WSS process $\{X(t)\}$ is given by $X(t) = 10\cos(100t + \theta)$, where θ is uniformly distributed over $(-\pi, \pi)$. Prove that $\{X(t)\}$ is correlation ergodic.

Since θ is uniformly distributed over $(-\pi, \pi)$, $f(\theta) = \frac{1}{2\pi}, -\pi < \theta < \pi$

We know that $R_{XX}(t_1, t_2) = E\{X(t_1) \cdot X(t_2)\}$

$$= E\left[\left(10\cos(100t_1 + \theta)\right)\left(10\cos(100t_2 + \theta)\right)\right]$$

$$= \frac{100}{2} E\left[\cos(100t_1 + \theta + 100t_2 + \theta) + \cos(100t_1 + \theta - 100t_2 - \theta)\right]$$

$$= 50\ E\left[\cos(100(t_1 + t_2) + 2\theta) + \cos(100(t_1 - t_2))\right]$$

$$= 50\ E\left[\cos(100(t_1 + t_2) + 2\theta)\right] + 50\cos(100(t_1 - t_2))$$

$$E\left[\cos(100(t_1 + t_2) + 2\theta)\right] = \int_{-\pi}^{\pi} \cos(100(t_1 + t_2) + 2\theta) \cdot f(\theta)\ d\theta$$

$$= \frac{2}{2\pi} \int_{0}^{\pi} \cos(100(t_1 + t_2) + 2\theta)\ d\theta$$

$$= \frac{2}{2\pi} \left[\frac{\sin(100(t_1 + t_2) + 2\theta)}{2}\right]_0^{\pi}$$

$$= \frac{1}{2\pi}\left[\sin 100(t_1 + t_2) - \sin 100(t_1 + t_2)\right]$$

$$= 0$$

Therefore $R_{XX}(t_1, t_2) = 50\cos(100(t_1 - t_2))$ ......(1)

Consider $Z_T = \frac{1}{2T}\int_{-T}^{T} X(t_1)\cdot X(t_2)\,dt$

$$Z_T = \frac{1}{2T}\int_{-T}^{T} \left(10\cos(100t_1+\theta)\right)\left(10\cos(100t_2+\theta)\right)\,dt$$

$$= \frac{100}{2T}\int_{-T}^{T} \left(\cos(100t_1+\theta)\right)\left(\cos(100t_2+\theta)\right)\,dt$$

$$= \frac{100}{2\times 2T}\int_{-T}^{T} \cos(100t_1+\theta+100t_2+\theta)+\cos(100t_1+\theta-100t_2-\theta)\,dt$$

$$= \frac{25}{T}\int_{-T}^{T} \cos(100(t_1+t_2)+2\theta)+\cos 100(t_1-t_2)\,dt$$

$$= \frac{25}{T}\int_{-T}^{T} \cos(100(t_1+t_2)+2\theta)\,dt + \frac{25}{T}(2T)\cos 100(t_1-t_2)$$

Now $\underset{T\to\infty}{Lt} Z_T = \underset{T\to\infty}{Lt} \frac{25}{T}\int_{-T}^{T} \cos(100(t_1+t_2)+2\theta)\,dt + \underset{T\to\infty}{Lt} 50\cos 100(t_1-t_2)$

$$= 50\cos 100(t_1-t_2) \text{.....(2)}$$

From (1) and (2), we have $R_{XX}(t_1, t_2) = \underset{T\to\infty}{Lt} \frac{1}{2T}\int_{-T}^{T} X(t_1)\cdot X(t_2)\,dt$

Therefore $\{X(t)\}$ is correlation ergodic.

Poisson Process

If $X(t)$ represents number of occurrences of an event in $(0, t)$ then the random process is $\{X(t)\}$ said to be Poisson process if the following postulates are satisfied.

Postulates of Poisson process.

If $X(t)$ represents the number of occurrences of an event in $(0, t)$ then

Probability of [1 occurrence in $(t, t+\Delta t)$] $= \lambda\Delta t$

Probability of [0 occurrence in $(t, t+\Delta t)$] $= 1-\lambda\Delta t$

Probability of [2 or more occurrences in $(t, t+\Delta t)$] $= 0$

$X(t)$ is independent at any interval

Probability Law for Poisson Process

The probability law for Poisson process is given by $P\left[X(t)=n\right] = \frac{e^{-\lambda t}(\lambda t)^n}{n!}$, $n = 0, 1, \ldots$.

Mean and Variance of Poisson Process

$$E\left[X(t)\right] = \sum_{n=0}^{\infty} n \cdot P\left[X(t)\right]$$

$$= \sum_{n=0}^{\infty} n \cdot \frac{e^{-\lambda t}\ (\lambda t)^{n}}{n!}$$

$$= \sum_{n=0}^{\infty} n \cdot \frac{e^{-\lambda t}\ (\lambda t)^{n}}{n!}$$

$$= e^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\lambda t)^{n}}{(n-1)!}$$

$$= e^{-\lambda t} \left[\frac{(\lambda t)^{1}}{1} + \frac{(\lambda t)^{2}}{1!} + \frac{(\lambda t)^{3}}{2!} + \ldots\ldots\right]$$

$$= e^{-\lambda t}(\lambda t)\left[1 + \frac{(\lambda t)^{1}}{1!} + \frac{(\lambda t)^{2}}{2!} + \ldots\ldots\right]$$

$$= e^{-\lambda t}(\lambda t)\, e^{\lambda t}$$

$$= \lambda t$$

$$E\left[X^{2}(t)\right] = \sum_{n=0}^{\infty} n^{2} \cdot P\left[X(t)\right]$$

$$= \sum_{n=0}^{\infty} \left[n(n-1)+1\right] \cdot P\left[X(t)\right]$$

$$= \sum_{n=0}^{\infty} n(n-1) \cdot \frac{e^{-\lambda t}\ (\lambda t)^{n}}{n!} + \sum_{n=0}^{\infty} n \cdot \frac{e^{-\lambda t}\ (\lambda t)^{n}}{n!}$$

$$= e^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\lambda t)^{n}}{(n-2)!} + \lambda t$$

$$= e^{-\lambda t} \left[\frac{(\lambda t)^{2}}{1} + \frac{(\lambda t)^{3}}{1!} + \frac{(\lambda t)^{4}}{2!} + \ldots\ldots\right] + \lambda t$$

$$= e^{-\lambda t}\ (\lambda t)^{2} \left[1 + \frac{(\lambda t)^{1}}{1!} + \frac{(\lambda t)^{2}}{2!} + \ldots\ldots\right] + \lambda t$$

$$= e^{-\lambda t}\ (\lambda t)^{2}\ e^{\lambda t} + \lambda t$$

$$= (\lambda t)^{2} + \lambda t$$

$$Var\left[X(t)\right] = E\left[X^2(t)\right] - \left\{E\left[X(t)\right]\right\}^2$$

$$= (\lambda t)^2 + \lambda t - (\lambda t)^2$$

$$= \lambda t$$

Autocorrelation of Poisson Process

By definition $R_{XX}(t_1, t_2) = E\left\{X(t_1)\cdot X(t_2)\right\}$

$$= E\left\{X(t_1)\cdot\left(X(t_2) - X(t_1)\right) + X^2(t_1)\right\}$$

$$= E\left[X(t_1)\right]\cdot E\left(X(t_2) - X(t_1)\right) + E\left\{X^2(t_1)\right\}$$

$$= \lambda t_1 \cdot \lambda\left(t_2 - t_1\right) + \left(\lambda t_1\right)^2, \ \ where \ \ t_2 > t_1$$

$$= \lambda^2 t_1 t_2 - \lambda^2 t_1^2 + \lambda^2 t_1^2 + \lambda t_1$$

$$= \lambda^2 t_1 t_2 + \lambda t_1, \quad if \ \ t_2 > t_1$$

$$= \lambda^2 t_1 t_2 + \lambda \min\left(t_1, t_2\right)$$

Properties of Poisson Process

i. Poisson process is Markov process

ii. Sum of two independent Poisson processes is Poisson processes

Proof: Let $X(t), Y(t)$ be two independent random processes with parameter λ and μ respectively and let $Z(t) = X(t) + Y(t)$

$$P\left[Z(t) = n\right] = \sum_{k=o}^{n} P\left[X(t) = k\right]\cdot P\left[Z(t) = n-k\right]$$

$$= \sum_{k=0}^{n} \frac{e^{-\lambda t}(\lambda t)^n}{k!}\cdot\frac{e^{-\mu t}(\mu t)^{n-k}}{(n-k)!}$$

$$= \frac{e^{-(\lambda+\mu)t}}{n!}\sum_{k=0}^{n} \frac{n!}{k!(n-k)!}\cdot(\lambda t)^n(\mu t)^{n-k}$$

$$= \frac{e^{-(\lambda+\mu)t}}{n!}(\lambda t + \mu t)^n$$

$$= \frac{e^{-(\lambda+\mu)t}\left[(\lambda+\mu)t\right]^n}{n!}$$

Hence $Z(t)$ is a Poisson process with parameter $(\lambda + \mu)$.

iii. Difference of two independent Poisson processes is not a Poisson processes

Proof: Let $X(t)$, $Y(t)$ be two independent random processes with parameter λ and μ respectively and let $Z(t)=X(t)-Y(t)$

$$E[Z(t)]=E[X(t)-Y(t)]$$

$$=\lambda(t)-\mu(t)$$

$$=(\lambda-\mu)(t)$$

Also $E[Z^2(t)]=E[X^2(t)+Y^2(t)-2X(t)Y(t)]$

$$=E[X^2(t)]+E[Y^2(t)]-2E[X(t)Y(t)]$$

$$=(\lambda t)^2+\lambda t+(\mu t)^2+\mu t-2\lambda t\,\mu t$$

$$=(\lambda+\mu)t+[(\lambda-\mu)t]^2$$

Here $E[Z^2(t)]$ is not of the form $(\lambda-\mu)t+[(\lambda-\mu)t]^2$.

Hence $Z(t)$ is not a Poisson process with parameter $(\lambda-\mu)$.

Note: If the number of occurrence of an event in an interval of length t is a Poisson process with parameter λ and if each occurrence of the event has a constant probability, then the number $N(t)$ of the recorded occurrences in t is also Poisson process with parameter λp.

i.e. $P[N(t)=n] = \dfrac{e^{-\lambda pt}\ (\lambda pt)^n}{n!}$, $n=0,1,\ldots.$

Example: An office receives 3 calls per minute on an average and it follows Poisson process. What is the probability of receiving (a) no calls in a one minute interval (b) at most 3 calls in a 5 minute interval?.

Given λ = 3 / min. Probability law for Poisson process is $P[X(t)=n]=\dfrac{e^{-\lambda t}(\lambda t)^n}{n!}$

a. Probability of receiving no calls $(n=0)$ in a one minute interval $(t=1)$

$$P[X(1)=0]=\frac{e^{-3}(3)^0}{0!} = e^{-3}$$

ii. Probability of receiving at most 3 calls $(n\le 3)$ in a five minute interval $(t=5)$

$$P[X(5)\le 3] = P[X(5)=0] + P[X(5)=1] + P[X(5)=2] + P[X(5)=3]$$

$$= \frac{e^{-15}\ (15)^0}{0\ !} + \frac{e^{-15}\ (15)^1}{1\ !} + \frac{e^{-15}\ (15)^2}{2\ !} + \frac{e^{-15}\ (15)^3}{3\ !}$$

$$= e^{-15}\left[1 + 15 + \frac{15^2}{2} + \frac{15^3}{6}\right]$$

Example: A radioactive source emits particles at a rate of 5 per minute according to a Poisson process. Each particle emitted has a probability 0.6 of being recorded. Find the probability that 10 particles are recorded in a 4 minute period.

Given $\lambda = 5$ / min Probability of recording $p = 0.6$

Probability law for Poisson process is $P[X(t) = n] = \dfrac{e^{-\lambda pt}(\lambda pt)^n}{n!}$

i. Probability of recording 10 particles $(n = 10)$ in four minute interval $(t = 4)$

$$P[X(4) = 10] = \frac{e^{-12}(12)^{10}}{10!} = 0.104$$

Example . If customers arrive at a counter in accordance with Poisson process with a mean rate of 2 per minute, find the probability that the interval between 2 consecutive arrivals is (a) more than 1 minute (b) between 1 minute and 2 minute and (c) 4 minute or less.

Let T be the random variable denoting inter arrival time. By a property it follows exponential distribution. Therefore $f(t) = \lambda e^{-\lambda t} = 2e^{-2t}$, $0 < t < \infty$.

(a) $P(T > 1) = \int_1^\infty 2\, e^{-2t}\, dt = 2\left[\dfrac{e^{-2t}}{-2}\right]_1^\infty = e^{-2} = 0.1353$

(b) $P(1 < T < 2) = \int_1^2 2\, e^{-2t}\, dt = 2\left[\dfrac{e^{-2t}}{-2}\right]_1^2 = -e^{-4} + e^{-2} = 0.1169$

(b) $P(T < 4) = \int_0^4 2\, e^{-2t}\, dt = 2\left[\dfrac{e^{-2t}}{-2}\right]_0^4 = -e^{-8} + 1 = 0.9996$

Markov Process

A random process in which the future value depends only on the present value and not on the past values, is called a Markov process.

Results

- If for all n, $P(X_n = a_n / X_{n-1} = a_{n-1},\, X_{n-2} = a_{n-2},\, \ldots,\, X_0 = a_0) = P(X_n = a_n / X_{n-1} = a_{n-1})$
- The process $(X_n : n = 0,1,2,\ldots)$ is called a Markov chain and $a_0, a_1, a_2, \ldots, a_n$ are called states of the Markov chain
- $P(X_n = a_j / X_{n-1} = a_i)$ is called the one step transition probability from state a_i to state a_j in the n^{th} step.
- The conditional probability $P(X_n = a_j / X_{n-1} = a_i) = P_{ij}^{(n)}$ is called the n step transition probability.

- $p^{(0)} = \left(p_1^{(0)},\ p_2^{(0)},\ p_3^{(0)}, \ldots,\ p_n^{(0)} \right)$ is called the initial probability distribution where $p_1^{(0)} = p\left[X_0 = 1\right],\ p_2^{(0)} = p\left[X_0 = 2\right]$........ are the initial probabilities for states $1, 2, 3, \ldots$.

Definitions

If P is the transition probability matrix of a homogeneous Markov chain, then the n step tpm $P^{(n)}$ is equal to P^n. i.e. $\left[p_{ij}^{(n)} \right] = \left[p_{ij} \right]^n$.

A stochastic matrix P is said to be regular if all the entries of P^m are positive for some positive integer m. Also a homogeneous Markov chain is regular if its tpm is regular.

If P is the tpm of the regular chain and π is the steady state distribution, then $\pi P = \pi$.

If for every $i,\ j$, we can find some n such that $p_{ij}^{(n)} > 0$, then every state can be reached from every other state and the Markov chain is said to be irreducible. Otherwise it is reducible.

The state i if a Markov chain is called a return state, if $p_{ii}^{(n)} > 0$ for some $n > 1$.

The period of a return state is the GCD of all m such that $p_{ii}^{(m)} > 0$.

The probability that the chain returns to state i, having started from state i for the first time at the n^{th} step is denoted by $f_{ii}^{(n)}$ and is called the first return time probability.

If $\sum f_{ii}^{(n)} = 1$, the return state i is said to be persistent or recurrent. Otherwise it is said to be transient.

Let $\mu_{ii} = \sum n \cdot f_{ii}^{(n)}$. If μ_{ii} is finite, the state i is non-null persistent, otherwise it is null persistent.

A non-null persistent and a periodic state are called ergodic.

If a Morkov chain is irreducible, all state are of the same type.

Example: The transition probability matrix of a Markov chain with three states 0, 1, 2 is $P = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{3}{4} & \frac{1}{4} \end{pmatrix}$ **and the initial state distribution of the chain is** $p\left[X_0 = i\right] = \frac{1}{3},\ i = 0, 1, 2.$**.**

Find $(i)\ p\left[X_0 = 2\right]$ $(ii)\ p\left[X_3 = 1,\ X_2 = 2,\ X_1 = 1,\ X_0 = 2\right]$ $(iii)\ p\left[X_2 = 1,\ X_0 = 0\right]$

$$\begin{array}{l} \\ \text{Given} \end{array}\qquad P = State\ X_{n-1}\begin{array}{c} State\ X_n \\ \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{pmatrix} \frac{3}{4} & \frac{1}{4} & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{3}{4} & \frac{1}{4} \end{pmatrix} \end{array} \quad and \quad P^{(2)} = P^2 = State\ X_{n-1}\begin{array}{c} State\ X_n \\ \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{pmatrix} \frac{5}{8} & \frac{5}{16} & \frac{1}{6} \\ \frac{5}{16} & \frac{8}{16} & \frac{3}{16} \\ \frac{3}{16} & \frac{9}{16} & \frac{4}{16} \end{pmatrix} \end{array}$$

By the definition of conditional probability

$(i)\ \ p[X_0 = 2] = \sum p[X_2 = 2 / X_0 = i] \cdot p[X_0 = i]$

$= p[X_2 = 2 / X_0 = 0] \cdot p[X_0 = 0] + p[X_2 = 2 / X_0 = 1] \cdot p[X_0 = 1] + p[X_2 = 2 / X_0 = 3] \cdot p[X_0 = 3]$

$= p_{02}^{(2)} \cdot p[X_0 = 0] + p_{12}^{(2)} \cdot p[X_0 = 1] + p_{22}^{(2)} \cdot p[X_0 = 3]$

$= \frac{1}{16} \cdot \frac{1}{3} + \frac{3}{16} \cdot \frac{1}{3} + \frac{4}{16} \cdot \frac{1}{3}$

$= \frac{1}{6}$

$(ii)\ \ p[X_3 = 1,\ X_2 = 2,\ X_1 = 1,\ X_0 = 2]$

$= p[X_3 = 1 / X_2 = 2] \cdot p[X_2 = 2 / X_1 = 1] \cdot p[X_1 = 1 / X_0 = 2] \cdot p[X_0 = 2]$

$= p_{23}^{(1)} \cdot p_{12}^{(1)} \cdot p_{21}^{(1)} \cdot p[X_0 = 2]$

$= \frac{3}{4} \cdot \frac{1}{4} + \frac{3}{4} \cdot \frac{1}{3}$

$= \frac{3}{64}$

$(iii)\ \ p[X_2 = 1,\ \ X_0 = 0] = p[X_2 = 1 / X_0 = 0] \cdot p[X_0 = 0]$

$= p_{01}^{(2)} \cdot p[X_0 = 0]$

$= \frac{5}{16} \cdot \frac{1}{3}$

$= \frac{5}{48}$

Example: A house wife buys 3 kinds of cereals A, B and C. She never buys the same cereal In Successive weeks. If she buys cereal A, the next day she buys B. However if she buys B Or C, the next week she is 3 times as likely to buy A as the other cereal. In the long run, how often does she buy each of the three cereals?

The transition probability matrix of the process is

$$P = \begin{array}{c} \\ A \\ B \\ C \end{array} \begin{array}{c} \begin{array}{ccc} A & B & C \end{array} \\ \begin{pmatrix} 0 & 1 & 0 \\ 3/4 & 0 & 1/4 \\ 3/4 & 1/4 & 0 \end{pmatrix} \end{array}$$

Let $\pi = (\pi_1 \quad \pi_2 \quad \pi_3)$ be the steady state distribution of the Markov chain.

Then

$$(\pi_1 \quad \pi_2 \quad \pi_3) \begin{pmatrix} 0 & 1 & 0 \\ 3/4 & 0 & 1/4 \\ 3/4 & 1/4 & 0 \end{pmatrix} = (\pi_1 \quad \pi_2 \quad \pi_3)$$

$$0\ \pi_1 + \frac{3}{4}\ \pi_2 + \frac{3}{4}\ \pi_3 = \pi_1 \quad -------(1)$$

$$1\ \pi_1 + 0\ \pi_2 + \frac{1}{4}\ \pi_3 = \pi_2 \quad -------(2)$$

$$0\ \pi_1 + \frac{1}{4}\ \pi_2 + 0\ \pi_3 = \pi_3 \quad -------(3)$$

$$and \quad \pi_1 + \pi_2 + \pi_3 = 1 \quad -------(4)$$

Solving (1), (2), (3), (4) , we get $\pi_1 = \frac{15}{35}, \quad \pi_2 = \frac{16}{35}, \quad \pi_3 = \frac{4}{35}$

$\therefore$ In the long run, probability of buying A = 15/35

$\therefore$ In the long run, probability of buying B = 16/35

$\therefore$ In the long run, probability of buying C = 4/35

Example: The probability of a dry day following a rainy day is 1/3 and that the probability of a rainy day following a dry day is 1/2. Given that May 1st is a dry day. Find the probability that May 3rd is a dry day and also May 5th is a dry day.

Here the states are Dry day(0) and Rainy day(1).

Given that $P_{01} = \frac{1}{2}$ and $P_{10} = \frac{1}{3}$. Therefore the $t.p.m$ is

$$P = M1 \begin{array}{c} \\ \\ 0 \\ 1 \end{array} \begin{array}{c} M2 \\ \begin{array}{cc} 0 & 1 \end{array} \\ \begin{bmatrix} 1/2 & 1/2 \\ 1/3 & 2/3 \end{bmatrix} \end{array} \qquad P^2 = M2 \begin{array}{c} \\ \\ 0 \\ 1 \end{array} \begin{array}{c} M3 \\ \begin{array}{cc} 0 & 1 \end{array} \\ \begin{bmatrix} 5/12 & 7/12 \\ 7/18 & 11/18 \end{bmatrix} \end{array}$$

$$
P^3 = M3 \quad \begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 29/72 & 43/72 \\ 1 & 43/108 & 65/108 \end{array} \qquad P^4 = M4 \quad \begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 173/432 & 259/432 \\ 1 & 259/648 & 389/648 \end{array}
$$

(column headings: $M4$ over P^3, $M5$ over P^4)

Probability for May 3 is a dry day is $P_{00}^{(2)} = \dfrac{5}{12}$

Probability for May 5 is a dry day is $P_{00}^{(4)} = \dfrac{173}{432}$

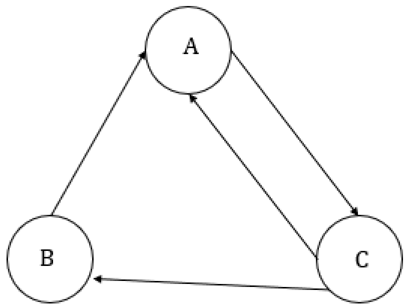
Example: Consider a Markov chain having the transition probability matrix

$$
P = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 1/2 & 1/2 & 0 \end{pmatrix}
$$

a. Show that the chain is irreducible b. Find the stationary distribution

The transition probability matrix of the process is

$$
P = \begin{array}{c} \\ A \\ B \\ C \end{array} \begin{array}{c} \begin{array}{ccc} A & B & C \end{array} \\ \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 1/2 & 1/2 & 0 \end{pmatrix} \end{array}
$$



Let $\pi = (\pi_1 \quad \pi_2 \quad \pi_3)$ be the steady state distribution of the Markov chain.

Then

$$
(\pi_1 \quad \pi_2 \quad \pi_3) \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 1/2 & 1/2 & 0 \end{pmatrix} = (\pi_1 \quad \pi_2 \quad \pi_3)
$$

$$0\ \pi_1 + \pi_2 + \frac{1}{2}\pi_3 = \pi_1 \quad -------(1)$$

$$0\ \pi_1 + 0\ \pi_2 + \frac{1}{2}\pi_3 = \pi_2 \quad -------(2)$$

$$1\ \pi_1 + 0\ \pi_2 + 0\ \pi_3 = \pi_3 \quad -------(3)$$

$$and \quad \pi_1 + \pi_2 + \pi_3 = 1 \quad -------- \ (4)$$

Solving (1), (2), (3), (4) , we get $\pi_1 = \frac{2}{5}, \quad \pi_2 = \frac{1}{5}, \quad \pi_3 = \frac{2}{5}$

$\therefore$ In the long run, probability of buying A = 15/35

$\therefore$ In the long run, probability of buying B = 16/35

$\therefore$ In the long run, probability of buying C = 4/35

$$P^2 = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \\ 1/2 & 0 & 1/2 \end{pmatrix} \qquad P^3 = \begin{pmatrix} 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \\ 1/4 & 1/4 & 1/2 \end{pmatrix}$$

$$P^4 = \begin{pmatrix} 1/4 & 1/4 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/4 & 1/4 \end{pmatrix} \qquad P^5 = \begin{pmatrix} 1/2 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/2 \\ 3/8 & 1/8 & 1/2 \end{pmatrix}$$

Since all $p_{ij} > 0$ *for* some i, j and each state is reachable from the other state, the chain is irreducible.

Example . There are 2 white marbles in $urn\ A$ and 3 red marbles in $urn\ B$. At each step of the process, a marble is selected from each urn and the 2 marbles selected are interchanged. The state of the related Markov chain is the number of red marbles in $urn\ A$ after the interchange. What is the probability that there are 2 red marbles in $urn\ A$ after 3 steps? In the long run, what the probability that there are 2 red marbles in $urn\ A$?

The Markov chain $\{X_n\}$ has state space 0, 1, 2 since the number of marbles in $urn\ A$ is always 2 and the number of red marbles may be 0,1,2.

The $t.p.m$ of the chain is

$$\begin{array}{cc} & (X_{n+1}) \\ & \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ P = (X_n) \begin{array}{c} 0 \\ 1 \\ 2 \end{array} & \begin{pmatrix} P_{00} & P_{01} & P_{02} \\ P_{10} & P_{11} & P_{12} \\ P_{20} & P_{21} & P_{22} \end{pmatrix} \end{array}$$

If $X_n = 0$ i.e., if the system is at state 0. $urn\ A$ contains 2 while marbles and $urn\ B$ contains 3 red marbles

After one interchange, $urn\ A$ has no red marbles, which is not true. Then $P_{00} = 0$

After one interchange, $urn\ A$ definitely contains 1 red marbles. $\therefore$ $P_{01} = 1$

After one interchange, $urn\ A$ definitely contains 2 red marbles. $\therefore\ P_{02}=0$

If $X_n = 1$ i.e., if the system is at state 1, $urn\ A$ contains 1 red and 1 white and $urn\ B$ contains 1 White and 2 Red marbles.

$P_{10} = P(X_{n+1}=1/X_n=0) = \frac{1}{2}\times\frac{1}{3} = \frac{1}{6}$ (one red at $urn\ A$ becomes no red after interchange)

$P_{12} = P(X_{n+1}=2/X_n=1) = \frac{2}{3}\times\frac{1}{2} = \frac{1}{3}$ (one red at $urn\ A$ becomes two red after interchange)

But $P_{10}+P_{11}+P_{12}=1.$ $\therefore$ $P_{11}=1-P_{12}-P_{10}=1-\frac{1}{6}-\frac{1}{3}=\frac{1}{2}$ (one red at $urn\ A$ becomes one red after interchange)

If $X_n = 2$ i.e., if the system is at state 2, $urn\ A$ contains 2 red and $urn\ B$ contains 2 White and 1 Red marbles.
After one interchange, $urn\ A$ definitely contains 0 red marbles. $\therefore\ P_{20}=0$

$P_{21} = P(X_{n+1}=1/X_n=2) = \frac{2}{3}\times\frac{2}{2} = \frac{2}{3}$ (two red at $urn\ A$ becomes one red after interchange)

But $P_{20}+P_{21}+P_{22}=1.$ $\therefore$ $P_{22}=1-P_{21}-P_{20}=1-\frac{2}{3}-0=\frac{1}{3}$ (two red at $urn\ A$ becomes two red after interchange)

Therefore the $t.p.m$ is $P = \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \\ o & \frac{2}{3} & \frac{1}{3} \end{pmatrix}$

The initial distribution is $\begin{array}{c} \quad\ 0 \quad 1 \quad 2 \\ P^{(0)} = \begin{pmatrix}1 & 0 & 0\end{pmatrix} \end{array}$. (There is no red marble initially, and it is sure event)

$$P^{(1)} = P^{(0)}P = \begin{pmatrix}1 & 0 & 0\end{pmatrix}\begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \\ o & \frac{2}{3} & \frac{1}{3} \end{pmatrix} = \begin{pmatrix}0 & 1 & 0\end{pmatrix}$$

$$P^{(2)} = P^{(1)}P = \begin{pmatrix}0 & 1 & 0\end{pmatrix}\begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \\ o & \frac{2}{3} & \frac{1}{3} \end{pmatrix} = \begin{pmatrix}\frac{1}{6} & \frac{1}{2} & \frac{1}{3}\end{pmatrix}$$

$$P^{(3)} = P^{(2)}P = \begin{pmatrix} \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \\ 0 & \frac{2}{3} & \frac{1}{3} \end{pmatrix} = \begin{pmatrix} \frac{1}{12} & \frac{23}{36} & \frac{5}{18} \end{pmatrix}$$

Therefore $P(\text{2 red marbles in } urn\ A \text{ after 3 interchanges}) = \frac{5}{18}$

The long run probability is limiting probability.

Let $\pi = \begin{pmatrix} \pi_0 & \pi_1 & \pi_2 \end{pmatrix}$ where $\pi_0 + \pi_1 + \pi_2 = 1 \cdots\cdots (1)$

Also $\pi P = \pi$ gives

$$\begin{pmatrix} \pi_0 & \pi_1 & \pi_2 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \\ 0 & \frac{2}{3} & \frac{1}{3} \end{pmatrix} = \begin{pmatrix} \pi_0 & \pi_1 & \pi_2 \end{pmatrix}$$

$$0\pi_0 + \frac{1}{6}\pi_1 + 0\pi_1 = \pi_0 \cdots\cdots (2)$$

$$1\pi_0 + \frac{1}{2}\pi_1 + \frac{2}{3}\pi_2 = \pi_1 \cdots\cdots (3)$$

$$0\pi_0 + \frac{1}{3}\pi_1 + \frac{1}{3}\pi_2 = \pi_2 \cdots\cdots (4)$$

Solving (1), (2), (3), (4) , we get $\pi_0 = \frac{1}{10}, \quad \pi_1 = \frac{6}{10}, \quad \pi_2 = \frac{3}{10}$

Therefore $P(\text{2 red marbles in } urn\ A \text{ after long run of interchanges}) = \frac{3}{10}$

Exercise

1. Show that the process $X(t) = A\cos\lambda t + B\sin\lambda t$ where A and B are random variables, is wide sense stationary process if $E(A) = E(B) = E(AB) = 0$ and $E\left(A^2\right) = E\left(B^2\right)$.

2. Customers arrive at a bank according to Poisson process with a mean rate of 3 per minute. Find the probability that during a time interval of 2 minutes (i) exactly 4 customers arrive (ii) less than 4 customers arrive (iii) more than 4 customers arrive.

3. Two boys and two girls are throwing a ball one to the other. Each boys throws the ball to the other boy with probability 0.5 and to each girl with probability 0.25. On the other hand each girls throws the ball to each boy with probability 0.5 and never to the other girl. In the long run, how often does each receive the ball?

ANSWERS

1. $E\left(X(t)\right) = 0$, $R_{XX}(t_1, t_2) = 2\cos\lambda(t_2 - t_1)$

2. (i) 0.1338, , $(ii)\ e^{-6}[1+6+18+36]$, $(iii)\ 1-e^{-6}[1+6+18+36]$

3. $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{6}, \frac{1}{6}\right)$

Unit-IV Vector Spaces

Introduction

Linear algebra helps to design and analyze MIMO systems used in wireless communication to increase data rates and transmission reliability. MIMO channels are modeled as matrices where each entry represents a signal path. It is used in control systems for state-space modeling, where the system is described using matrix equations to analyze stability and system dynamics. It is also a foundational in coding theory for error correction in digital communication systems. The main topics in linear algebra is vector spaces.

In engineering, a vector is characterized by two quantities (length and direction) and is represented by a directed line segment. A vector in the plane is represented geometrically by a directed line segment whose initial point is the origin $(0,0)$ and whose terminal point is the point (x_1, x_2). This vector is represented by the ordered $x=(x_1, x_2)$, the pair represented its terminal point.

Note:

- Vectors are denoted by lower case letters in bold type.
- Co-ordinates x_1 and x_2 are called components of the vector x.
- Two vectors $x=(x_1, x_2)$ and $y=(y_1, y_2)$ are equal if $x_1 = x_2$ and $y_1 = y_2$.

Standard Operations on Vectors

Vector Addition: To add two vectors $x=(x_1, x_2)$ and $y=(y_1, y_2)$ in the plane, add their corresponding components. i.e., $x+y=(x_1, x_2)+(y_1, y_2)=(x_1+y_1\ ,\ x_2+y_2)$.

Note:
- The difference of x and y is defined as $x-y=x+(-y)$.
- The vector $-y$ is called the negative of y.

Scalar Multiplication: To multiply a vector $x=(x_1, x_2)$ by a scalar c, multiply each of the components of the vector by c. i.e., $cx=c(x_1, x_2)=(cx_1, cx_2)$.

Note:
- for a scalar c, the vector cx will be $|c|$ times as long as x.
- If c is positive, then cx and x have the same direction.
- If c is negative, then cx and x have opposite directions.

https://doi.org/10.5281/zenodo.15090328

Example: If $x=(3,\ 4)$ and $y=(2,\ -3)$ find $(i)\ x-y \quad (ii)\ \frac{1}{2}x \quad (iii)\ x+\frac{1}{2}y$

$(i)\ x-y=x+(-y)=(3,\ 4)+(-2,\ 3)=(3-2,\ 4+3)=(1,\ 7)$

$(ii)\ \frac{1}{2}x=\frac{1}{2}(3,\ 4)=\left(\frac{3}{2},\ \frac{4}{2}\right)$

$(iii)\ x+\frac{1}{2}y=(3,\ 4)+\frac{1}{2}(-2,\ 3)=(3,\ 4)+\left(-1,\ \frac{3}{2}\right)=\left(3-1,\ 4+\frac{3}{2}\right)=\left(2,\ \frac{11}{2}\right)$

Properties of Vector Operations

Let x, y and z be vectors in the plane and let α and β be scalars. Then

Closure under addition $x+y$ is a vector in the plane	**Commutative property of addition** $x+y=y+x$
Associative property under addition $(x+y)+z=x+(y+z)$	**Additive Identity** $x+0=x$
Additive Inverse $x+(-x)=0$	**Closure under scalar multiplication** αx is a vector in the plane
Distributive Property $\alpha(x+y)=\alpha x+\alpha y$	**Distributive Property** $(\alpha+\beta)x=\alpha x+\beta x$
Associative property of multiplication $(\alpha\beta)x=\alpha(\beta x)$	**Multiplicative Identity** $1(x)=x$
Additive inverse is unique If $x+y=0$, then $x=-y$	**Additive Identity is unique** If $x+y=x$, then $y=0$
$0x=0$ (first one scalar zero, next one vector zero)	$\alpha 0=0$
If $\alpha x=0$, then $\alpha=0$ (scalar) or $x=0$ (vector)	$-(-x)=x$

Note: The proof of each property is a direct application of the definition of vector addition and scalar multiplication combined with the corresponding properties of addition and multiplication of real numbers.

For example, Consider

$$\begin{aligned} 0 &= 0x+(-0x) \\ &= (0+0)x+(-0x) \\ &= 0x+0x+(-0x) \\ &= 0x+(0x+(-0x)) \\ &= 0x+0 \\ &= 0x \end{aligned}$$

For example, Consider

$$\begin{aligned} 0 &= \alpha 0+(-\alpha 0) \\ &= \alpha(0+0)+(-\alpha 0) \\ &= \alpha 0+\alpha 0+(-\alpha 0) \\ &= \alpha 0+(\alpha 0+(-\alpha 0)) \\ &= \alpha 0+0 \\ &= \alpha 0 \end{aligned}$$

Cancellation Law for vector addition: If x, y and z are vectors in a vector space V such that $x+z=y+z$ then $x=y$.

Given $x+z=y+z$

We know that for any vector z in V there exists a vector w in V such that $z+w=0$.

Also by additive property
$$\begin{aligned} x &= x+0 \\ &= x+(z+w) \\ &= (x+z)+w \\ &= (y+z)+w \\ &= y+(z+w) \\ &= y+0 \\ &= y \end{aligned}$$

Vectors in R^n.

Like the vectors defined in the plane R^2 (2-space), we can define the vectors in other spaces also.

The vectors in R^3 (3-space) is of the form $x=(x_1,\ x_2,\ x_3)$

The vectors in R^4 (4-space) is of the form $x=(x_1,\ x_2,\ x_3,\ x_4)$

In the same way, the vectors in R^n (n-space) is of the form $x=(x_1,\ x_2,\ \cdots\cdots,x_n)$

Note: The vector operations and its properties remain holds good in the above spaces also.

Example: Let $x=(3,4,1)$, $y=(-2,0,1)$ and $z=(0,-3,4)$ be vectors in R^4. Solve for x if $x+y=2z-x$.

$$2x=2z-y$$

$$x=\frac{1}{2}(2z-y)$$

$$x=\frac{1}{2}(2(0,-3,4)-(-2,0,1))$$

$$x=\frac{1}{2}((0,-6,8)+(2,0,-1))$$

$$x=\frac{1}{2}(0+2,-6+0,8-1)$$

$$x=\frac{1}{2}(2,-6,7)$$

$$x=\left(1,-3,\frac{7}{2}\right)$$

Definition: A vector x is said to be a linear combination of the vectors $x_1,\ x_2,\ \cdots\cdots,x_n$ if it can be expressed as the sum of scalar multiples of the vectors. i.e. $x=\alpha_1 x_1+\alpha_2\, x_2+\cdots\cdots+\alpha_n x_n$.

Example: Show that $x=(10,1,4)$ is a linear combination of the vectors $u=(2,3,5)$, $v=(1,2,4)$ and $w=(-2,2,3)$.

Since x is a linear combination of the vectors $u,\ v$ and w then there exists constants $a,\ b$ and c such that $x=au+\ bv+cw$.

Therefore $(10,\ 1,\ 4)=a(2,\ 3,\ 5)+b(1,\ 2,\ 4)+c(-2,\ 2,\ 3)$

Equating the corresponding components, we get

$$2a+1b-2c=10$$
$$3a+2b+2c=1$$
$$5a+4b+3c=4$$

Solving the simultaneous equations, we have $a=1,\ b=2,\ c=-3$
Therefore x can be written as a linear combination of $u,\ v,\ w$ as $x=u+2v-3w$.

Example: If $x=(-1,-2-2)$ is a linear combination of the vectors $u=(0,1,4)$, $v=(-1,1,2)$ and $w=(3,1,2)$ find the scalars $a,\ b$ and c.
Given that $x=au+bv+cw$
Therefore $(-1,-2-2)=a(0,1,4)+b(-1,1,2)+c(3,1,2)$
Equating the corresponding components, we get

$$0a-1b+3c=-1$$
$$1a+1b+1c=-2$$
$$4a+2b+2c=-2$$

Solving the simultaneous equations (by Crammer's rule), we have $a=1,\ b=-2,\ c=-1$
Therefore x can be written as a linear combination of $u,\ v,\ w$ as $x=u-2v-w$.

Example: Is the vector $(1,1)$ a linear combination of the vectors $(1,2)$ and $(-2,-4)$?

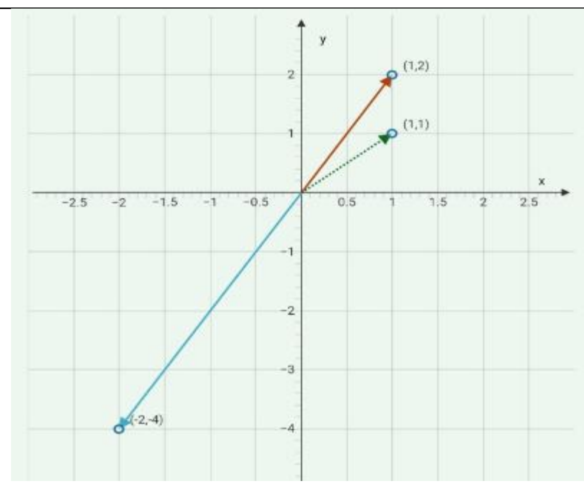
Suppose the vector $(1,1)$ a linear combination of the vectors $(1,2)$ and $(-2,-4)$. Then

$(1,1)=a(1,2)+b(-2,-4)$
Equating the corresponding components, we get

$$a-2b=1$$
$$2a-4b=1$$

Since the equations are inconsistent, it has no solution and hence $(1,1)$ is not a linear combination of the given vectors.

Try this: Is the vector $(1,1)$ a linear combination of the vectors $(1,2)$ and $(2,1)$?



The vectors $(1,2)$ and $(-2,-4)$ form a straight line. So any combination of these two vectors lies on the same line only. But $(1,1)$ is lying away from the line. Therefore it is not the scalar multiple of the given vectors.

Let us recall the following rules in solving the system of equations in matrices

AX=0 |A|≠0 Trivial Solution – Independent $\left[r(A)=3=no.\ of\ \ vectors\right]$	AX=B |A|≠0 Non trivial Solution - Unique $\left[r(A)=3=no.\ of\ \ vectors\right]$
AX=0 |A|=0 Non trivial Solution – Dependent $\left[r(A)=2<no.\ of\ \ vectors\right]$	AX=B |A|=0 No Solution $\left[r(A)=2<no.\ of\ \ vectors\right]$

Vector Spaces

Definition: Let V be a set of vectors together with vector addition and scalar multiplication. Then V is said to be a vector space if it satisfies the following conditions for all $x, y, w \in V$ and any scalars $\alpha, \beta \in R$.

Vector Addition	Scalar Multiplication
$x+y$ is in V	αx is in V
$x+y=y+x$ (commutative law)	$\alpha(x+y)=\alpha x+\alpha y$
$(x+y)+z=x+(y+z)$ (associative law)	$(\alpha+\beta)x=\alpha x+\beta x$
For all $x \in V$ there exists a unique $0 \in V$ such that $x+0=x$ (additive identity)	$(\alpha\beta)x=\alpha(\beta x)$
For all $x \in V$ there exists a unique $y \in V$, denoted by $y=-x$, such that $x+y=0$ (additive inverse)	$1(x)=x$ (scalar identity)

Example:

- The set of all ordered pairs of real numbers R^2 with the standard operations is a vector space.

- The set of all ordered n-tuples of real numbers R^n with the standard operations is a vector space
- The set of all $m \times n$ matrices with the operations of matrix addition and scalar multiplication is a vector space.
- The set of all polynomials of degree 2 or less is a vector space.

Note:

- The set of second degree polynomials is not a vector space. Because if $p(x) = x^2 - 2x + 3, \quad q(x) = -x^2 - 2x + 3$ then $p(x) + q(x) = -4x + 6$ which is not a second degree polynomial.
- Consider the set of all ordered pairs of real numbers, with the scalar multiplication as $\alpha(x_1, x_2) = (0, \alpha x_2)$. It is not a vector space because $1(1,1) = (0,1) \neq (1,1)$.
- The set of integers is not a vector space because $\frac{1}{2}(5)$ is not an integer

Try this: Are these sets a vector space? (i) The set of 2×2 all singular matrices with the standard operations (No) (ii) The set of all 2×2 diagonal matrices with the standard operations.(Yes).

Example: Show that $V = \{a + b\sqrt{2} : a,\, b \in Q\}$ is a vector space over Q under usual addition and scalar multiplication.

Let $V_1 = a_1 + b_1\sqrt{2},\ V_2 = a_2 + b_2\sqrt{2},\ \ V_3 = a_3 + b_3\sqrt{2}$ where $a_1,\, b_1,\, a_2,\, b_2,\, a_3,\, b_3 \in Q$.

(i) $V_1 + V_2 = \left(a_1 + b_1\sqrt{2}\right) + \left(a_2 + b_2\sqrt{2}\right) = \left((a_1 + a_2) + (b_1 + b_2)\sqrt{2}\right) = \left(k_1 + k_2\sqrt{2}\right) \in V$

(ii) $V_1 + V_2 = \left(a_1 + b_1\sqrt{2}\right) + \left(a_2 + b_2\sqrt{2}\right) = \left(a_2 + b_2\sqrt{2}\right) + \left(a_1 + b_1\sqrt{2}\right) = V_2 + V_1$

(iii) $V_1 + V_2 = \left(a_1 + b_1\sqrt{2}\right) + \left(a_2 + b_2\sqrt{2}\right) = \left((a_1 + a_2) + (b_1 + b_2)\sqrt{2}\right)$

$V_2 + V_3 = \left(a_2 + b_2\sqrt{2}\right) + \left(a_3 + b_3\sqrt{2}\right) = \left((a_2 + a_3) + (b_2 + b_3)\sqrt{2}\right)$

$$(V_1+V_2)+V_3=\left[\left(a_1+b_1\sqrt{2}\right)+\left(a_2+b_2\sqrt{2}\right)\right]+\left(a_3+b_3\sqrt{2}\right)$$

$$=\left((a_1+a_2)+(b_1+b_2)\sqrt{2}\right)+\left(a_3+b_3\sqrt{2}\right)$$

$$=\left((a_1+a_2+a_3)+(b_1+b_2+b_3)\sqrt{2}\right)$$

$$V_1+(V_2+V_3)=\left(a_1+b_1\sqrt{2}\right)+\left[\left(a_2+b_2\sqrt{2}\right)+\left(a_3+b_3\sqrt{2}\right)\right]$$

$$=\left(a_1+b_1\sqrt{2}\right)+\left((a_2+a_3)+(b_2+b_3)\sqrt{2}\right)$$

$$=\left((a_1+a_2+a_3)+(b_1+b_2+b_3)\sqrt{2}\right)$$

(iv) $V_1+0=\left(a_1+b_1\sqrt{2}\right)+\left(0+0\sqrt{2}\right)=\left((a_1+0)+(b_1+0)\sqrt{2}\right)=\left(a_1+b_1\sqrt{2}\right)=V_1$

(v) Let $V_1=a_1+b_1\sqrt{2}$ and $-V_1=-a_1-b_1\sqrt{2}$.

$$V_1+(-V_1)=\left(a_1+b_1\sqrt{2}\right)+\left(-a_1-b_1\sqrt{2}\right)=\left((a_1-a_1)+(b_1-b_1)\sqrt{2}\right)=\left(0+0\sqrt{2}\right)=0\in V$$

(vi) $\alpha V_1=\alpha\left(a_1+b_1\sqrt{2}\right)=\left(\alpha a_1+\alpha b_1\sqrt{2}\right)\in V$

(vii) $\alpha(V_1+V_2)=\alpha\left[\left(a_1+b_1\sqrt{2}\right)+\left(a_2+b_2\sqrt{2}\right)\right]=\left[\alpha\left(a_1+b_1\sqrt{2}\right)+\alpha\left(a_2+b_2\sqrt{2}\right)\right]=\alpha V_1+\alpha V_2$

(viii) $(\alpha+\beta)V_1=(\alpha+\beta)\left(a_1+b_1\sqrt{2}\right)=\left((\alpha+\beta)a_1+(\alpha+\beta)b_1\sqrt{2}\right)\in V$

$\alpha V_1=\alpha\left(a_1+b_1\sqrt{2}\right)=\left(\alpha a_1+\alpha b_1\sqrt{2}\right)\in V$ and $\beta V_1=\beta\left(a_1+b_1\sqrt{2}\right)=\left(\beta a_1+\beta b_1\sqrt{2}\right)\in V$

$$\alpha V_1+\beta V_1=\left(\alpha a_1+\alpha b_1\sqrt{2}\right)+\left(\beta a_1+\beta b_1\sqrt{2}\right)=\left((\alpha+\beta)a_1+(\alpha+\beta)b_1\sqrt{2}\right)$$

(ix) $\alpha\beta V_1=\alpha\beta\left(a_1+b_1\sqrt{2}\right)=\left(\alpha\beta a_1+\alpha\beta b_1\sqrt{2}\right)$

$$\alpha(\beta V_1)=\alpha\left(\beta a_1+\beta b_1\sqrt{2}\right)=\left(\alpha\beta a_1+\alpha\beta b_1\sqrt{2}\right)$$

(x) Let 1 be a scalar. Then $1\cdot V_1=1\cdot\left(a_1+b_1\sqrt{2}\right)=\left(1\cdot a_1+1\cdot b_1\sqrt{2}\right)=\left(a_1+b_1\sqrt{2}\right)=V_1$

Since all the conditions are satisfied, V is a vector space.

Example: Show $V=\{(x,y)\in R^2 : y=2x\}$ that a vector space.

Let $u=(x,2x),\ v=(p,2p),\ w=(m,2m)\in V$ and $\alpha,\ \beta\in R$ are any scalars.

(i) $u+v=(x,2x)+(p,2p)=(x+p,\ 2x+2p)=(x+p,\ 2(x+p))\in V$

(ii) $u+v=(x,2x)+(p,2p)=(x+p,\ 2x+2p)=(p+x,\ 2p+2x)=(p,2p)+(x,2x)=v+u$

(iii) $u+v=(x,2x)+(p,2p)=(x+p,\ 2x+2p)$

$$v+w=(p,2p)+(m,2m)=(p+m,\ 2p+2m)$$

$$(u+v)+w=(x+p,\ 2x+2p)+(m,2m)$$
$$=(x+p+m,\ 2x+2p+2m)$$
$$u+(v+w)=(x,2x)+(p+m,\ 2p+2m)$$
$$=(x+p+m,\ 2x+2p+2m)$$

(iv) L et $u=(x,2x),\ 0=(0,0)\in V$. Here

$$u+0=(x,2x)+(0,\ 0)=(x+0,\ 2x+0)=(x,2x)=u$$

(v) Let $u=(x,\ 2x)\in V$ and let $-u=(-x,-2x)\in V$. Now

$$u+(-u)=(x,2x)+(-x,-2x)=(x-x,\ 2x-2x)=(0,0)=0.$$

(vi) $\alpha u=\alpha(x,2x)=(\alpha x,2\alpha x)\in V$

(vii) $\alpha(u+v)=\alpha\left[(x,2x)+(p,2p)\right]=\alpha(x+p,\ 2x+2p)=(\alpha x+\alpha p,\ 2\alpha x+2\alpha p)=(\alpha x+\alpha p,\ 2(\alpha x+\alpha$

$$\alpha u+\alpha v=\alpha(x,2x)+\alpha(p,2p)=(\alpha x,2\alpha x)+(\alpha p,2\alpha p)=(\alpha x+\alpha p,\ 2(\alpha x+\alpha p))$$

(viii) $(\alpha+\beta)u=(\alpha+\beta)(x,2x)=((\alpha+\beta)x,\ (\alpha+\beta)2x)=(\alpha x+\beta x,\ 2\alpha x+2\beta x)$

$$\alpha u+\beta u=\alpha(x,2x)+\beta(x,2x)=(\alpha x,\ 2\alpha x)+(\beta x,\ 2\beta x)=(\alpha x+\beta x,\ 2\alpha x+2\beta x)$$

(ix) $(\alpha\beta)u=(\alpha\beta)(x,2x)=(\alpha\beta x,\ 2\alpha\beta x)$

$$(\alpha)\beta u=(\alpha)\beta(x,2x)=(\alpha)(\beta x,\ 2\beta x)=(\alpha\beta x,\ 2\alpha\beta x)$$

(x) Let 1 be a scalar. Then $1u=1(x,2x)=(x,2x)=u$

Since all the conditions are satisfied, V is a vector space.

Example: Show that the set of all 2×2 matrices of the form $\begin{pmatrix} a & b \\ c & 0 \end{pmatrix}$ with the standard operations is a vector space.

Let $V =$ the set of all 2×2 matrices of the form $\begin{pmatrix} a & b \\ c & 0 \end{pmatrix}$ with the standard operations matrix addition and scalar multiplication.

Let $x=\begin{pmatrix} a & b \\ c & 0 \end{pmatrix}$, $y=\begin{pmatrix} d & e \\ f & 0 \end{pmatrix}$, $z=\begin{pmatrix} g & h \\ i & 0 \end{pmatrix}\in V$ and $\alpha,\ \beta$ are any scalars.

(i) $x+y=\begin{pmatrix} a & b \\ c & 0 \end{pmatrix}+\begin{pmatrix} d & e \\ f & 0 \end{pmatrix}=\begin{pmatrix} a+d & b+e \\ c+f & 0 \end{pmatrix}\in V$

(ii) $x+y=\begin{pmatrix} a & b \\ c & 0 \end{pmatrix}+\begin{pmatrix} d & e \\ f & 0 \end{pmatrix}=\begin{pmatrix} a+d & b+e \\ c+f & 0 \end{pmatrix}=\begin{pmatrix} d+a & e+b \\ f+c & 0 \end{pmatrix}=\begin{pmatrix} d & e \\ f & 0 \end{pmatrix}+\begin{pmatrix} a & b \\ c & 0 \end{pmatrix}=y+x$

(iii) $x+y=\begin{pmatrix} a & b \\ c & 0 \end{pmatrix}+\begin{pmatrix} d & e \\ f & 0 \end{pmatrix}=\begin{pmatrix} a+d & b+e \\ c+f & 0 \end{pmatrix}$

$$y+z=\begin{pmatrix} d & e \\ f & 0 \end{pmatrix}+\begin{pmatrix} g & h \\ i & 0 \end{pmatrix}=\begin{pmatrix} d+g & e+h \\ f+i & 0 \end{pmatrix}$$

$$(x+y)+z=\begin{pmatrix} a+d & b+e \\ c+f & 0 \end{pmatrix}+\begin{pmatrix} g & h \\ i & 0 \end{pmatrix}$$

$$=\begin{pmatrix} a+d+g & b+e+h \\ c+f+i & 0 \end{pmatrix}$$

$$x+(y+z)=\begin{pmatrix} a & b \\ c & 0 \end{pmatrix}+\begin{pmatrix} d+g & e+h \\ f+i & 0 \end{pmatrix}$$

$$=\begin{pmatrix} a+d+g & b+e+h \\ c+f+i & 0 \end{pmatrix}$$

(iv) Let $x=\begin{pmatrix} a & b \\ c & 0 \end{pmatrix}$, $0=\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$. Here $x+0=x$. But $0=\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\notin V$ **(condition fails).**

(v) Let $x=\begin{pmatrix} a & b \\ c & 0 \end{pmatrix}\in V$ then $-x=\begin{pmatrix} -a & -b \\ -c & 0 \end{pmatrix}\in V$. Now $x+(-x)=\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

(vi) $\alpha x = \alpha \begin{pmatrix} a & b \\ c & 0 \end{pmatrix} = \begin{pmatrix} \alpha a & \alpha b \\ \alpha c & 0 \end{pmatrix} \in V$

(vii) $\alpha(x+y) = \alpha\left(\begin{pmatrix} a & b \\ c & 0 \end{pmatrix} + \begin{pmatrix} d & e \\ f & 0 \end{pmatrix}\right) = \alpha \begin{pmatrix} a+d & b+e \\ c+f & 0 \end{pmatrix} = \begin{pmatrix} \alpha a + \alpha d & \alpha b + \alpha e \\ \alpha c + \alpha f & 0 \end{pmatrix} \in V$

$\alpha x + \alpha y = \begin{pmatrix} \alpha a & \alpha b \\ \alpha c & 0 \end{pmatrix} + \begin{pmatrix} \alpha d & \alpha e \\ \alpha f & 0 \end{pmatrix} = \begin{pmatrix} \alpha a + \alpha d & \alpha b + \alpha e \\ \alpha c + \alpha f & 0 \end{pmatrix} \in V$

(viii) $(\alpha+\beta)x = (\alpha+\beta)\begin{pmatrix} a & b \\ c & 0 \end{pmatrix} = \begin{pmatrix} (\alpha+\beta)a & (\alpha+\beta)b \\ (\alpha+\beta)c & 0 \end{pmatrix} = \begin{pmatrix} \alpha a + \beta a & \alpha b + \beta b \\ \alpha c + \beta c & 0 \end{pmatrix}$

$(\alpha x + \beta x) = \begin{pmatrix} \alpha a & \alpha b \\ \alpha c & 0 \end{pmatrix} + \begin{pmatrix} \beta a & \beta b \\ \beta c & 0 \end{pmatrix} = \begin{pmatrix} \alpha a + \beta a & \alpha b + \beta b \\ \alpha c + \beta c & 0 \end{pmatrix}$

(ix) $(\alpha\beta)x = (\alpha\beta)\begin{pmatrix} a & b \\ c & 0 \end{pmatrix} = \begin{pmatrix} (\alpha\beta)a & (\alpha\beta)b \\ (\alpha\beta)c & 0 \end{pmatrix} = \begin{pmatrix} \alpha\beta a & \alpha\beta b \\ \alpha\beta c & 0 \end{pmatrix}$

$(\alpha)\beta x = (\alpha)\begin{pmatrix} \beta a & \beta b \\ \beta c & 0 \end{pmatrix} = \begin{pmatrix} \alpha\beta a & \alpha\beta b \\ \alpha\beta c & 0 \end{pmatrix}$

(x) Clearly $1x = 1 \cdot \begin{pmatrix} a & b \\ c & 0 \end{pmatrix} = \begin{pmatrix} a & b \\ c & 0 \end{pmatrix}$

Since all conditions are satisfied, V is a vector space.

Example: Determine whether the set of all pairs of real numbers (x,y) with the operations $(x,y)+(p,q)=(x+p+1, y+q+1)$ and $k(x,y)=(kx,ky)$ is a vector space or not. If not, list all the axioms that fail to hold.

Let $V =$ the set of all pairs of real numbers (x,y) with the standard operations as given below.

$(x,y)+(p,q)=(x+p+1, y+q+1)$ and $k(x,y)=(kx,ky)$

Let $u=(x,y)$, $v=(p,q)$, $w=(m,n) \in V$ and α, β are any scalars.

(i) $u+v=(x,y)+(p,q)=(x+p+1,\ y+q+1)\in V$

(ii) $u+v=(x,y)+(p,q)=(x+p+1,\ y+q+1)=(p+x+1,\ q+y+1)=(p,q)+(x,y)=v+u$

(iii) $u+v=(x,y)+(p,q)=(x+p+1,\ y+q+1)$

$$v+w=(p,q)+(m,n)=(p+m+1,\ q+n+1)$$

$$(u+v)+w=(x+p+1,\ y+q+1)+(m,n)$$

$$=(x+p+m+2,\ y+q+n+2)$$

$$u+(v+w)=(x,y)+(p+m+1,\ q+n+1)$$

$$=(x+p+m+2,\ y+q+n+2)$$

(iv) Let $u=(x,y),\ 0=(-1,-1)\in V$. Here $u+0=(x,y)+(-1,-1)=(x-1+1,\ y-1+1)=(x,y)=u$

(v) Let $u=(x,y)\in V$ and let $-u=(-x-2,-y-2)\in V$. Now

$$u+(-u)=(x,y)+(-x-2,-y-2)=(x-x-2+1,\ y-y-2+1)=(-1,-1)=0$$

(vi) $\alpha u=\alpha(x,y)=(\alpha x,\alpha y)\in V$

(vii) $\alpha(u+v)=\alpha\left[(x,y)+(p,q)\right]=\alpha(x+p+1,\ y+q+1)=(\alpha x+\alpha p+\alpha,\ \alpha y+\alpha q+\alpha)$

$$\alpha u+\alpha v=\alpha(x,y)+\alpha(p,q)=(\alpha x,\alpha y)+(\alpha p,\alpha q)=(\alpha x+\alpha p+1,\ \alpha y+\alpha q+1)$$

(viii) $(\alpha+\beta)u=(\alpha+\beta)(x,y)=((\alpha+\beta)x,\ (\alpha+\beta)y)=(\alpha x+\beta x,\ \alpha y+\beta y)$

$$\alpha u+\beta u=\alpha(x,y)+\beta(x,y)=(\alpha x,\ \alpha y)+(\beta x,\ \beta y)=(\alpha x+\beta x+1,\ \alpha y+\beta y+1)$$

(ix) $(\alpha\beta)u=(\alpha\beta)(x,y)=(\alpha\beta x,\ \alpha\beta y)$

$$(\alpha)\beta u=(\alpha)\beta(x,y)=(\alpha)(\beta x,\ \beta y)=(\alpha\beta x,\ \alpha\beta y)$$

(x) Let 1 be a scalar. Then $1u=1(x,y)=(x,y)=u$

Here $\alpha(u+v)\neq\alpha u+\alpha v$ and $(\alpha+\beta)u\neq\alpha u+\beta u$

Since condition (vii) and (viii) fails, V is not a vector space.

Verify that the set V of all ordered triplets of real numbers of the form (x_1, y_1, z_1) and defined the operations $+$ and $\cdot$ by (i) $(x_1, y_1, z_1)+(x_2, y_2, z_2)=(x_1+x_2, y_1+y_2, z_1+z_2)$, (ii) $c\cdot(x_1, y_1, z_1)=(cx_1, y_1, z_1)$ is a vector space or not.

Verify that the set V of all ordered triplets of real numbers of the form $(x_1, y_1, 0)$ and defined the operations $+$ and $\cdot$ by (i) $(x_1, y_1, 0)+(x_2, y_2, 0)=(x_1+x_2, y_1+y_2, 0)$, (ii) $c\cdot(x_1, y_1, 0)=(cx_1, cy_1, 0)$ is a vector space or not.

Example: Let $V=\{(a_1, a_2): a_1, a_2 \in R\}$. Define addition of elements of a_1, a_2 coordinate wise, and for any scalar c, define $c(a_1, a_2)=(a_1, 0)$. Is V a vector space with these operations? Justify your answer

No. Since $0(a_1, a_2)=(a_1, 0)$ is the zero vector but this will make the zero vector not be unique, it cannot be a vector space.

Theorem: In any vector space V, the following statements are true: (i) $0x=0$ for each $x\in V$.
(ii) $(-a)x=-(ax)=a(-x)$ for each $a\in R$ and each $x\in V$. (iii) $a0=0$ for each $a\in R$ (iv) $ax=0$ then $a=0$ or $x=0$.
(i) Consider $0x+0x=(0+0)x=0x=0x+0=0+0x$. Hence by cancellation law $0x=0$ for each $x\in V$.

(ii) If $ax\in V$ then $-(ax)\in V$ such that $(ax)+[-(ax)]=0$.

Consider $ax+(-a)x=[a+(-a)]x=0x=0$.

Comparing, we get $(-a)x=-(ax)$

Now $a(-x)=a[(-1)x]=[a(-1)x]=(-a)x=-(ax)$

(iii) Consider $a0+a0=a(0+0)=a0+0=0+a0$. Hence by cancellation law $a0=0$ for each $a\in R$.

(iv) Suppose $ax=0$ and $a\neq 0$. Since a is a scalar, a^{-1} exists.

Therefore $a^{-1}(ax)=a^{-1}(0)$

$$(a^{-1}a)x=a^{-1}(0)$$

$(1) x = 0$

$x = 0$

Theorem: In any vector space V, the following statements are true: (i) The zero vector is unique
(ii) The inverse identity is unique.

(i) Let $x \in V$ and $0, 0' \in V$ such that $x + 0 = x$ and $x + 0' = x, \forall x$.

Therefore $x + 0 = x + 0'$

$0 = 0'$ {by cancellation law}

(ii) Let $x \in V$ and $y, y' \in V$ such that $x + y = 0$ and $x + y' = 0, \forall x$.

Therefore $x + y = x + y'$

$y = y'$ {by cancellation law}

Try this: In any vector space, Is $ax = bx$ implies that $a = b$?. {Hint: No. Consider if the vector is zero}

Subspaces

Definition: A nonempty subset W of a vector space V is called a subspace of V if W is a vector space under the operations of addition and scalar multiplication defined in V.

Note: Let V be a vector space. A subset W of V is called a subspace of V if it satisfies the following properties:

- The zero vector of V is also in W.
- If x and y are in W then $x + y$ is in W. (W is closed under addition)
- If x is in W and α is a scalar then αx is in W. (W is closed under scalar multiplication)

Note:

- The trivial subspaces of a vector space V are V itself and zero vector.
- The empty set is not a subspace of every vector space. Because any subspace contains 0.

Example: Let $W = \{(x, y) \in R^2 : y = 2x\}$. Now W a subspace of $V = \mathbb{R}^2$.

Example: Let V be the vector space of $n \times n$ matrices. Let W be the se of all $n \times n$ matrices whose $tr(A) = 0$. Is W a subspace of V?.

(i) The zero vector of V is a $n \times n$ zero matrix. The trace of this zero matrix is zero. Therefore zero vector is also in W.

(ii) Let A, B are in W. Then $tr(A) = a_{11} + a_{22} + \cdots + a_{nn} = 0$ and $tr(B) = b_{11} + b_{22} + \cdots + b_{nn} = 0$.

Now $tr(A+B) = (a_{11} + b_{11}) + (a_{22} + b_{22}) + \cdots + (a_{nn} + b_{nn})$

$$= (a_{11} + a_{22} + \cdots + a_{nn}) + (b_{11} + b_{22} + \cdots + b_{nn})$$

$$= 0 + 0$$

$$= 0$$

Therefore $A + B$ is in W. i.e. W is closed under addition

(iii) Let α be a scalar and consider the matrix αA.

Now $tr(\alpha A) = \alpha a_{11} + \alpha a_{22} + \cdots + \alpha a_{nn}$

$$= \alpha (a_{11} + a_{22} + \cdots + a_{nn})$$

$$= \alpha \times 0$$

$$= 0$$

Therefore αA is in W. i.e. W is closed under scalar multiplication

Therefore W is a subspace of V.

Example: Show that the set $W = \{(x_1, x_2, 0); x_1, x_2 \in \mathbb{R}\}$ is a subspace of $V = \mathbb{R}^3$ with the standard operations.

(i) The zero vector of $\mathbb{R}^3$ is $(0, 0, 0)$ which is in W.

(ii) Let $u = (x_1, x_2, 0)$, $v = (y_1, y_2, 0) \in W$.

Now $u + v = (x_1, x_2, 0) + (y_1, y_2, 0) = (x_1 + y_1,\ x_2 + y_2,\ 0) \in W$ i.e. W is closed under addition.

(iii) Let α be a scalar and let $u = (x_1, x_2, 0) \in W$.

Now $\alpha u = \alpha (x_1, x_2, 0) = (\alpha x_1,\ \alpha x_2,\ 0) \in W$ i.e. W is closed under scalar multiplication

Therefore W is a subspace of V.

Example: Show that the set $W = \{(x_1, x_2); x_1 \geq 0, \ x_2 \geq 0\}$, with the standard operations, is not a subspace of $V = \mathbb{R}^2$.

Let $\alpha = -1$ be a scalar and let $u = (1, 1) \in W$.

Now $\alpha u = -1(1, 1) = (-1, -1) \notin W$ i.e. W is not closed under scalar multiplication

Therefore W is not a subspace of V.

Theorem: Intersection of two subspaces of a vector space V is also a subspace of V.

Let A and B be two subspaces of a vector space V over a field F.

Since A and B be two subspaces $0 \in A$ and $0 \in B$ hence $0 \in A \cap B$.

Let $x, y \in A \cap B$ and $\alpha, \ \beta \in F$.

Therefore $x, y \in A$ and $x, y \in B$.

Therefore $\alpha x + \beta y \in A$ and $\alpha x + \beta y \in B$. Hence $\alpha x + \beta y \in A \cap B$.

Hence $A \cap B$ is a subspace of V.

Result:

- Let $A = \{(a, 0, 0) : a \in R\}$ and $A = \{(0, b, 0) : b \in R\}$ be subspaces of R^3. But $(1, 0, 0) + (0, 1, 0) = (1, 1, 0) \notin A \cup B$. Therefore $A \cup B$ is not a subspace of R^3.
- But $A \cup B$ is a subspace if and only if one is contained in another.

Example: Show that the set $w = \{(a_1, a_2, a_3) \in R^3 : 2a_1 - 7a_2 + a_3 = 1\}$ is not a subspace.

Consider the zero vector $(0, 0, 0) \in R^3$

But given that $2a_1 - 7a_2 + a_3 = 1$

$$2(0) - 7(0) + (0) \neq 1$$

Therefore $(0, 0, 0) \notin w$ and hence w is not a sub space R^3.

Example: Show that the set $w = \{(a_1, a_2, a_3) \in R^3 : a_1 = 3a_2, \ \ a_3 = -a_2\}$ is a subspace of R^3.

Consider the zero vector $(0,\ 0,\ 0) \in R^3$

Here $a_1 = 3a_2,\ \ a_3 = -a_2$

$0 = 3(0), \quad 0 = -0$

Therefore $(0,\ 0,\ 0) \in w$

Let $x = (a_1,\ a_2,\ a_3),\ \ y = (b_1,\ b_2,\ b_3) \in w$. By definition of w, we have

$a_1 = 3a_2,\ \ a_3 = -a_2$ and $b_1 = 3b_2,\ \ b_3 = -b_2$

Now $x + y = (a_1,\ a_2,\ a_3) + (b_1,\ b_2,\ b_3) = (a_1 + b_1,\ a_2 + b_2,\ a_3 + b_3)$

Here $a_1 + b_1 = 3a_2 + 3b_2$ and $a_3 + b_3 = -a_2 - b_2$

Therefore $x + y \in w$

Let c be a scalar and $x = (a_1,\ a_2,\ a_3) \in w$. Therefore $a_1 = 3a_2,\ \ a_3 = -a_2$.

Consider $cx = c(a_1,\ a_2,\ a_3) = (ca_1,\ ca_2,\ ca_3)$

Here $ca_1 = 3ca_2,\ \ ca_3 = -ca_2$

Therefore $cx \in w$ and hence w is a sub space R^3.

Spanning Sets

Definition: Let $S = \{v_1,\ v_2,\ \cdots\cdots,\ v_m\}$ be a subset of a vector space V. The set S is called a spanning set of V if every vector in V can be written as a linear combination of vectors in S. In such cases it is said that S spans V.

Note:

- $\text{Span}(S) = \{\alpha_1 v_1 + \alpha_2 v_2 + \cdots\cdots + \alpha_m v_m\}$ where $\alpha_1,\ \alpha_2,\ \cdots\cdots,\ \alpha_m$ are real numbers.
- If $\text{Span}(S) = V$, it is said that V is spanned by S or S spans V.
- The $\text{Span}(S)$ is a subspace of V.

Example: The set $S = \{(1,0,0),\ (0,1,0),\ (0,0,1)\}$ spans R^3 because any vector $v = (v_1,\ v_2,\ v_3)$ in R^3 can be written as $v = v_1(1,0,0) + v_2(0,1,0) + v_3(0,0,1) = (v_1,\ v_2,\ v_3)$. Here S is known as standard spanning set of R^3.

Example: The set $S = \{1, x, x^2\}$ spans the set P_2 of all polynomials of degree 2 because any polynomial function $f(x) = a + bx + cx^2$ in P_2 can be written as $f(x) = a(1) + b(x) + c(x^2) = a + bx + cx^2$. Here S is known as standard spanning set of P_2.

Example: Show that the set $S = \{(1,2,3), (0,1,2), (-2,0,1)\}$ spans R^3.

Let $v = (v_1, v_2, v_3)$ be any vector in R^3.

We have to find scalars $\alpha_1, \alpha_2, \alpha_3$ such that

$$(v_1, v_2, v_3) = \alpha_1(1,2,3) + \alpha_2\ (0,1,2) + \alpha_3\ (-2,0,1)$$

$$(v_1, v_2, v_3) = (\alpha_1 - 2\alpha_3,\ 2\alpha_1 + \alpha_2,\ 3\alpha_1 + 2\alpha_2 + \alpha_3)$$

Equating the components of the vectors, we have

$$\begin{aligned} \alpha_1 + 0\alpha_2 - 2\alpha_3 &= v_1 \\ 2\alpha_1 + \alpha_2 + 0\alpha_3 &= v_2 \\ 3\alpha_1 + 2\alpha_2 + \alpha_3 &= v_3 \end{aligned}$$

Consider the determinant value of the coefficient matrix of the system

$$\begin{vmatrix} 1 & 0 & -2 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{vmatrix} = 1(1-0) - 2(4-3) = -1 \neq 0$$

Hence the system has a unique solution. So, any vector in R^3 can be written as a linear combination of the vectors in S. Therefore the set S spans R^3.

Or (apply $X = A^{-1}B$ to find $\alpha_1, \alpha_2, \alpha_3$ in terms of v_1, v_2, v_3)

Example: Determine whether the set $S = \{(1,2,3), (0,1,2), (-1,0,1)\}$ spans R^3.

Let $v = (v_1, v_2, v_3)$ be any vector in R^3.

We have to find scalars $\alpha_1, \alpha_2, \alpha_3$ such that

$$(v_1, v_2, v_3) = \alpha_1(1,2,3) + \alpha_2\ (0,1,2) + \alpha_3\ (-1,0,1)$$

$$(v_1, v_2, v_3) = (\alpha_1 - \alpha_3,\ 2\alpha_1 + \alpha_2,\ 3\alpha_1 + 2\alpha_2 + \alpha_3)$$

Equating the components of the vectors, we have

$$\begin{aligned} \alpha_1 + 0\alpha_2 - \alpha_3 &= v_1 \\ 2\alpha_1 + \alpha_2 + 0\alpha_3 &= v_2 \\ 3\alpha_1 + 2\alpha_2 + \alpha_3 &= v_3 \end{aligned}$$

Consider the determinant value of the coefficient matrix of the system

$$\begin{vmatrix} 1 & 0 & -1 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{vmatrix} = 1(1-0) - 1(4-3) = 1 - 1 = 0$$

Hence the system has no solution. So, all vectors in R^3 cannot be written as a linear combination of the vectors in S. Therefore the set S does not span R^3.

Ex: The vector $(1,-2,2)$ in R^3 cannot be expressed as a linear combination of vectors of S.

Linear Dependence and Independence of vectors

Definition: A set of vectors $S = \{v_1, v_2, \cdots\cdots, v_m\}$ in a vector space V is called linearly independent if the vector equation $\alpha_1 v_1 + \alpha_2 v_2 + \cdots\cdots + \alpha_m v_m = 0$ has the trivial solution $\alpha_1 = \alpha_2 = \cdots\cdots = \alpha_m = 0$. If the vector equation has a non trivial solution, then the vectors are linearly dependent.

Property:

- If the vectors $S = \{v_1, v_2, \cdots\cdots, v_m\}$ are linearly dependent, then **some** vector can be expressed as linear combination of other vectors i.e. $v_j = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_{j-1} v_{j-1} + \alpha_{j+1} v_{j+1} + \cdots + \alpha_m v_m$.
- It is not true that, if S is a linearly dependent set, then each vector in S is a linear combination of other vectors in S. **Ex:** $S = \{(1,0), (2,0), (0,1)\}$. Here $(0,1)$ cannot be expressed in terms of the other two.

Note:

- A set consisting of a single non-zero vector is linearly independent.
- Any set containing the zero vector, is linearly dependent.
- The empty set is linearly independent
- A set consisting of two non-zero vectors is linearly independent if and only if neither of the vectors is a multiple of the other.
- If a set of vectors is linearly independent, any non empty sub set of these vectors is also linearly independent.
- Consider the set of vectors $\{v_1, v_2, \cdots\cdots, v_m\}$ in $\mathbb{R}^n$. If $m > n$ then the vectors are linearly dependent.
- If $m \le n$ then the vectors are linearly independent.

Example: Show that the vectors $v_1 = \begin{pmatrix}1\\2\\3\end{pmatrix}$, $v_2 = \begin{pmatrix}4\\5\\6\end{pmatrix}$ and $v_3 = \begin{pmatrix}2\\1\\0\end{pmatrix}$ are linearly dependent. Also verify the property.

Consider the vector equation $\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 = 0$. That is

$$\alpha_1 + 4\alpha_2 + 2\alpha_3 = 0$$

$$2\alpha_1 + 5\alpha_2 + \alpha_3 = 0$$

$$3\alpha_1 + 6\alpha_2 + 0\,\alpha_3 = 0$$

To solve the system of equation by Gauss Elimination method, consider the augmented matrix

$$(A,B) = \begin{pmatrix} 1 & 4 & 2 & 0 \\ 2 & 5 & 1 & 0 \\ 3 & 6 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 4 & 2 & 0 \\ 0 & -3 & -3 & 0 \\ 0 & -6 & -6 & 0 \end{pmatrix} \begin{matrix} \\ R_2 \to R_2 - 2R_1 \\ R_3 \to R_3 - 3R_1 \end{matrix}$$

$$= \begin{pmatrix} 1 & 4 & 2 & 0 \\ 0 & -3 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{matrix} \\ \\ R_3 \to R_3 - 2R_2 \end{matrix}$$

This $\left[r(A) = 2 < no.\ of\ \ vectors \right]$ implies that the system has an infinite number of solutions. So, the system must have nontrivial solutions. Let $\alpha_3 = 1$.
From the second row, $-3\alpha_2 - 3\alpha_3 = 0$

From the first row, $\alpha_1 + 4\alpha_2 + 2\alpha_3 = 0$

$$-\alpha_2 = \alpha_3$$
$$\alpha_2 = -1$$
$$\alpha_1 - 4 + 2 = 0$$
$$\alpha_1 = 2$$

Therefore the vectors are linearly dependent and hence the vector equation becomes $2v_1 - v_2 + v_3 = 0$

By a property, we have $v_2 = 2v_1 + v_3$

$$\begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix} + \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$$

Example: Show that the vectors $v_1 = \begin{pmatrix} 0 \\ 1 \\ 5 \end{pmatrix}$, $v_2 = \begin{pmatrix} 1 \\ 2 \\ 8 \end{pmatrix}$ and $v_3 = \begin{pmatrix} 4 \\ -1 \\ 0 \end{pmatrix}$ are linearly independent.

Consider the vector equation $\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 = 0$. That is

$$0\alpha_1 + 1\alpha_2 + 4\alpha_3 = 0$$

$$\alpha_1 + 2\alpha_2 - \alpha_3 = 0$$

$$5\alpha_1 + 8\alpha_2 + 0\alpha_3 = 0$$

To solve the system of equation by Gauss Elimination method, consider the augmented matrix

$$(A, B) = \begin{pmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & 4 & 0 \\ 5 & 8 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & -2 & 5 & 0 \end{pmatrix} R_3 \to R_3 - 5R_1$$

$$= \begin{pmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 13 & 0 \end{pmatrix} R_3 \to R_3 + 2R_2$$

This $\left[r(A) = 3 = \textit{no. of vectors} \right]$ implies that the system has unique trivial solution $\alpha_1 = \alpha_2 = \alpha_3 = 0$. So, the vectors are linearly independent.

Example: Determine whether the set $S = \{2 - x,\ 2x - x^2,\ 6 - 5x + x^2\}$ in $\mathcal{P}_2$ is linearly independent.

Consider the vector equation $\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 = 0$. That is

$$\alpha_1 (2 - x) + \alpha_2 (2x - x^2) + \alpha_3 (6 - 5x + x^2) = (0 + 0x + 0x^2)$$

$$(2\alpha_1 + 6\alpha_3) + (-\alpha_1 + 2\alpha_2 - 5\alpha_3) x + (-\alpha_2 + \alpha_3) x^2 = (0 + 0x + 0x^2)$$

Equating the corresponding coefficients, we have the system of equations

$$2\alpha_1 + 6\alpha_3 = 0$$
$$-\alpha_1 + 2\alpha_2 - 5\alpha_3 = 0$$
$$-\alpha_2 + \alpha_3 = 0$$

To solve the system of equation by Gauss Elimination method, consider the augmented matrix

$$(A,B) = \begin{pmatrix} 2 & 0 & 6 & 0 \\ -1 & 2 & -5 & 0 \\ 0 & -1 & 1 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 2 & 0 & 6 & 0 \\ 0 & 4 & -4 & 0 \\ 0 & -1 & 1 & 0 \end{pmatrix} R_2 \to 2R_2 + R_1$$

$$= \begin{pmatrix} 2 & 0 & 6 & 0 \\ 0 & 4 & -4 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} R_3 \to 4R_3 + R_2$$

This $\left[r(A) = 2 < no.\ of\ \ vectors \right]$ implies that the system has an infinite number of solutions. So, the system must have nontrivial solutions. Let $\alpha_3 = 1$.
From the second row, $4\alpha_2 - 4\alpha_3 = 0$

From the first row, $2\alpha_1 + 0\alpha_2 + 6\alpha_3 = 0$

$$\alpha_2 = \alpha_3$$
$$\alpha_2 = 1$$

$$2\alpha_1 + 6 = 0$$
$$\alpha_1 = -3$$

Therefore the vectors are linearly dependent and hence the vector equation becomes $-3v_1 + v_2 + v_3 = 0$

Example: Determine whether the matrices $\begin{pmatrix} 1 & -1 \\ 4 & 5 \end{pmatrix}, \begin{pmatrix} 4 & 3 \\ -2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & -8 \\ 22 & 23 \end{pmatrix}$ form a linearly independent set.

Consider the vector equation $\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 = 0$. That is

$$\alpha_1\begin{pmatrix}1 & -1\\ 4 & 5\end{pmatrix}+\alpha_2\begin{pmatrix}4 & 3\\ -2 & 3\end{pmatrix}+\alpha_3\begin{pmatrix}1 & -8\\ 22 & 23\end{pmatrix}=\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}$$

Equating the corresponding elements, we have the system of equations

$$\begin{aligned}\alpha_1+4\alpha_2+\alpha_3&=0\\ -\alpha_1+3\alpha_2-8\alpha_3&=0\\ 4\alpha_1-2\alpha_2+22\alpha_3&=0\\ 5\alpha_1+3\alpha_2+23\alpha_3&=0\end{aligned}$$

To solve the system of equation by Gauss Elimination method, consider the augmented matrix

$$(A,B)=\begin{pmatrix}1 & 4 & 1 & 0\\ -1 & 3 & -8 & 0\\ 4 & -2 & 22 & 0\\ 5 & 3 & 23 & 0\end{pmatrix}$$

$$=\begin{pmatrix}1 & 4 & 1 & 0\\ 0 & 7 & -7 & 0\\ 0 & -18 & 18 & 0\\ 0 & -17 & 18 & 0\end{pmatrix}\begin{matrix}\\ R_2\to R_2+R_1\\ R_3\to R_3-4R_1\\ R_4\to R_4-5R_1\end{matrix}$$

$$=\begin{pmatrix}1 & 4 & 1 & 0\\ 0 & 7 & -7 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 7 & 0\end{pmatrix}\begin{matrix}\\ \\ R_3\to 7R_3+18R_2\\ R_4\to 7R_4+17R_2\end{matrix}$$

$$=\begin{pmatrix}1 & 4 & 1 & 0\\ 0 & 7 & -7 & 0\\ 0 & 0 & 7 & 0\\ 0 & 0 & 0 & 0\end{pmatrix}$$

This $\left[r(A)=3=\textit{no. of vectors}\right]$ implies that the system has unique trivial solution $\alpha_1=\alpha_2=\alpha_3=0$. So, the vectors are linearly independent.

Example: Show that the matrices $\begin{pmatrix}1 & -3 & 2\\ -4 & 0 & 5\end{pmatrix},\begin{pmatrix}-3 & 7 & 4\\ 6 & -2 & -7\end{pmatrix},\begin{pmatrix}-2 & 3 & 11\\ -1 & -3 & 2\end{pmatrix}$ form a linearly dependent set in $M_{2\times3}$.

Consider the vector equation $\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 = 0$. That is

$$\alpha_1 \begin{pmatrix} 1 & -3 & 2 \\ -4 & 0 & 5 \end{pmatrix} + \alpha_2 \begin{pmatrix} -3 & 7 & 4 \\ 6 & -2 & -7 \end{pmatrix} + \alpha_3 \begin{pmatrix} -2 & 3 & 11 \\ -1 & -3 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Equating the corresponding elements, we have the system of equations

$$\begin{aligned} \alpha_1 - 3\alpha_2 - 2\alpha_3 &= 0 \\ -3\alpha_1 + 7\alpha_2 + 3\alpha_3 &= 0 \\ 2\alpha_1 + 4\alpha_2 + 11\alpha_3 &= 0 \\ -4\alpha_1 + 6\alpha_2 - \alpha_3 &= 0 \\ 0\alpha_1 - 2\alpha_2 - 3\alpha_3 &= 0 \\ 5\alpha_1 - 7\alpha_2 + 2\alpha_3 &= 0 \end{aligned}$$

To solve the system of equation by Gauss Elimination method, consider the augmented matrix

$$(A,B) = \begin{pmatrix} 1 & -3 & -2 & 0 \\ -3 & 7 & 3 & 0 \\ 2 & 4 & 11 & 0 \\ -4 & 6 & -1 & 0 \\ 0 & -2 & -3 & 0 \\ 5 & -7 & 2 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & -3 & -2 & 0 \\ 0 & -2 & -3 & 0 \\ 0 & 10 & 15 & 0 \\ 0 & -6 & -9 & 0 \\ 0 & -2 & -3 & 0 \\ 0 & 8 & 12 & 0 \end{pmatrix} \begin{matrix} \\ R_2 \to R_2 + 3R_1 \\ R_3 \to R_3 - 2R_1 \\ R_4 \to R_4 + 4R_1 \\ \\ R_5 \to R_5 - 5R_1 \end{matrix}$$

$$= \begin{pmatrix} 1 & -3 & -2 & 0 \\ 0 & -2 & -3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{matrix} \\ \\ R_3 \to R_3 + 5R_2 \\ R_4 \to R_3 - 3R_2 \\ R_5 \to R_5 - R_2 \\ R_6 \to R_6 + 4R_2 \end{matrix}$$

This $\left[r(A)=2< no.\ of\ \ vectors\right]$ implies that the system has an infinite number of solutions. So, the system must have nontrivial solutions. Let $\alpha_3=1$.
From the second row, $-2\alpha_2-3\alpha_3=0$
From the first row, $\alpha_1-3\alpha_2-2\alpha_3=0$

$$-2\alpha_2=3\alpha_3$$
$$-2\alpha_2=3$$
$$\alpha_2=-\frac{3}{2}$$

$$\alpha_1=3\alpha_2+2\alpha_3$$
$$\alpha_1=3\left(-\frac{3}{2}\right)+2$$
$$\alpha_1=-\frac{5}{2}$$

$\therefore$ the vectors are linearly dependent and hence the vector equation becomes

$$-\frac{5}{2}v_1-\frac{3}{2}v_2+v_3=0$$

i.e. $5v_1+3v_2-2\ v_3=0$

Example: Check whether the set $\left\{\begin{pmatrix}1 & -3 & 2\\ -4 & 0 & 5\end{pmatrix}, \begin{pmatrix}-3 & 7 & 4\\ 6 & -2 & -7\end{pmatrix}, \begin{pmatrix}-2 & 3 & 11\\ -1 & -3 & 2\end{pmatrix}\right\}$ is linearly dependent or not.

Let $\alpha,\ \beta,\ \gamma$ be the scalars such that

$$\alpha\begin{pmatrix}1 & -3 & 2\\ -4 & 0 & 5\end{pmatrix}+\beta\begin{pmatrix}-3 & 7 & 4\\ 6 & -2 & -7\end{pmatrix}+\gamma\begin{pmatrix}-2 & 3 & 11\\ -1 & -3 & 2\end{pmatrix}=\begin{pmatrix}0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}$$

Equating the corresponding elements, we have

$$\begin{aligned}\alpha-3\beta-2\gamma&=0 \\ -3\alpha+7\beta+3\gamma&=0 \\ 2\alpha+4\beta+11\gamma&=0\end{aligned} \qquad \text{and} \qquad \begin{aligned}-4\alpha+6\beta-\gamma&=0 \\ 0\alpha-2\beta-3\gamma&=0 \\ 5\alpha-7\beta+2\gamma&=0\end{aligned}$$

Solving, we have $\alpha=5,\ \ \beta=3,\ \ \gamma=-2.$ Hence the given set is linearly dependent.

Example: Subsets of linearly dependent sets need not be linearly dependent.

Here $S=\{(1,0),\ (2,0),\ (0,1)\}$ linearly dependent but $\{(1,0),\ (0,1)\}$ is linearly independent.

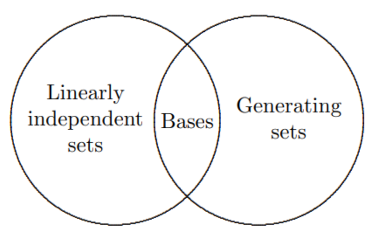
Basis and Dimension

Definition: A basis B for a vector space V is a linearly independent subset of V that generates V.

Note:

- B Spans V.
- A basis may have finite or infinite number of vectors. The set of all polynomials has no finite basis.
- A vector space V has a basis consisting of a finite number of vectors, then V is said to be finite dimensional.
- All bases of V have the same number of vectors.
- The empty set is the basis for zero vector space.
- If B is the basis for a vector space V, then every vector of V can be uniquely represented as a linear combination of vectors of the basis
- If a basis for a vector space V contains k vectors, then any set containing more than k vectors of V is linearly dependent
- Any set containing more vectors than the basis possess, will not be a basis as they are linearly dependent.
- A vector space may have more than one basis. $\{(1,0),\ (0,1)\}$ and $\{(1,0),\ (1,1)\}$ are the bases of R^2.

The relationship between the generating set, linearly independent set and bases is depicted in the Venn diagram.



Example: Standard basis for the vector space of 2×2 matrices is $\left\{\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}, \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}\right\}$.

Example: ϕ is the basis for the vector space $V=\{0\}$, as $Span(\phi)=\{0\}$.

Example: Check whether the set $B=\{(1,0,0),\ (0,1,0),\ (0,0,1)\}$ is the basis for R^3.

Any vector $v=(v_1,\ v_2,\ v_3)$ in R^3 can be written as
$(v_1,\ v_2,\ v_3)=v_1(1,0,0)+v_2(0,1,0)+v_3(0,0,1)$.
Therefore B spans R^3.

Clearly the vector equation $\alpha_1 v_1+\alpha_2 v_2+\alpha_3 v_3=0$ has a solution $\alpha_1=\alpha_2=\alpha_3=0$.
Therefore the vectors of B is linearly independent.

Therefore B is the basis for R^3.

Example: The set $B=\{(1,2,3),\ (0,1,2),\ (-2,0,1)\}$ is the basis of R^3.

It is proved that B spans R^3. Refer above.
Consider the vector equation
$\alpha_1 v_1+\alpha_2 v_2+\alpha_3 v_3=0$

$$\alpha_1-2\alpha_3=0$$

$$2\alpha_1+\alpha_2=0$$

$$3\alpha_1+2\alpha_2+\alpha_3=0$$

Solving, we have $\alpha_1=\alpha_2=\alpha_3=0$.
Therefore the vectors of B is linearly independent.

Therefore B is the basis for R^3.

Example: Show that the set $B=\{(1,0,0),\ (0,1,0),\ (1,1,1)\}$ is the basis for $V_3(R)$.

Consider the vector equation $\alpha_1 v_1+\alpha_2 v_2+\alpha_3 v_3=0$. That is

$$\alpha_1(1,\ 0,\ 0)+\alpha_2(0,\ 1,\ 0)+\alpha_3(1,\ 1,\ 1)=(0,\ 0,\ 0)$$

The system of equations are

$$\alpha_1+0\alpha_2+\alpha_3=0$$

$$0\alpha_1+\alpha_2+\alpha_3=0$$

$$0\alpha_1+0\alpha_2+\alpha_3=0$$

To solve the system of equation by Gauss Elimination method, consider the augmented matrix

$$(A,B)=\begin{pmatrix}1 & 0 & 1 & 0\\ 0 & 1 & 1 & 0\\ 0 & 0 & 1 & 0\end{pmatrix}$$

This $\left[r(A)=3=\textit{no. of vectors}\right]$ implies that the system has unique trivial solution $\alpha_1=\alpha_2=\alpha_3=0$. So, the vectors are linearly independent.

To prove the vectors forms the basis, we have to find scalars $l,\ m,\ n$ such that

$l(1, 0, 0)+m(0, 1, 0)+n(1, 1, 1)=(a, b, c)$

Equating the coefficients, we have $(l+n)=a,\ (m+n)=b,\ (n)=c$.

$$\begin{array}{lll} & l+n=a & m+n=b & \\ \text{Here} & l+c=a & m+c=b & \text{and} \quad n=c \\ & l=a-c & \therefore\ m=b-c & \end{array}$$

Therefore any vector in $V_3(R)$ can be uniquely expressed as a linear combination of vectors of B. Therefore B spans $V_3(R)$.

Try this: Show that the set $B=\{(1,0,0), (0,1,0), (1,1,1), (1,1,0)\}$ spans the vector space $V_3(R)$ but is not a basis.

Let $B'=\{(1,0,0), (0,1,0), (1,1,1)\}$. Since $B'\subseteq B$, then $L(B')=V_3(R)$. Thus B' spans $V_3(R)$. But B' is linearly dependent, because $(1, 1, 0)=(1, 0, 0)+(0, 1, 0)$. Hence B' is not a basis.

Try this: Show that the set $B=\{(1,0,0), (1,1,0)\}$ is linearly independent but not spans the vector space $V_3(R)$.

Try this: Show that the vector space P_2, the set of all polynomial functions of degree 2, has the basis $B=\{1, x, x^2\}$.

Example: Consider the vector space $V = R^4$. Let $B=\left\{\begin{pmatrix}1\\3\\0\\-2\end{pmatrix}, \begin{pmatrix}2\\-1\\-2\\1\end{pmatrix}, \begin{pmatrix}-1\\4\\2\\-3\end{pmatrix}\right\}$. Is B a basis for R^4.

Form the matrix

$$A=\begin{pmatrix}1 & 2 & -1\\ 3 & -1 & 4\\ 0 & -2 & 2\\ -2 & 1 & -3\end{pmatrix}$$

$$=\begin{pmatrix}1 & 2 & -1\\ 0 & -7 & 7\\ 0 & -2 & 2\\ 0 & 5 & -5\end{pmatrix}\begin{matrix}\\ R_2\to R_2-3R_1\\ \\ R_4\to R_4+2R_1\end{matrix}$$

$$=\begin{pmatrix}1 & 2 & -1\\ 0 & -7 & 7\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}\begin{matrix}\\ \\ R_3\to 7R_3-2R_2\\ R_4\to 7R_4+5R_2\end{matrix}$$

Here $r(A)=2<$ number of vectors. Hence the vectors are linearly dependent and hence it cannot form the basis for R^4. Note that $v_1-v_2=v_3$.

Example: Show that the set $\{1+2x+x^2,\ 3+x^2,\ x+x^2\}$ form the basis for the vector space $V=P_2$, the set of all polynomials function of degree 2.

Let $B=\{1+2x+x^2,\ 3+x^2,\ x+x^2\}$.

Consider the vector equation $\alpha_1 v_1+\alpha_2 v_2+\alpha_3 v_3=0$. That is

$$\alpha_1\left(1+2x+x^2\right)+\alpha_2\left(3+x^2\right)+\alpha_3\left(x+x^2\right)=\left(0+0x+0x^2\right)$$

$$\left(\alpha_1+3\alpha_2+\alpha_3\right)+\left(2\alpha_1+0\alpha_2+\alpha_3\right)x+\left(\alpha_1+\alpha_2+\alpha_3\right)x^2=\left(0+0x+0x^2\right)$$

Equating the corresponding coefficients, we have the system of equations

$$\alpha_1+3\alpha_2+\alpha_3=0$$

$$2\alpha_1+0\alpha_2+\alpha_3=0$$

$$\alpha_1+\alpha_2+\alpha_3=0$$

To solve the system of equation by Gauss Elimination method, consider the augmented matrix

$$(A,B)=\begin{pmatrix}1 & 3 & 1 & 0\\ 2 & 0 & 1 & 0\\ 1 & 1 & 1 & 0\end{pmatrix}$$

$$= \begin{pmatrix} 1 & 3 & 1 & 0 \\ 0 & -6 & -1 & 0 \\ 0 & -2 & 0 & 0 \end{pmatrix} \begin{matrix} \\ R_2 \to R_2 - 2R_1 \\ R_3 \to R_3 - R_1 \end{matrix}$$

$$= \begin{pmatrix} 1 & 3 & 1 & 0 \\ 0 & -6 & -1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} R_3 \to 3R_3 - R_2$$

This $\left[r(A) = 3 = no.\ of\ \ vectors \right]$ implies that the system has unique trivial solution $\alpha_1 = \alpha_2 = \alpha_3 = 0$. So, the vectors are linearly independent.

To prove the vectors forms the basis, we have to find scalars $l,\ m,\ n$ such that

$l\left(1+2x+x^2\right)+m\left(3+x^2\right)+n\left(x+x^2\right)=\left(a+bx+cx^2\right)$

$\left(l+3m\right)+\left(2l+n\right)x+\left(l+m+n\right)x^2=\left(a+bx+cx^2\right)$

Equating the coefficients, we have $\left(l+3m\right)=a,\ \left(2l+n\right)=b,\ \left(l+m+n\right)=c$.

Now solving we get $l,\ m,\ n$ in terms of $a,\ b,\ c$ as

$l=\frac{1}{4}\left(7a+3b-3c\right),\ \ m=\frac{1}{4}\left(-a-b+c\right),\ \ n=\frac{1}{4}\left(-14a-2b+6c\right)$. (to solve apply $X = A^{-1}B$)

There the set of vectors $\left\{1+2x+x^2,\ 3+x^2,\ x+x^2\right\}$ form the basis for the vector space $V = P_2$.

Example: Show that the matrices $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ and $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ generate $M_{2\times2}(R)$.

We have the show that the given vectors form the basis for the vector space $M_{2\times2}(R)$.

i.e. any general matrix $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ of $M_{2\times2}(R)$ can be expressed as a linear combination of the given vectors.

Consider $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = m\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} + n\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} + c\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} + k\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ where $m,\ n,\ c,\ k$ are scalars.

It forms the simultaneous equations, as

$$a_{11} = n + c + k$$
$$a_{12} = m + c + k$$
$$a_{21} = m + n + c$$
$$a_{22} = m + n + k$$

Solving, we get
$$m = \frac{1}{3}\left(-2a_{11} + a_{12} + a_{21} + a_{22}\right)$$
$$n = \frac{1}{3}\left(a_{11} - 2a_{12} + a_{21} + a_{22}\right)$$
$$c = \frac{1}{3}\left(a_{11} + a_{12} + a_{21} - 2a_{22}\right)$$
$$k = \frac{1}{3}\left(a_{11} + a_{12} - 2a_{21} + a_{22}\right)$$

Hence the given vectors generates $M_{2\times 2}(R)$.

The Dimension of a Vector Space

Definition : If a vector space has a basis consisting of n vectors, then the number n is called the dimension of V, denoted by $\dim(V) = n$.

Example:

1. The dimension of $\mathbb{R}^n$ with the standard operations is n.
2. The dimension of P_n with the standard operations is $n+1$.
3. The dimension of $M_{m\times n}$ with the standard operations is mn.

Note:

- A vector space is called finite-dimensional if it has a basis consisting of a finite number of vectors
- A vector space that is not finite-dimensional is called infinite-dimensional.
- If V consists of the zero vector alone, the dimension of V is defined as zero.

Result: If W is a subspace of an n-dimensional vector space V, then the dimension of W is less than or equal to n.

Example: Determine the dimension of the sub space $W = \{(b,\ a-b,\ a) : a, b \in \mathbb{R}\}$ of $\mathbb{R}^3$.

Given vector can be expressed as

Example: Determine the dimension of the sub space $W = \{(3a,\ a,\ 0) :\ a \in \mathbb{R}\}$ of $\mathbb{R}^3$.

Given vector can be expressed as

$(3a,\ a,\ 0) = a(3, 1, 0)$

$$(b,\ a-b,\ a)=(0,\ a,\ a)+(b,\ -b,\ 0)$$
$$=a(0,1,1)+b(1,-1,0)$$

i.e. W is spanned by the set

$S=\{(0,1,1),\ (1,-1,0)\}$

Clearly the vectors are linearly independent.

$(\because 2<3)$

Therefore S is the basis for W.
Therefore W is a 2 dimensional subspace of R^3.

i.e. W is spanned by the set $S=\{(3,1,0)\}$

Clearly the vectors are linearly independent. $(\because 1<3)$

Therefore S is the basis for W.

Therefore W is a 1 dimensional subspace of R^3.

Example: Determine the dimension of the sub space $W=$ {The set of symmetric matrices} of $M_{2\times 2}$.

Any symmetric matrix can be expressed as $\begin{pmatrix} a & b \\ b & c \end{pmatrix}=\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}+\begin{pmatrix} 0 & b \\ b & 0 \end{pmatrix}+\begin{pmatrix} 0 & 0 \\ 0 & c \end{pmatrix}$

$$=a\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}+b\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}+c\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

i.e. W is spanned by the set $S=\left\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right\}$

Clearly the vectors are linearly independent. Therefore S is the basis for W.
Therefore W is a 3 dimensional subspace of $M_{2\times 2}$.

EXERCISE

1. Determine whether the set of vectors $s=\{(1,2,3),\ (0,1,2),\ (-2,0,1)\}$ in R^3 is linearly independent or linearly dependent.
2. Are the vectors below linearly independent?
   $s=\{(8,3,0,-2),\ (4,11,-4,6),\ (2,0,1,1),\ (0,-2,-7,7)\}$
3. Determine whether the set $S=\{1+x-2x^2,\ 2+5x-x^2,\ x+x^2\}$ in P_2 is dependent. If yes, express v_2 in terms of other vectors.
4. Determine whether the set $B=\{(1,1),\ (1,-1)\}$ is the basis for R^2.
5. Determine whether the set $\{-1-x+2x^2,\ 2+x-2x^2,\ 1-2x+4x^2\}$ form the basis for the vector space $V=P_2$, the set of all polynomials function of degree 2.
6. Determine whether the set $\{x^2+3x-2,\ 2x^2+5x-3,\ -x^2-4x+4\}$ form the basis for the vector space $V=P_2$, the set of all polynomials function of degree 2.
7. Determine whether the set of all 2×2 matrices of the form $\begin{pmatrix} a & a+b \\ a+b & b \end{pmatrix}$, $a,\ b\in R$ with respect to standard matrix addition and scalar multiplication is a vector space or not?.
8. Let $V=R^3$ and $W=\{(x,\ y,\ z)\ :\ 2x-7y+z=0\}$. Verify whether W is a subspace or not.
9. For what value of k, the vector $(1,\ -2,\ k)$ in R^3 is a liner combination of the vectors $(3,\ 0,\ -2)$ and $(2,\ -1,\ -5)$.

ANSWERS

1. Independent 2. Dependent 3. $v_2=2v_1+3v_3$

4. Yes 5. No 6. Yes 7. Yes 8. Yes 9. $k=-8$

Unit-V Linear Transformation and Inner Product Spaces

Function Terminologies

Let $T : V \to W$ be a map from a vector space V into a vector space W. Here V is called the domain of T and W is called the codomain of T.

If $v \in V$ and $w \in W$ such that $T(v) = w$ then w is called the image of v under T and v is called preimage of w under T.

The set of all images in W is called the range of T.

Example: For any vector $v = (a,\ b)$ in R^2, let $T : R^2 \to R^2$ be defined by $T(a,\ b) = (a+b,\ a-2b)$.
(i) Find the image of $v = (1,\ -2)$. (ii) Find the preimage of $w = (3,\ 0)$.

(i) From the definition, $T(1,\ -2) = (1-2,\ 1+4) = (-1,\ 5)$, which is the image.

(ii) Given that $T(a,\ b) = (3,\ 0)$

$$(a+b,\ a-2b) = (3,\ 0)$$

$$a+b = 3$$
$$a-2b = 0$$

Solving these equations, we get $a = 2,\ b = 1$.
Hence the preimage is $(2,\ 1)$.

Linear Transformations

Definition : Let V and W be vector spaces over F. The function $T : V \to W$ is called a linear transformation of V into W if the following two properties are true for all u and v in V and for any scalar c.
(i) $T(u+v) = T(u) + T(v)$ (ii) $T(cu) = cT(u)$

Note:

- A linear transformation is said to be operation preserving, because the same result occurs whether the operations of addition and scalar multiplication are performed before or after the linear transformation is applied.

https://doi.org/10.5281/zenodo.15090459

- A linear transformation $T: V \to V$ from a vector space into itself is called a linear operator.

Example: The zero transformation $T: V \to W$ such that $T(v) = 0$ for all $v \in V$ and the **identity** transformation $T: V \to V$ such that $T(v) = v$ for all $v \in V$ are linear transformations.

Example: Show that the function $T(a, b) = (a + b, a - 2b)$ is a linear transformation from R^2 into R^2.

let $u = (u_1, u_2)$ and $v = (v_1, v_2)$ be any two vectors in R^2 and let c be any real number.

Then by definition $T(u) = T(u_1, u_2) = (u_1 + u_2, u_1 - 2u_2)$.

Also $u + v = (u_1, u_2) + (v_1, v_2) = (u_1 + v_1, u_2 + v_2)$ and $cu = c(u_1, u_2) = (cu_1, cu_2)$

Now

$T(u + v) = T(u_1 + v_1, u_2 + v_2)$

$= ((u_1 + v_1) + (u_2 + v_2), (u_1 + v_1) - 2(u_2 + v_2))$

$= (u_1 + v_1 + u_2 + v_2, u_1 + v_1 - 2u_2 - 2v_2)$

$= (u_1 + u_2 + v_1 + v_2, u_1 - 2u_2 + v_1 - 2v_2)$

$= (u_1 + u_2, u_1 - 2u_2) + (v_1 + v_2, v_1 - 2v_2)$

$= T(u) + T(v)$

Also

$T(cu) = T(cu_1, cu_2)$

$= (cu_1 + cu_2, cu_1 - 2cu_2)$

$= c(u_1 + u_2, u_1 - 2u_2)$

$= cT((u_1, u_2))$

$= cT(u)$

Hence T is a linear transformation

Try this: Show that the function $T(a, b) = (2a - 3b, a + 4b)$ is a linear transformation from R^2 into R^2.

Example: $T(x) = \sin x$ is not a linear transformation from R into R because, in general $\sin(u + v) \neq \sin u + \sin v$.

Example: $T(x) = x + 1$ is a linear function in R but it is not a linear transformation from R into R. Why?

Properties

- $T(0) = 0$. Because $T(0) = T(0v) = 0T(v) = 0$

- $T(-v) = -T(v)$. Because $T(-v) = T((-1)v) = (-1)T(v) = -T(v)$
- $T(u-v) = T(u) - T(v)$. Because $T(u-v) = T(u+(-1)v) = T(u) + (-1)T(v) = T(u) - T(v)$
- If $v = c_1v_1 + c_2v_2 + \cdots + c_nv_n$ then $T(v) = T(c_1v_1 + c_2v_2 + \cdots + c_nv_n) = c_1T(v_1) + c_2T(v_2) + \cdots + c_nT(v_n)$

Example: T is linear if and only if $T(cx+y) = cT(x) + T(y)$ for all $x, y \in V$ and $c \in F$.

Suppose T is linear.

Then $T(cx+y) = T(cx) + T(y)$

$= cT(x) + T(y)$

Conversely, suppose $T(cx+y) = cT(x) + T(y)$

Put $c = 1$. Then $T(x+y) = T(x) + T(y)$

Let $y = 0 \in V$.

$T(cx+0) = cT(x) + T(0)$

$= cT(x) + 0$

$= cT(x)$

Example: Let $T : R^3 \to R^3$ be a linear transformation such that $T(1,0,0) = (0,1,2)$, $T(0,1,0) = (1,1,2)$ and $T(0,0,1) = (-1,1,0)$. Find $T(1,-1,3)$..

Here $(1,-1,3)$. can be written as $(1,-1,3) = 1(1,0,0) - 1(0,1,0) + 3(0,0,1)$. Therefore by the above property,

$$\begin{aligned}
T(1,-1,3) &= 1T(1,0,0) - 1T(0,1,0) + 3T(0,0,1) \\
&= 1(0,1,2) - 1(1,1,2) + 3(-1,1,0) \\
&= (0,1,2) + (-1,-1,-2) + (-3,3,0) \\
&= (-4,3,0)
\end{aligned}$$

Note: The properties help us to identify the functions that are not linear transformations.

For example, consider the transformation $T(u_1, u_2) = (u_1, u_2 - 1)$.

Here $T(0, 0) = (0, -1) \neq (0, 0)$. Therefore T is not linear.

Matrix Representation of a Linear Transformation

Let A be an $m\times n$ matrix. Let $T: R^n \to R^m$ be a linear transformation defined by $T(v) = Av$ where v is an $n\times 1$ matrix.

By definition, $T(u+v) = A(u+v) = Au + Av = T(u) + T(v)$ and $T(cu) = A(cu) = cAu = cT(u)$

Therefore T is linear.

Example: Let the linear transformation $T: R^2 \to R^3$ is defined as $T(v) = Av = \begin{pmatrix} 3 & 0 \\ 2 & 1 \\ -1 & -2 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$.

Find $T(v)$, where $v = (2, -1)$.

$$T(v) = Av = \begin{pmatrix} 3 & 0 \\ 2 & 1 \\ -1 & -2 \end{pmatrix}\begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 6 \\ 3 \\ 0 \end{pmatrix}. \text{ Therefore } T(2, -1) = (6, 3, 0)$$

Example: The linear transformation $T: R^n \to R^m$ is defined by $T(v) = Av$. Find the dimensions of R^n and R^m for the linear transformation represented by the matrix $\begin{pmatrix} 1 & 2 & -1 & 1 \\ 2 & 3 & 3 & 2 \end{pmatrix}$.

Since the order of this matrix is 2×4 it defines a linear transformation from R^4 into R^2.

Example: Let the linear transformation $T: R^2 \to R^3$ is defined as $T(v) = Av$ where $A = \begin{pmatrix} 1 & 2 \\ -2 & 4 \\ -2 & 2 \end{pmatrix}$.

Find the image of $v = (2, 4)$. Also find the preimage of $w = (-1, 2, 2)$. Explain why the vector $w = (1, 1, 1)$ has no preimage under this transformation.

Given linear transformation transforms an ordered pair into a triplet i.e. $T(v_1, v_2) = (u_1, u_2, u_3)$ and $T^{-1}(u_1, u_2, u_3) = (v_1, v_2)$.

The image of the vector is v given by

$$T(v) = Av$$

$$T(v) = w$$

$$\begin{pmatrix} 1 & 2 \\ -2 & 4 \\ -2 & 2 \end{pmatrix}\begin{pmatrix} 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 10 \\ 12 \\ 4 \end{pmatrix}.$$ Therefore $T(2,\ 4) = (10,\ 12,\ 4)$

Let $v = (v_1, v_2)$ be the preimage of $w = (-1,\ 2,\ 2)$. Then by definition,

$$T(v) = w$$

$$\begin{pmatrix} 1 & 2 \\ -2 & 4 \\ -2 & 2 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \\ 2 \end{pmatrix}$$

$$v_1 + 2v_2 = -1$$

$$-2v_1 + 4v_2 = 2$$

$$-2v_1 + 2v_2 = 2$$

Solving the above equations, we get $v = (v_1, v_2) = (-1,\ 0)$

To find the preimage for $w = (1, 1, 1)$.

Then by definition,

$$T(v) = w$$

$$\begin{pmatrix} 1 & 2 \\ -2 & 4 \\ -2 & 2 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$v_1 + 2v_2 = 1$$

$$-2v_1 + 4v_2 = 1$$

$$-2v_1 + 2v_2 = 1$$

Since the equations are inconsistent, it has no solution and hence $w = (1, 1, 1)$ has no preimage.

Example: Find the matrix of the linear transformation $T : V_3(R) \rightarrow V_3(R)$ given by $T(a,\ b,\ c) = (3a,\ a-b,\ 2a+b+c)$ with respect to the standard basis $\{e_1,\ e_2,\ e_3\}$.

$T(e_1) = T(1,\ 0,\ 0) = (3, 1,\ 2)$

$T(e_2) = T(0, 1, 0) = (0, -1, 1)$

$T(e_3) = T(0, 0, 1) = (0, 0, 1)$

Thus the matrix representing T is $\begin{pmatrix} 3 & 1 & 2 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$.

Example: Find the linear transformation $T : V_3(R) \to V_3(R)$ determined by the matrix $\begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ -1 & 3 & 4 \end{pmatrix}$ with respect to the standard basis $\{e_1, e_2, e_3\}$.

$T(e_1) = e_1 + 2e_2 + e_3 = (1, 2, 1)$

$T(e_2) = 0e_1 + e_2 + e_3 = (0, 1, 1)$

$T(e_3) = -e_1 + 3e_2 + 4e_3 = (-1, 3, 4)$

$$\begin{aligned}(a, b, c) &= a(1, 0, 0) + b(0, 1, 0) + c(0, 0, 1) \\ &= ae_1 + be_2 + ce_3\end{aligned}$$

$$\begin{aligned}T(a, b, c) &= T(ae_1 + be_2 + ce_3) \\ &= a\,T(e_1) + b\,T(e_2) + c\,T(e_3) \\ &= a(1, 2, 1) + b(0, 1, 1) + c(-1, 3, 4) \\ &= (a - c, 2a + b + 3c, a + b + 4c)\end{aligned}$$

The Kernel and Range of a Linear Transformation

Definition : Let $T : V \to W$ be a linear transformation. Then the set of all vectors $v \in V$ such that $T(v) = 0$ is called the Kernel of T and is denoted by $\ker(T)$.

Note:

- The kernel of a linear transformation $T : V \to W$ is a subspace of the domain V.
- The kernel of T is sometimes called the null space of T.

Example:

1. Consider the zero transformation $T:V \to W$ such that $T(v)=0$ for all $v \in V$. Hence $\ker(T)=V$.

2. Consider the identity transformation $T:V \to V$ such that $T(v)=v$ for all $v \in V$. Hence $\ker(T)=\{0\}$.

3. Consider the linear transformation $T:M_{1\times 3} \to M_{3\times 1}$ such that $T(A)=A^T$ for all $A \in M_{1\times 3}$.

   i.e. all matrices of order 1×3 is mapped to its transpose. Hence $\ker(T)=$ the zero matrix of order 1×3.

4. Consider the transformation $T:R^3 \to R^3$ such that $T(x,y,z)=(x,y,0)$. Hence $\ker(T)=\{(0,0,z);\, z\in R\}$

Example: Find the kernel of the linear transformation $T:R^2 \to R^2$ represented by $T(x_1, x_2)=(x_1-x_2,\, x_2-x_1)$.

The $\ker(T)$ is the set of all $x=(x_1, x_2)$ in R^2 such that $T(x_1, x_2)=(x_1-x_2,\, x_2-x_1)=(0,\,0)$.

This is equivalent to the system of equations

$$x_1-x_2=0$$
$$x_2-x_1=0$$

i.e. the solution is all points of the form $x_1=x_2$..

Therefore $\ker(T)=\{(x_1,\, x_1):\, x_1\in R\}$

Example: Find the kernel of the linear transformation $T:R^3 \to R^2$ represented by $T(v)=Av$, where $A=\begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & 2 \end{pmatrix}$.

The $\ker(T)$ is the set of all $v=(v_1, v_2, v_3)$ in R^3 such that $T(v_1, v_2, v_3)=(0,\,0)$.

But $T(v)=Av$

Therefore $Av = (0, 0)$

$$\begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Writing the augmented matrix of this system in reduced row-echelon form gives

$$\begin{pmatrix} 1 & 0 & 4 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\begin{matrix} v_1 + 4v_3 = 0 \\ v_2 + 2v_3 = 0 \end{matrix} \quad \text{i.e.} \quad \begin{matrix} v_1 = -4v_3 \\ v_2 = -2v_3 \end{matrix}$$

Let $v_3 = k$. Therefore $v_1 = -4v_3 = -4k, \quad v_2 = -2v_3 = -2k$

The family of solutions are

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} -4k \\ -2k \\ k \end{pmatrix} = k \begin{pmatrix} -4 \\ -2 \\ 1 \end{pmatrix}$$

Therefore $\ker(T) = \{k(-4, -2, 1) : k \in R\}$. $= \text{Span}\{(-4, -2, 1)\}$

Remark: Here one basis for $\ker(T)$ is $B = \{(-4, -2, 1)\}$

Definition: Let $T : V \to W$ be a linear transformation. Then the set of all vectors $w \in W$ such that $T(v) = w$ is called the Range of T and is denoted by $\text{range}(T)$.

Example:

1. Consider the zero transformation $T : V \to W$ such that $T(v) = 0$ for all $v \in V$. Hence $\text{range}(T) = \{0\}$.
2. Consider the identity transformation $T : V \to V$ such that $T(v) = v$ for all $v \in V$. Hence $\text{range}(T) = V$.

Theorem : The range of a linear transformation $T : V \to W$ is a subspace of W.

Corollary: Let $T: R^n \to R^m$ be the linear transformation given by $T(v) = Av$. Then the column space of A is equal to the range of T. (In the echelon form, the columns which contains 1 in the diagonal is called column space)

Definition : Let $T: V \to W$ be a linear transformation. The dimension of the kernel of T is called the nullity of T. The dimension of the range of T is called the rank of T.

Note: If T is provided by a matrix A then the rank of T is equal to the rank of A.

Dimension Theorem: Let $T: V \to W$ be a linear transformation from an n-dimensional vector space V into a vector space W. Then the sum of the dimensions of the range and kernel is equal to the dimension of the domain.

i.e. dim(range) + dim(kernel) = dim(domain)

$$\text{rank}(T) + \text{nullity}(T) = n$$

Assume that the linear transformation T is represented by an $m \times n$ matrix A.

Assume that the matrix A has a rank r.

Now $\text{rank}(T) = \dim(\text{range of } T) = \dim(\text{column space}) = \text{rank}(A) = r$

Since matrix A has a rank r, its equivalent row reduced echelon matrix B is $r \times r$ identity matrix.

Also $\text{nullity}(T) = \dim(\text{kernel of } T) = \dim(\text{solution space}) = n - r$.

Therefore $\text{rank}(T) + \text{nullity}(T) = n$

Example: Let $T: R^5 \to R^4$ be a linear transformation defined by $T(v) = Av$ where v is in R^5 and $A = \begin{pmatrix} 1 & 2 & 0 & 1 & -1 \\ 2 & 1 & 3 & 1 & 0 \\ -1 & 0 & -2 & 0 & 1 \\ 0 & 0 & 0 & 2 & 8 \end{pmatrix}$.

i. Find the basis for kernel of T and hence find $\text{nullity}(T)$.

ii. Find the basis for range of T and hence find $\text{rank}(T)$.

The $\ker(T)$ is the set of all $v = (v_1, v_2, v_3, v_4, v_5)$ in R^5 such that $T(v_1, v_2, v_3, v_4, v_5) = (0, 0, 0, 0)$.

But $T(v) = Av$. Therefore $Av = (0,\ 0,\ 0,\ 0)$

$$\begin{pmatrix} 1 & 2 & 0 & 1 & -1 \\ 2 & 1 & 3 & 1 & 0 \\ -1 & 0 & -2 & 0 & 1 \\ 0 & 0 & 0 & 2 & 8 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Writing the augmented matrix of this system in reduced row-echelon form gives

$$\begin{pmatrix} 1 & 2 & 0 & 1 & -1 \\ 0 & -3 & 3 & -1 & 2 \\ 0 & 2 & -2 & 1 & 0 \\ 0 & 0 & 0 & 2 & 8 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 3 & 0 & 6 & 1 & 1 \\ 0 & -3 & 3 & -1 & 2 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 2 & 8 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 3 & 0 & 6 & 0 & -3 \\ 0 & -3 & 3 & 0 & 6 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 2 & 0 & -1 \\ 0 & 1 & -1 & 0 & -2 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \cdots\cdots(1)$$

$$v_1 + 2v_3 - v_5 = 0$$

$$v_2 - v_3 - 2v_5 = 0$$

$$v_4 + 4v_5 = 0$$

Let $v_3 = m,\ \ v_5 = n$. Therefore

$v_4 = -4v_5 = -4n$

$v_2 = v_3 + 2v_5 = m + 2n$

$v_1 = -2v_3 + v_5 = -2m + n$

The family of solutions are

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{pmatrix} = \begin{pmatrix} -2m+n \\ m+2n \\ m \\ -4n \\ n \end{pmatrix} = \begin{pmatrix} -2m \\ m \\ m \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} n \\ 2n \\ 0 \\ -4n \\ n \end{pmatrix} = m\begin{pmatrix} -2 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + n\begin{pmatrix} 1 \\ 2 \\ 0 \\ -4 \\ 1 \end{pmatrix}$$

Therefore $\ker(T) = \{ m(-2, 1, 1, 0, 0),\ n(1, 2, 0, -4, 1)\ ;\ m,\ n \in R \}$.

Therefore one basis for $\ker(T)$ is $B = \{ (-2, 1, 1, 0, 0), (1, 2, 0, -4, 1) \}$

Therefore $\text{nullity}(T) = \dim(\text{kernal}) = 2$

Since the leading 1's appear in columns 1, 2, and 4 of the reduced echelon form of A, which was calculated in equation (1), the corresponding column vectors of A form a basis for the range of T.

One basis for the range of T is $B = \{ (1, 2, -1, 0), (2, 1, 0, 0), (1, 1, 0, 2) \}$.

Therefore $\text{rank}(T) = \dim(\text{range}) = 3$

Also $\text{rank}(T) + \text{nullity}(T) = 3 + 2 = 5 = \dim(\text{domain})$.

Example: Find the rank and nullity of the linear transformation $T : R^3 \to R^3$ defined by the matrix $A = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}$.

Since A is in row-echelon form and has two nonzero rows, it has a rank 2. So, the rank of T is 2, and $\dim(\text{domain}) = 3$.

We know that

$$\text{rank}(T) + \text{nullity}(T) = \dim(\text{domain})$$

$$2 + \text{nullity}(T) = 3$$

$$\text{nullity}(T) = 1$$

Matrices for Linear Transformations

Standard Matrix for a Linear Transformation: Let $T : R^n \to R^m$ be a linear transformation such that $T(e_1) = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}$, $T(e_2) = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}$, ..., $T(e_n) = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}$.

Then the $m \times n$ matrix whose n columns correspond to $T(e_i)$, $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$, is such that $T(v) = Av$ for every v in R^n. A is called the standard matrix for T.

Note:

- The standard matrix for the zero transformation from R^n into R^m is the $m \times n$ zero matrix.
- The standard matrix for the identity transformation from R^n into R^n is I_n.

Example: Find the standard matrix for the linear transformation $T : R^3 \to R^2$ defined by $T(x, y, z) = (x - 2y, 2x + y)$

First find the images of e_1, e_2, e_3.

$$T(e_1) = T\left[\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right] = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad T(e_2) = T\left[\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right] = \begin{pmatrix} -2 \\ 1 \end{pmatrix}, \quad T(e_3) = T\left[\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right] = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Therefore standard matrix for the given linear transformation is $A = \left(T(e_1) : T(e_2) : T(e_3)\right) = \begin{pmatrix} 1 & -2 & 0 \\ 2 & 1 & 0 \end{pmatrix}$.

Transformation Matrix for Nonstandard Bases: Let V and W be finite-dimensional vector spaces with bases B and B' respectively, where $B = \{v_1, v_2, \ldots\ldots, v_n\}$. If $T : V \to W$ is a linear transformation such that

$\left[T(v_1)\right]_{B'} = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}$, $\left[T(v_2)\right]_{B'} = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}$, ………, $\left[T(v_n)\right]_{B'} = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}$, then the $m \times n$ matrix whose n columns corresponds to $\left[T(v_i)\right]_{B'}$, $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$ is such that

$\left[T(v)\right]_{B'} = A[v]_B$ for every v in V.

Example: Let $T : R^2 \to R^3$ be a linear transformation defined by $T(x_1, x_2) = (x_1 + x_2, x_1, x_2)$. Find the matrix for T relative to the bases $B = \{(1, -1), (0, 1)\}$ and $B' = \{(1, 1, 0), (0, 1, 1), (1, 0, 1)\}$. Also find the image of $v = (5, 4)$

Let $B = \{v_1, v_2\}$ and $B' = \{w_1, w_2, w_3\}$.

From the definition of $T(x_1, x_2) = (x_1 + x_2, x_1, x_2)$, we have

$$T(v_1) = T(1, -1) = (0, 1, -1) = 1w_1 + 0w_2 - 1w_3 \quad \text{(solving } (0, 1, -1) = aw_1 + bw_2 + cw_3\text{)}$$

$$T(v_2) = T(0, 1) = (1, 0, 1) = 0w_1 + 0w_2 + 1w_3$$

Therefore the coordinate matrices for $T(v_1)$ and $T(v_2)$ relative to B' are

$$\left[T(v_1)\right]_{B'} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \quad \text{and} \quad \left[T(v_2)\right]_{B'} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

The matrix for T relative to the bases B and B' is formed by using the coordinate matrices as columns..

Therefore $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ -1 & 1 \end{pmatrix}$.

The image of $v = (5, 4)$ is $Av = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 5 \\ 4 \end{pmatrix} = \begin{pmatrix} 5 \\ 0 \\ -1 \end{pmatrix}$

Example: Let $T: R^2 \to R^2$ be a linear transformation defined by $T(x_1, x_2) = (x_1 + x_2,\ 2x_1 - x_2)$. Find the matrix for T relative to the bases $B = \{(1, 2), (-1, 1)\}$ and $B' = \{(1, 0), (0, 1)\}$.

Let $B = \{v_1, v_2\}$ and $B' = \{w_1, w_2\}$.

From the definition of $T(x_1, x_2) = (x_1 + x_2,\ 2x_1 - x_2)$, we have

$$T(v_1) = T(1, 2) = (3, 0) = 3w_1 + 0w_2$$

$$T(v_2) = T(-1, 1) = (0, -3) = 0w_1 - 3w_2$$

Therefore the coordinate matrices for $T(v_1)$ and $T(v_2)$ relative to B' are

$$\left[T(v_1)\right]_{B'} = \begin{pmatrix} 3 \\ 0 \end{pmatrix} \quad \text{and} \quad \left[T(v_2)\right]_{B'} = \begin{pmatrix} 0 \\ -3 \end{pmatrix}$$

The matrix for T relative to the bases B and B' is formed by using the coordinate matrices as columns.

Therefore $A = \begin{pmatrix} 3 & 0 \\ 0 & -3 \end{pmatrix}$.

Inner Product Spaces

Let V be a vector space over F. An inner product on V is a function which assigns to each ordered pairs of vectors u, v in V a scalar in F denoted by $\langle u, v \rangle$ satisfying the following conditions.

i. $\langle u + v,\ w \rangle = \langle u,\ w \rangle + \langle v,\ w \rangle$

ii. $\langle \alpha u, v \rangle = \alpha \langle u, v \rangle$

iii. $\langle u, v \rangle = \overline{\langle v, u \rangle}$, *where* $\overline{\langle v, u \rangle}$ *is the complex conjugate*

iv. $\langle u, u \rangle \geq 0$ *and* $\langle u, u \rangle = 0$ *iff* $u = 0$

A vector space with an inner product defined on it is called an inner product space.

Definition: Let V be an inner product space and let $x \in V$. The norm of x, denoted by $\|x\|$, is defined by $\|x\| = \sqrt{\langle x, x \rangle}$.

Result 1:

$$\langle u, \alpha v\rangle = \overline{\langle \alpha v, u\rangle}$$
$$= \overline{\alpha\langle v, u\rangle}$$
$$= \bar{\alpha}\,\overline{\langle v, u\rangle}$$
$$= \bar{\alpha}\,\langle u, v\rangle$$

Result 2:

$$\langle u, v+w\rangle = \overline{\langle v+w, u\rangle}$$
$$= \overline{\langle v, u\rangle + \langle w, u\rangle}$$
$$= \overline{\langle v, u\rangle} + \overline{\langle w, u\rangle}$$
$$= \langle u, v\rangle + \langle u, w\rangle$$

Result 3:

$$\langle u, 0\rangle = \langle u, 0 0\rangle$$
$$= 0\langle u, 0\rangle$$
$$= 0$$

Similarly $\langle 0, v\rangle = 0$

Example: Show that in an inner product space

$$\langle \alpha u + \beta v, w\rangle = \langle \alpha u, w\rangle + \langle \beta v, w\rangle$$
$$= \alpha\langle u, w\rangle + \beta\langle v, w\rangle$$

$$\langle u, \alpha v + \beta w\rangle = \overline{\langle \alpha v + \beta w, u\rangle}$$
$$= \overline{\langle \alpha v, u\rangle} + \overline{\langle \beta w, u\rangle}$$
$$= \overline{\alpha\langle v, u\rangle} + \overline{\beta\langle w, u\rangle}$$
$$= \bar{\alpha}\,\overline{\langle v, u\rangle} + \bar{\beta}\,\overline{\langle w, u\rangle}$$
$$= \bar{\alpha}\,\langle u, v\rangle + \bar{\beta}\,\langle u, w\rangle$$

Example: Show that $V_2(R)$ is an inner product space with inner product defined by $\langle x, y\rangle = x_1y_1 + x_2y_1 - x_1y_2 + 4x_2y_2$ where $x = (x_1, x_2)$ and $y = (y_1, y_2)$.

Let $x, y, z \in V_2(R)$ *and* $\alpha \in R$.

(i) $\langle x+y, z\rangle = (x_1+y_1)z_1 + (x_2+y_2)z_1 - (x_1+y_1)z_2 + 4(x_2+y_2)z_2$

$$= (x_1z_1 + y_1z_1) + (x_2z_1 + y_2z_1) - (x_1z_2 + y_1z_2) + 4(x_2z_2 + y_2z_2)$$
$$= (x_1z_1 + x_2z_1 - x_1z_2 + 4x_2z_2) + (y_1z_1 + y_2z_1 - y_1z_2 + 4y_2z_2)$$
$$= \langle x, z\rangle + \langle y, z\rangle$$

(ii) $\langle \alpha x, y\rangle = \alpha x_1y_1 + \alpha x_2y_1 - \alpha x_1y_2 + \alpha\, 4x_2y_2$

$$= \alpha\,(x_1y_1 + x_2y_1 - x_1y_2 + 4x_2y_2)$$

(iii) $\langle x, y\rangle = x_1y_1 + x_2y_1 - x_1y_2 + 4x_2y_2$

$$= y_1x_1 + y_1x_2 - y_2x_1 + 4y_2x_2$$
$$= \langle y, x\rangle$$

(iv) $\langle x, x\rangle = x_1 x_1 + x_2 x_1 - x_1 x_2 + 4x_2 x_2 = x_1^2 + 4x_2^2 \geq 0$

$\langle x, x\rangle = 0 \;\; iff \;\; x_1 = x_2 = 0$

Example: Let V be the vector space of polynomials with inner product given by $\langle f, g\rangle = \int_0^1 f(t)\, g(t)\, dt$. Let $f(t) = t+2$ and $g(t) = t^2 - 2t - 3$. Find (*i*) $\langle f, g\rangle$ (ii) $\| f\|$.

$$(i) \quad \langle f, g\rangle = \int_0^t f(t) g(t)\, dt$$

$$= \int_0^t (t+2)(t^2 - 2t - 3)\, dt$$

$$= \int_0^t (t^3 - 7t - 6)\, dt$$

$$= \left(\frac{t^4}{4} - 7\frac{t^2}{2} - 6t \right)_0^1$$

$$= -\frac{37}{4}$$

$$(ii) \quad \langle f, f\rangle = \int_0^t f(t) f(t)\, dt$$

$$= \int_0^t (t+2)^2\, dt$$

$$= \int_0^t (t^2 + 4t + 4)\, dt$$

$$= \left(\frac{t^3}{3} + 4\frac{t^2}{2} + 4t \right)_0^1$$

$$= \frac{19}{3}$$

Therefore $\| f\| = \sqrt{\frac{19}{3}}$

EXERCISE

1. Apply the Gram-Schmidt orthonormalization process to transform the following basis (i) $B=\{(1,1),(0,1)\}$ for R^2 (ii) $B=\{(0,1,0),(1,1,1)\}$ for R^3 into an orthonormal basis. Use the Euclidean inner product for the given vector space.

2. Find an orthonormal basis for the solution space of the homogeneous system of linear equations: $x_1+x_2+0x_3+7x_4=0;\ \ 2x_1+x_2+2x_3+6x_4=0.$

3. Find the least square solution of the system of the system $Ax=b$ where $A=\begin{pmatrix}1 & 1\\ 1 & 2\\ 1 & 3\end{pmatrix}$ and $b=\begin{pmatrix}0\\ 1\\ 3\end{pmatrix}$.

4. Let $T:R^3\to R^3$ be a linear transformation such that $T(1,0,0)=(2,4,-1)$, $T(0,1,0)=(1,3,-2)$ and $T(0,0,1)=(0,-2,2)$. Find $T(0,3,-1)$..

5. Let $T:R^3\to R^3$ be a linear transformation such that $T(1,1,1)=(2,0,-1)$, $T(0,-1,2)=(-3,2,-1)$ and $T(1,0,1)=(1,1,0)$. Find $T(2,1,0)$. Hint: Linear Independence vectors

6. Let the linear transformation $T:R^4\to R^4$ is defined as $T(v)=Av$ where $A=\begin{pmatrix}-1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 2 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}$.
Find the image of $v=(1,1,1,1)$. Also find the preimage of $w=(1,1,1,1)$.

7. Find the kernel of the linear transformation $T:R^2\to R^3$ represented by $T(x_1,x_2)=(x_1+2x_2,-x_1,0)$.

8. Find the kernel of the linear transformation $T:R^3\to R^2$ represented by $T(v)=Av$, where $A=\begin{pmatrix}1 & -1 & -2\\ -1 & 2 & 3\end{pmatrix}$.

9. The linear transformation T is defined by $T(v) = Av$. Find (a) $\ker(T)$ (b) $\text{nullity}(T)$ (c) $\text{range}(T)$ and (d) $\text{rank}(T)$ if (i) $A = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}$ (ii) $A = \begin{pmatrix} 0 & -2 & 3 \\ 4 & 0 & 11 \end{pmatrix}$ (iii) $A = \begin{pmatrix} 5 & -3 \\ 1 & 1 \\ 1 & -1 \end{pmatrix}$

10. Let $T : R^2 \to R^2$ be a linear transformation defined by $T(x_1, x_2, x_3) = (x_1 - x_2, x_2 - x_3)$. Find the matrix for T (i) relative to standard basis (ii) relative to the bases $B = \{(1, 1, 1), (1, 1, 0), (0, 1, 1)\}$ and $B' = \{(1, 2), (1, 1)\}$. Also find the image of $v = (1, 2, -3)$

ANSWERS

(1)(i) $\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right), \left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$ (ii) $(0,1,0), \left(\frac{\sqrt{2}}{2}, 0, \frac{\sqrt{2}}{2}\right)$

2. $\left(-\frac{2}{3}, \frac{2}{3}, -\frac{1}{3}, 0\right); \left(-\frac{3}{\sqrt{30}}, -\frac{4}{\sqrt{30}}, \frac{2}{\sqrt{30}}, \frac{1}{\sqrt{30}}\right)$

3. $X = \begin{pmatrix} -5/3 \\ 3/2 \end{pmatrix}$ 4. $(3, 11, -8)$ 5. $(5, 0, 1)$ 6. $(-1, 1, 2, 1)$, $(-1, 1, 1/2, 1)$

7. $\{(0, 0)\} = \{0\}$

8. $k(-4, -2, 1)$ 9.(i) (a) $\{(0,0)\}$ (b) 0 (c) R^2 (d) 2

9.(ii) (a) $\{(-11k, 6k, 4k) : k \in R\}$ (b) 1 (c) R^2 (d) 2

9.(iii) (a) $\{(0,0)\}$ (b) 0 (c) $\{(4m, 4n, m-n) : m, n \in R\}$ (d) 2

10. (i) $(6, -7)$ (ii) $(-1, 4)$

Inner Product

Definition: Let u, v and w be vectors in a vector space , and let c be any scalar. An inner product on V is a function that associates a real number $\langle u,\ v\rangle$ and satisfying the following axioms.

1. $\langle u,\ v\rangle = \langle v,\ u\rangle$
2. $\langle u,\ v+w\rangle = \langle u,\ v\rangle + \langle u,\ w\rangle$
3. $c\langle u,\ v\rangle = \langle cu,\ v\rangle$
4. $\langle v,\ v\rangle \geq 0$ and $\langle v,\ v\rangle = 0$ if and only if $v = 0$

Note : A vector space with an inner product is called an inner product space.

Example: Euclidean Inner Product (dot product) in R^2.

Let $u = (u_1,\ u_2),\ v = (v_1,\ v_2),\ w = (w_1,\ w_2)$

1. $\langle u,v\rangle = (u_1,\ u_2)\cdot(v_1,\ v_2)$

$$= u_1v_1 + u_2v_2$$
$$= v_1u_1 + v_2u_2$$
$$= \langle v,u\rangle$$

2. $v + w = (v_1 + w_1,\ v_2 + w_2)$

$$\langle u, v+w\rangle = (u_1, u_2)\cdot(v_1 + w_1, v_2 + w_2)$$
$$= u_1v_1 + u_1w_1 + u_2v_2 + u_2w_2$$
$$= u_1v_1 + u_2v_2 + u_1w_1 + u_2w_2$$
$$= \langle u,\ v\rangle + \langle u,\ w\rangle$$

3. $c\langle u,v\rangle = c\left[(u_1, u_2)\cdot(v_1, v_2)\right]$

$$= c[u_1v_1 + u_2v_2]$$
$$= cu_1v_1 + cu_2v_2$$

$$\langle cu,v\rangle = \left[(cu_1, cu_2)\cdot(v_1, v_2)\right]$$
$$= cu_1v_1 + cu_2v_2$$

4. $\langle v,v\rangle = (v_1,\ v_2)\cdot(v_1,\ v_2)$

$$= v_1v_1 + v_2v_2$$
$$= v_1^2 + v_2^2 > 0$$

Example: Show that the function defined as $\langle u,v\rangle = (u_1,\ u_2)\cdot(v_1,\ v_2) = u_1v_1 + 2u_2v_2$ is an Inner Product in R^2.

1. $\langle u,v\rangle = (u_1,\ u_2)\cdot(v_1,\ v_2)$

$$= u_1v_1 + 2u_2v_2$$
$$= v_1u_1 + 2v_2u_2$$
$$= \langle v,u\rangle$$

2. $v + w = (v_1 + w_1,\ v_2 + w_2)$

$$\langle u, v+w\rangle = (u_1, u_2)\cdot(v_1 + w_1, v_2 + w_2)$$
$$= u_1v_1 + u_1w_1 + 2u_2v_2 + 2u_2w_2$$
$$= u_1v_1 + 2u_2v_2 + u_1w_1 + 2u_2w_2$$
$$= \langle u,\ v\rangle + \langle u,\ w\rangle$$

3. $c\langle u,v\rangle = c\left[(u_1, u_2)\cdot(v_1, v_2)\right]$

$$= c[u_1v_1 + 2u_2v_2]$$
$$= cu_1v_1 + c2u_2v_2$$

$$\langle cu,v\rangle = \left[(cu_1, cu_2)\cdot(v_1, v_2)\right]$$
$$= cu_1v_1 + 2cu_2v_2$$

4. $\langle v,v\rangle = (v_1,\ v_2)\cdot(v_1,\ v_2)$

$$= v_1v_1 + 2v_2v_2$$
$$= v_1^2 + 2v_2^2 > 0$$

If $v=0$, then $(v_1, v_2)=(0, 0)$ Therefore $\langle v,v\rangle=0$	If $v=0$, then $(v_1, v_2)=(0, 0)$ Therefore $\langle v,v\rangle=0$

Example: Show that the function defined as $\langle u,v\rangle=(u_1, u_2)\cdot(v_1, v_2)=u_1v_1-2u_2v_2$ is not an Inner Product in R^2.

Consider the fourth axiom: $\langle v,v\rangle=(v_1, v_2)\cdot(v_1, v_2)=v_1v_1-2v_2v_2$

Let $v=(1, 1)$. Then $\langle v,v\rangle=(1, 1)\cdot(1, 1)=1-2=-1<0$, the axiom fails.

Example: For polynomials $p=a_0+a_1x^1+a_2x^2+\cdots+a_nx^n$ and $q=b_0+b_1x^1+b_2x^2+\cdots+b_nx^n$ in the vector space P_n, the function $\langle p, q\rangle=a_0b_0+a_1b_1+a_2b_2+\cdots+a_nb_n$ is an inner product.

Try this: Let $A=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22}\end{pmatrix}$ and $B=\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22}\end{pmatrix}$ be matrices in the vector space $M_{2\times2}$. The function $\langle A, B\rangle=a_{11}b_{11}+a_{12}b_{12}+a_{21}b_{21}+a_{22}b_{22}$ is an inner product on $M_{2\times2}$.

Properties of Inner Products

Let u, v and w be vectors in an inner product space , and let be any real number. Then

1. $\langle 0, v\rangle=\langle u, 0\rangle=0$
2. $\langle u+v, w\rangle=\langle u, w\rangle+\langle v, w\rangle$
3. $\langle u, cv\rangle=c\langle u, v\rangle$

Definition: Let u, v be vectors in an inner product space V.

1. The norm (length) of u is $\|u\|=\sqrt{\langle u, u\rangle}$.
2. The distance between u and v is $d(u,v)=\|u-v\|$
3. The angle between two non zero vectors u and v is $\cos\theta=\dfrac{\langle u, v\rangle}{\|u\|\,\|v\|}$, $0\le\theta\le\pi$
4. u and v are orthogonal if $\langle u, v\rangle=0$.

Note:

- If u is orthogonal to v, then αu is orthogonal to v.
- If u_1 and u_2 is orthogonal to v, then $u_1 + u_2$ is orthogonal to v.
- Zero vector is orthogonal to every vector in V.

Note:

- If $\|u\| = 1$, then u is called a unit vector.
- $\dfrac{u}{\|u\|}$ is called unit vector in the direction of u.

Properties:

- $\|u\| \geq 0$
- $\|u\| = 0$ if and only if $u = 0$
- $\|cu\| = |c|\|u\|$
- $d\,(u,v) \geq 0$
- $d\,(u,v) = 0$ if and only if $u = v$
- $d\,(u,v) = d\,(v,u)$

Example: Let $p = 1 - x + 3x^2$, $q = x - x^2$, be polynomials in P_2. Determine
(i) $\langle p,\ q\rangle$ (ii) $\|p\|$ (iii) $d\,(p,q)$

(i) $\langle p,\ q\rangle = a_0b_0 + a_1b_1 + a_2b_2 = (1)(0) + (-1)(1) + (3)(-1) = -1 - 4 = -5$

(ii) $\langle p,\ p\rangle = a_0a_0 + a_1a_1 + a_2a_2 = (1)(1) + (-1)(-1) + (3)(3) = 1 + 1 + 9 = 11$

$$\|p\| = \sqrt{\langle p,\ p\rangle} = \sqrt{11}$$

(iii) $p - q = \left(1 - x + 3x^2\right) - \left(x - x^2\right) = 1 - 2x + 4x^2$

$$\langle p-q,\ p-q\rangle = c_0c_0 + c_1c_1 + c_2c_2 = (1)(0) + (-2)(-2) + (4)(4) = 4 + 16 = 20$$

$$d\,(p,q) = \|p-q\| = \sqrt{\langle p-q,\ p-q\rangle} = \sqrt{20}$$

Orthonormal Bases

Definition: A set of vectors S in an inner product space V is called orthogonal if every pair of vectors in S is orthogonal. If each vector in the set is a unit vector, then S is called orthonormal.

Note:

- If S is a basis, then it is called an orthogonal basis or an orthonormal basis, respectively.
- If S is an orthogonal set of nonzero vectors in an inner product space V, then S is linearly independent.(Theorem)
- If V is an inner product space of dimension n then any orthogonal set of n nonzero vectors is a basis for V.(Corollary)

Example: The standard basis for R^n is orthonormal.

Example: Show that the set $S = \left\{ \left(\frac{\sqrt{2}}{2}, 0, \frac{\sqrt{2}}{2} \right), \left(-\frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{3}, \frac{\sqrt{6}}{6} \right), \left(\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}, -\frac{\sqrt{3}}{3} \right) \right\}$ is an orthonormal basis for R^3.

To show that the three vectors are mutually orthogonal.

$$v_1 \cdot v_2 = \left(\frac{\sqrt{2}}{2}, 0, \frac{\sqrt{2}}{2} \right) \bullet \left(-\frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{3}, \frac{\sqrt{6}}{6} \right) = -\frac{\sqrt{12}}{12} + 0 + \frac{\sqrt{12}}{12} = 0$$

$$v_1 \cdot v_3 = \left(\frac{\sqrt{2}}{2}, 0, \frac{\sqrt{2}}{2} \right) \bullet \left(\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}, -\frac{\sqrt{3}}{3} \right) = \frac{\sqrt{6}}{6} + 0 - \frac{\sqrt{6}}{6} = 0$$

$$v_2 \cdot v_3 = \left(-\frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{3}, \frac{\sqrt{6}}{6} \right) \bullet \left(\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}, -\frac{\sqrt{3}}{3} \right) = -\frac{\sqrt{18}}{18} + \frac{\sqrt{18}}{9} - \frac{\sqrt{18}}{18} = 0$$

To show that each vector is unit vector

$$\|v_1\| = \sqrt{\langle v_1, v_1 \rangle} = \sqrt{v_1 \cdot v_1} = \sqrt{\left(\frac{\sqrt{2}}{2}, 0, \frac{\sqrt{2}}{2} \right) \cdot \left(\frac{\sqrt{2}}{2}, 0, \frac{\sqrt{2}}{2} \right)} = \sqrt{\frac{2}{4} + 0 + \frac{2}{4}} = 1$$

$$\|v_2\| = \sqrt{\langle v_2, v_2 \rangle} = \sqrt{v_2 \cdot v_2} = \sqrt{\left(-\frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{3}, \frac{\sqrt{6}}{6} \right) \cdot \left(-\frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{3}, \frac{\sqrt{6}}{6} \right)} = \sqrt{\frac{6}{36} + \frac{6}{9} + \frac{6}{36}} = 1$$

$$\|v_3\| = \sqrt{\langle v_3, v_3 \rangle} = \sqrt{v_3 \cdot v_3} = \sqrt{\left(\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}, -\frac{\sqrt{3}}{3} \right) \cdot \left(\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}, -\frac{\sqrt{3}}{3} \right)} = \sqrt{\frac{3}{9} + \frac{3}{9} + \frac{3}{9}} = 1$$

Therefore the given set is an orthonormal basis for R^3.

Example: Show that the set $S = \{(2, 3, 2, -2), (1, 0, 0, 1), (-1, 0, 2, 1), (-1, 2, -1, 1)\}$ is a basis for R^4.

We know that the dimension of R^4 is 4. Also R^4 is an inner product space. Here the set S has 4 non zero vectors.

To show that the four vectors are mutually orthogonal.

$$v_1 \cdot v_2 = (2, 3, 2, -2) \cdot (1, 0, 0, 1) = 2 + 0 + 0 - 2 = 0$$

$$v_1 \cdot v_3 = (2, 3, 2, -2) \cdot (-1, 0, 2, 1) = -2 + 0 + 4 - 2 = 0$$

$$v_1 \cdot v_4 = (2, 3, 2, -2) \cdot (-1, 2, -1, 1) = -2 + 6 - 2 - 2 = 0$$

$$v_2 \cdot v_3 = (1, 0, 0, 1) \cdot (-1, 0, 2, 1) = -1 + 0 + 0 + 1 = 0$$

$$v_2 \cdot v_4 = (1, 0, 0, 1) \cdot (-1, 2, -1, 1) = -1 + 0 + 0 + 1 = 0$$

$$v_3 \cdot v_4 = (-1, 0, 2, 1) \cdot (-1, 2, -1, 1) = 1 + 0 - 2 + 1 = 0$$

Since the set S is orthogonal, by corollary (Note 3), it is the basis for R^4 .

Gram-Schmidt Orthonormalization Process

The procedure for finding a orthonormal basis is known as Gram-Schmidt orthonormalization process, It consists of the following three steps.

1. Begin with a basis for the inner product space. It need not be orthogonal nor consist of unit vectors.
2. Convert the given basis to an orthogonal basis.
3. Normalize each vector in the orthogonal basis to form an orthonormal basis

1. Let $B = \{v_1, v_2, \cdots, v_n\}$ be a basis for an inner product space V .
2. Let $B' = \{w_1, w_2, \cdots, w_n\}$ be the orthogonal basis for V where w_i is given by

$$w_1 = v_1$$

$$w_2 = v_2 - \frac{\langle v_2, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1$$

$$w_3 = v_3 - \frac{\langle v_3, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 - \frac{\langle v_3, w_2 \rangle}{\langle w_2, w_2 \rangle} w_2$$

•

•

•

alternative form of the Gram-Schmidt orthonormalization process

$$u_1 = \frac{w_1}{\|w_1\|} = \frac{v_1}{\|v_1\|}$$

$$u_2 = \frac{w_2}{\|w_2\|} \quad where \ w_2 = v_2 - \langle v_2, u_1 \rangle u_1$$

$$u_3 = \frac{w_3}{\|w_3\|} \quad where \ w_3 = v_3 - \langle v_3, u_1 \rangle u_1 - \langle v_3, u_2 \rangle u_2$$

•

•

•

$$u_n = \frac{w_n}{\|w_n\|} \quad where$$

$$w_n = v_n - \langle v_n, u_1 \rangle u_1 - \langle v_n, u_2 \rangle u_2 - \cdots - \langle v_n, u_{n-1} \rangle u_{n-1}$$

$$w_n = v_n - \frac{\langle v_n, w_1\rangle}{\langle w_1, w_1\rangle} w_1 - \frac{\langle v_n, w_2\rangle}{\langle w_2, w_2\rangle} w_2 - \cdots$$

$$- \frac{\langle v_n, w_{n-1}\rangle}{\langle w_{n-1}, w_{n-1}\rangle} w_{n-1}$$

3. Let $u_i = \frac{w_i}{\|w_i\|}$. Then the set $B'' = \{u_1, u_2, \cdots, u_n\}$ is an orthonormal basis for V.

Note: An orthonormal set derived by the Gram-Schmidt orthonormalization process depends on the order of the vectors in the basis

<table>
<tr><td colspan="2">Apply the Gram-Schmidt orthonormalization process to transform the basis $B = \{(3, 4), (1, 0)\}$ for R^2 into an orthonormal basis. Use the Euclidean inner product for R^2.</td></tr>
<tr><td>Let $B = \{v_1, v_2\}$ be a basis for an inner product space R^2 where $v_1 = (3, 4)$, $v_2 = (1, 0)$

Let $w_1 = v_1 = (3, 4)$

$\langle v_2, w_1\rangle = \langle (1, 0), (3, 4)\rangle = 3 + 0 = 3$

$\langle w_1, w_1\rangle = \langle (3, 4), (3, 4)\rangle = 9 + 16 = 25$

$w_2 = v_2 - \frac{\langle v_2, w_1\rangle}{\langle w_1, w_1\rangle} w_1$
$= (1, 0) - \frac{3}{25}(3, 4)$
$= (1, 0) - \left(\frac{9}{25}, \frac{12}{25}\right)$
$= \left(\frac{16}{25}, -\frac{12}{25}\right)$</td><td>Alternative form of Gram-Schmidt process:

$u_1 = \frac{w_1}{\|w_1\|} = \frac{v_1}{\|v_1\|}$
$= \frac{(3, 4)}{\sqrt{\langle (3, 4), (3, 4)\rangle}}$
$= \frac{(3, 4)}{\sqrt{9 + 16}}$
$= \left(\frac{3}{5}, \frac{4}{5}\right)$

$u_2 = \frac{w_2}{\|w_2\|}$ where</td></tr>
</table>

<table>
<tr>
<td>

Therefore orthonormal basis is $B'' = \{u_1, u_2\}$ and its orthonormal vectors are given by

$$u_1 = \frac{w_1}{\|w_1\|} = \frac{(3, 4)}{\sqrt{25}} = \left(\frac{3}{5}, \frac{4}{5}\right)$$

$$\langle w_2, w_2\rangle = \left\langle \left(\frac{16}{25}, -\frac{12}{25}\right), \left(\frac{16}{25}, -\frac{12}{25}\right)\right\rangle = \frac{256}{625} + \frac{144}{225} = \frac{400}{625}$$

$$u_2 = \frac{w_2}{\|w_2\|} = \frac{\left(\frac{16}{25}, -\frac{12}{25}\right)}{\sqrt{\frac{400}{625}}}$$

$$= \left(\frac{4}{5}, -\frac{3}{5}\right)$$

</td>
<td>

$$w_2 = v_2 - \langle v_2, u_1\rangle u_1$$

$$= (1, 0) - \left\langle (1, 0), \left(\frac{3}{5}, \frac{4}{5}\right)\right\rangle \left(\frac{3}{5}, \frac{4}{5}\right)$$

$$= (1, 0) - \left(\frac{3}{5} + 0\right)\left(\frac{3}{5}, \frac{4}{5}\right)$$

$$= (1, 0) - \left(\frac{9}{25}, \frac{12}{25}\right)$$

$$= \left(\frac{16}{25}, -\frac{12}{25}\right)$$

$$u_2 = \frac{w_2}{\|w_2\|}$$

$$= \frac{\left(\frac{16}{25}, -\frac{12}{25}\right)}{\sqrt{\left\langle \left(\frac{16}{25}, -\frac{12}{25}\right), \left(\frac{16}{25}, -\frac{12}{25}\right)\right\rangle}}$$

$$= \frac{\left(\frac{16}{25}, -\frac{12}{25}\right)}{\sqrt{\frac{256}{625} + \frac{144}{625}}} = \frac{\left(\frac{16}{25}, -\frac{12}{25}\right)}{\frac{20}{25}}$$

$$= \left(\frac{4}{5}, -\frac{3}{5}\right)$$

Therefore orthonormal basis is $B'' = \{u_1, u_2\}$

</td>
</tr>
</table>

.

<table>
<tr>
<td colspan="2">

Apply the Gram-Schmidt orthonormalization process to transform the basis $B = \{(1, 1, 0), (1, 2, 0), (0, 1, 2)\}$ for R^3 into an orthonormal basis. Use the Euclidean inner product for R^3.

</td>
</tr>
<tr>
<td>

Let $B = \{v_1, v_2, v_3\}$ be a basis for an inner product space R^3 where
$v_1 = (1, 1, 0)$, $v_2 = (1, 2, 0)$, $v_3 = (0, 1, 2)$

Let $w_1 = v_1 = (1, 1, 0)$

</td>
<td>

Alternative form of Gram-Schmidt process:

$$u_1 = \frac{w_1}{\|w_1\|} = \frac{v_1}{\|v_1\|}$$

$$= \frac{(1, 1, 0)}{\sqrt{\langle (1, 1, 0), (1, 1, 0)\rangle}}$$

</td>
</tr>
</table>

$\langle v_2, w_1\rangle = \langle(1, 2, 0), (1, 1, 0)\rangle = 1+2+0 = 3$

$\langle w_1, w_1\rangle = \langle(1, 1, 0), (1, 1, 0)\rangle = 1+1+0 = 2$

$$w_2 = v_2 - \frac{\langle v_2, w_1\rangle}{\langle w_1, w_1\rangle} w_1$$
$$= (1, 2, 0) - \frac{3}{2}(1, 1, 0)$$
$$= (1, 2, 0) - \left(\frac{3}{2}, \frac{3}{2}, 0\right)$$
$$= \left(-\frac{1}{2}, \frac{1}{2}, 0\right)$$

$\langle v_3, w_1\rangle = \langle(0, 1, 2), (1, 1, 0)\rangle = 1$

$\langle v_3, w_2\rangle = \left\langle(0, 1, 2), \left(-\frac{1}{2}, \frac{1}{2}, 0\right)\right\rangle = \frac{1}{2}$

$\langle w_1, w_1\rangle = \langle(1, 1, 0), (1, 1, 0)\rangle = 2$

$\langle w_2, w_2\rangle = \left\langle\left(-\frac{1}{2}, \frac{1}{2}, 0\right), \left(-\frac{1}{2}, \frac{1}{2}, 0\right)\right\rangle = \frac{1}{2}$

$$w_3 = v_3 - \frac{\langle v_3, w_1\rangle}{\langle w_1, w_1\rangle} w_1 - \frac{\langle v_3, w_2\rangle}{\langle w_2, w_2\rangle} w_2$$
$$= (0, 1, 2) - \frac{1}{2}(1, 1, 0) - \frac{1}{2}\cdot\frac{2}{1}\left(-\frac{1}{2}, \frac{1}{2}, 0\right)$$
$$= (0, 1, 2) - \left(\frac{1}{2}, \frac{1}{2}, 0\right) - \left(-\frac{1}{2}, \frac{1}{2}, 0\right)$$
$$= (0, 0, 2)$$

Therefore orthonormal basis is $B'' = \{u_1, u_2, u_3\}$ and its orthonormal vectors are given by

$$u_1 = \frac{w_1}{\|w_1\|} = \frac{(1, 1, 0)}{\sqrt{2}} = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right)$$

$$= \frac{(1, 1, 0)}{\sqrt{1+1+0}}$$
$$= \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right)$$

$u_2 = \frac{w_2}{\|w_2\|}$ where

$$w_2 = v_2 - \langle v_2, u_1\rangle u_1$$
$$= (1, 2, 0) - \left\langle(1, 2, 0), \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right)\right\rangle\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right)$$
$$= (1, 2, 0) - \left(\frac{1}{\sqrt{2}} + \frac{2}{\sqrt{2}} + 0\right)\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right)$$
$$= (1, 2, 0) - \left(\frac{3}{2}, \frac{3}{2}, 0\right)$$
$$= \left(-\frac{1}{2}, \frac{1}{2}, 0\right)$$

$$u_2 = \frac{w_2}{\|w_2\|}$$
$$= \frac{\left(-\frac{1}{2}, \frac{1}{2}, 0\right)}{\sqrt{\left\langle\left(-\frac{1}{2}, \frac{1}{2}, 0\right), \left(-\frac{1}{2}, \frac{1}{2}, 0\right)\right\rangle}}$$
$$= \frac{\left(-\frac{1}{2}, \frac{1}{2}, 0\right)}{\sqrt{\frac{1}{4} + \frac{1}{4} + 0}} = \frac{\left(-\frac{1}{2}, \frac{1}{2}, 0\right)}{\frac{1}{\sqrt{2}}}$$
$$= \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right)$$

$u_3 = \frac{w_3}{\|w_3\|}$ where

<table>
<tr>
<td>

$u_2 = \dfrac{w_2}{\|w_2\|} = \dfrac{\left(-\frac{1}{2}, \frac{1}{2}, 0\right)}{\sqrt{\frac{1}{2}}} = \left(-\dfrac{1}{\sqrt{2}}, \dfrac{1}{\sqrt{2}}, 0\right)$

$\langle w_3, w_3 \rangle = (0, 0, 2)\cdot(0, 0, 2) = 4$

$u_3 = \dfrac{w_3}{\|w_3\|} = \dfrac{(0, 0, 2)}{\sqrt{4}} = (0, 0, 1)$

</td>
<td>

$w_3 = v_3 - \langle v_3, u_1 \rangle u_1 - \langle v_3, u_2 \rangle u_2$

$= (0, 1, 2) - \left\langle (0, 1, 2), \left(\dfrac{1}{\sqrt{2}}, \dfrac{1}{\sqrt{2}}, 0\right)\right\rangle \left(\dfrac{1}{\sqrt{2}}, \dfrac{1}{\sqrt{2}}, 0\right)$

$- \left\langle (0, 1, 2), \left(-\dfrac{1}{\sqrt{2}}, \dfrac{1}{\sqrt{2}}, 0\right)\right\rangle \left(-\dfrac{1}{\sqrt{2}}, \dfrac{1}{\sqrt{2}}, 0\right)$

$= (0, 1, 2) - \dfrac{1}{\sqrt{2}}\left(\dfrac{1}{\sqrt{2}}, \dfrac{1}{\sqrt{2}}, 0\right) - \dfrac{1}{\sqrt{2}}\left(-\dfrac{1}{\sqrt{2}}, \dfrac{1}{\sqrt{2}}, 0\right)$

$= (0, 1, 2) - \left(\dfrac{1}{2}, \dfrac{1}{2}, 0\right) - \left(-\dfrac{1}{2}, \dfrac{1}{2}, 0\right)$

$= (0, 0, 2)$

$u_3 = \dfrac{w_3}{\|w_3\|}$

$= \dfrac{(0, 0, 2)}{\sqrt{\langle (0, 0, 2), (0, 0, 2) \rangle}}$

$= \dfrac{(0, 0, 2)}{\sqrt{4}}$

$= (0, 0, 1)$

Therefore orthonormal basis is $B'' = \{u_1, u_2, u_3\}$

</td>
</tr>
</table>

Try this: Apply Gram Schmidt process to construct an orthonormal basis for $V_3(R)$ with the standard inner product for the basis $\{v_1, v_2, v_3\}$ where $v_1 = (1, 0, 1)$, $v_2 = (1, 3, 1)$ and $v_3 = (3, 2, 1)$.

Answer: The orthogonal basis is $\{(1, 0, 1), (0, 3, 0), (1, 0, -1)\}$.

Example: Let V be the set of all polynomials of degree ≤ 2 together with the zero polynomial. V is a real inner product space with inner product defined by $\langle f, g \rangle = \int_{-1}^{1} f(x)\, g(x)\, dx$, Starting with the basis $\{1, x, x^2\}$, obtain an orthonormal basis for V.

Let $v_1 = 1$, $v_2 = x$ and $v_3 = x^2$.

Let $w_1 = v_1 = 1$

$$\langle v_2, w_1\rangle = \int_{-1}^{1} v_2 \cdot w_1 \, dx = \int_{-1}^{1} x \, dx = 0$$

$$\langle w_1, w_1\rangle = \int_{-1}^{1} w_1 \cdot w_1 \, dx = \int_{-1}^{1} 1 \, dx = 2$$

$$\begin{aligned} w_2 &= v_2 - \frac{\langle v_2, w_1\rangle}{\langle w_1, w_1\rangle} w_1 \\ &= x - \frac{0}{2} \cdot 1 \\ &= x \end{aligned}$$

$$\langle v_3, w_1\rangle = \int_{-1}^{1} v_3 \cdot w_1 \, dx = \int_{-1}^{1} x^2 \cdot 1 \, dx = 2\int_{0}^{1} x^2 \cdot 1 \, dx = \frac{2}{3}$$

$$\langle v_3, w_2\rangle = \int_{-1}^{1} v_3 \cdot w_2 \, dx = \int_{-1}^{1} x^2 \cdot x \, dx = 0$$

$$\langle w_1, w_1\rangle = \int_{-1}^{1} w_1 \cdot w_1 \, dx = \int_{-1}^{1} 1 \cdot 1 \, dx = 2$$

$$\langle w_2, w_2\rangle = \int_{-1}^{1} w_2 \cdot w_2 \, dx = \int_{-1}^{1} x \cdot x \, dx = 2\int_{0}^{1} x^2 \, dx = \frac{2}{3}$$

$$\begin{aligned} w_3 &= v_3 - \frac{\langle v_3, w_1\rangle}{\langle w_1, w_1\rangle} w_1 - \frac{\langle v_3, w_2\rangle}{\langle w_2, w_2\rangle} w_2 \\ &= x^2 - \frac{2/3}{2} \cdot 1 - \frac{0}{2/3} \cdot x \\ &= x^2 - \frac{1}{3} \end{aligned}$$

$$\langle w_3, w_3\rangle = \int_{-1}^{1} w_3 \cdot w_3 \, dx = \int_{-1}^{1} \left(x^2 - \frac{1}{3}\right)^2 dx = \frac{8}{45}$$

Hence the orthogonal basis is $\left\{u_1 = 1,\ u_2 = x,\ u_3 = x^2 - \frac{1}{3}\right\}$.

Therefore the orthonormal basis is

$$u_1 = \frac{w_1}{\|w_1\|} = \frac{w_1}{\sqrt{\langle w_1, w_1 \rangle}} = \frac{1}{\sqrt{2}}$$

$$u_2 = \frac{w_2}{\|w_2\|} = \frac{w_2}{\sqrt{\langle w_2, w_2 \rangle}} = \frac{x}{\sqrt{2/3}} = \frac{\sqrt{3}}{\sqrt{2}}x$$

$$u_3 = \frac{w_3}{\|w_3\|} = \frac{w_3}{\sqrt{\langle w_3, w_3 \rangle}} = \frac{3x^2 - 1}{3\sqrt{8/45}} = \frac{\sqrt{5}}{\sqrt{8}}\left(3x^2 - 1\right)$$

Example: Find an orthonormal basis for the solution space of the homogeneous system of linear equations: $2x_1 + x_2 - 6x_3 + 2x_4 = 0;\ x_1 + 2x_2 - 3x_3 + 4x_4 = 0;\ x_1 + x_2 - 3x_3 + 2x_4 = 0$.

Given system of equations can be expressed as $AX = B$ where $A = \begin{pmatrix} 2 & 1 & -6 & 2 \\ 1 & 2 & -3 & 4 \\ 1 & 1 & -3 & 2 \end{pmatrix}$, $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$ and $B = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$. Apply Gauss-Jordan method to reduce the augmented matrix (A, B).

$$(A,B) = \begin{pmatrix} \boxed{2} & 1 & -6 & 2 & 0 \\ 1 & 2 & -3 & 4 & 0 \\ 1 & 1 & -3 & 2 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 2 & 1 & -6 & 2 & 0 \\ 0 & \boxed{3} & 0 & 6 & 0 \\ 0 & 1 & 0 & 2 & 0 \end{pmatrix} \begin{matrix} \\ R_2 \to 2R_2 - R_1 \\ R_3 \to 2R_3 - R_1 \end{matrix}$$

$$= \begin{pmatrix} 6 & 0 & -18 & 0 & 0 \\ 0 & 3 & 0 & 6 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{matrix} R_1 \to 3R_1 - R_2 \\ \\ R_3 \to 3R_3 - R_2 \end{matrix}$$

Let $x_3 = m,\ x_4 = n$, then $3x_2 = 6n,\ x_2 = 2n$ and $6x_1 = -18m,\ x_1 = -3m$.

Therefore the solution of the system has the form

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -3m \\ 2n \\ m \\ n \end{pmatrix} = \begin{pmatrix} -3m \\ 0 \\ m \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 2n \\ 0 \\ n \end{pmatrix} = m\begin{pmatrix} -3 \\ 0 \\ 1 \\ 0 \end{pmatrix} + n\begin{pmatrix} 0 \\ 2 \\ 0 \\ 1 \end{pmatrix}$$

So, one basis for the solution space is $B = \{v_1, v_2\}$ where $v_1 = (-3, 0, 1, 0)$; $v_2 = (0, 2, 0, 1)$

To find an orthonormal basis, use the alternative form of the Gram-Schmidt orthonormalization process, as follows.

$$u_1 = \frac{w_1}{\|w_1\|} = \frac{v_1}{\|v_1\|}$$

$$= \frac{(-3, 0, 1, 0)}{\sqrt{\langle(-3, 0, 1, 0), (-3, 0, 1, 0)\rangle}}$$

$$= \frac{(-3, 0, 1, 0)}{\sqrt{9+0+1+0}}$$

$$= \left(-\frac{3}{\sqrt{10}}, 0, \frac{1}{\sqrt{10}}, 0\right)$$

$$u_2 = \frac{w_2}{\|w_2\|} \quad \text{where}$$

$$w_2 = v_2 - \langle v_2, u_1\rangle u_1$$

$$= (0, 2, 0, 1) - \left\langle (0, 2, 0, 1), \left(-\frac{3}{\sqrt{10}}, 0, \frac{1}{\sqrt{10}}, 0\right)\right\rangle \left(-\frac{3}{\sqrt{10}}, 0, \frac{1}{\sqrt{10}}, 0\right)$$

$$= (0, 2, 0, 1) - (0+0+0+0)\left(-\frac{3}{\sqrt{10}}, 0, \frac{1}{\sqrt{10}}, 0\right)$$

$$= (0, 2, 0, 1)$$

$$u_2 = \frac{w_2}{\|w_2\|}$$

$$= \frac{(0, 2, 0, 1)}{\sqrt{\langle (0, 2, 0, 1), (0, 2, 0, 1) \rangle}}$$

$$= \frac{(0, 2, 0, 1)}{\sqrt{0+4+0+1}}$$

$$= \left(0, \frac{2}{\sqrt{5}}, 0, \frac{1}{\sqrt{5}} \right)$$

Therefore orthonormal basis is $B'' = \{u_1, u_2\}$

Example: Find the least square solution of the system of the system $Ax = b$ where $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \\ 1 & 1 \end{pmatrix}$ and $b = \begin{pmatrix} 2 \\ 0 \\ -3 \end{pmatrix}$.

$$A^T A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 6 & 5 \\ 5 & 6 \end{pmatrix} \quad \text{and} \quad A^T b = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \\ -3 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

The normal equations are $A^T A x = A^T b$

$$\begin{pmatrix} 6 & 5 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$\begin{pmatrix} 6 & 5 \\ 0 & 11 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -11 \end{pmatrix}$$

$$\begin{pmatrix} -66 & 0 \\ 0 & 11 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -66 \\ -11 \end{pmatrix}$$

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Example: Find the least square solution of the system of the system $Ax = b$ where

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \text{ and } b = \begin{pmatrix} 4 \\ -1 \\ 0 \\ 1 \end{pmatrix}.$$

$$A^T A = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & 2 \\ 2 & 2 & 3 \end{pmatrix} \quad \text{and} \quad A^T b = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 4 \\ -1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \\ 3 \end{pmatrix}$$

The normal equations are $A^T A x = A^T b$

$$\begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & 2 \\ 2 & 2 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \\ 3 \end{pmatrix}$$

$$\begin{pmatrix} 3 & 2 & 2 \\ 0 & 5 & 2 \\ 0 & 2 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 \\ -8 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 15 & 0 & 6 \\ 0 & 5 & 2 \\ 0 & 0 & 21 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 36 \\ -8 \\ 21 \end{pmatrix}$$

$$\begin{pmatrix} 315 & 0 & 0 \\ 0 & 105 & 0 \\ 0 & 0 & 21 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 630 \\ -210 \\ 21 \end{pmatrix}$$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}$$

Fit the least square line for the data $(1,2)$, $(2,3)$, $(3,5)$ and $(4,7)$.

Let $y = mx + c$ be the required straight line where m is the slope and c is the y intercept.

Now $\bar{x} = \dfrac{\sum x}{n} = \dfrac{10}{4} = 2.5$ and $\bar{y} = \dfrac{\sum y}{n} = \dfrac{17}{4} = 4.25$

x	y	$x - x$	$y - \bar{y}$	$(x-\bar{x})(y-\bar{y})$	$(x-\bar{x})^2$
1	2	-1.5	-2.25	3.375	2.25
2	3	-0.5	-1.25	0.625	0.25
3	5	0.5	0.75	0.375	0.25
4	7	1.5	2.75	4.125	2.25
10	17			8.5	5

But $m = \dfrac{\sum (x-\bar{x})(y-\bar{y})}{\sum (x-\bar{x})^2} = \dfrac{8.5}{5} = 1.7$

Since $\bar{x}$ and $\bar{y}$ passes through the line $y = mx + c$, we have $\bar{y} = m\bar{x} + c$.

Therefore $4.25 = (1.7)(2.5) + c$

$c = 0$

Therefore the required line equation is $y = (1.7)x$.

EDITORS

Dr.I.Neethi Manickam is currently working as the Principal in St. Joseph's College of Engineering and Technology, Thanjavur. He obtained his Ph.D degree from Anna University, Chennai. He is very active in research and published more than 56 articles in reputed journals and presented 19 papers in conferences. He has acted as the reviewer of many international journals. He has organized many seminars, conferences, FDP and workshops. He has published 5 patents to his credit.

Dr.K.R.Kavitha is currently working as an Assistant Professor in Mechanical Engineering at St. Joseph's College of Engineering and Technology, Thanjavur. She has Completed her Doctorate in Sathyabama Institute of Science and Technology, Chennai. She has more than 20 years of teaching and research experience in both Engineering college and Deemed University. She has Published many articles in reputed indexed journals and international conferences.

ISBN 978-81-981760-0-4

PUBLISHER

DMI FOUNDATIONS

FULLY HUMAN & FULLY ALIVE

ST. JOSEPH'S COLLEGE OF ENGINEERING AND TECHNOLOGY

ELUPATTI, THANJAVUR – 613 403.