# UNIT III

# RANDAM VARIABLES

# RANDOM VARIABLES

A function which takes any value from the sample space and its range is some set of real numbers is called a random variable of the experiment.

* A random variable is not random since it takes value from well defined sample space. It is a fixed value.

Ex

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

Random Variable $X = \{-1, 1\}$

Ex

Let $X = S^2$

$S = \{1, 2, 3, 4, 5, 6\}$

$X = S^2$

$X = \{1, 4, 9, 16, 25, 26\}$

TWO TYPS OF RANDOM VARIABLES

1. Discrete random Variable
2. Continuous random variable

Discrete Random Variable:

A random variable $X$ is a discrete random variable if $X$ can take on only finite number of values in any finite observation interval. Thus discrete random variable has countable number of distinct

Discrete

Ex :

If 3 of 20 tubes are deffective and 4 of them are randomly chosen for inspection, what is the probability that only one of defective tubes will be included ?

Soln :

Four tubes can be selected out of 20 in $20C_4$ ways.

$20C_4$

$nC_r$

Possible ways $= 20C_4$

$n = 20$

$r = 4$

W.K.T

$$nC_r = \frac{n!}{(n-r)!\,r!}$$

$$nC_r = \frac{20!}{(20-4)!\,4!} = \frac{20!}{16!\,4!}$$

$$= \frac{20 \times 19 \times 18 \times 17 \times 16!}{16! \times 4 \times 3 \times 2 \times 1}$$

$$nC_r = 4845$$

Now there are three defective tubes. Only one defective tube should be included in set of four. This can be chosen in $3C_1$ ways.

Thus in a set of four defective tubes one tube should be defective tubes and 3 tubes should be non defective.

and 3 tubes can be selected in

6. Property : 3

If events A and B are not mutually exclusive events

$$P(A+B) = P(A) + P(B) - P(AB)$$

where

$$P(AB) = \lim_{N \to \infty} \frac{N_{AB}}{N}$$

If A and B are mutually exclusive, then the joint probability $P(AB) = 0$.

## CONDITIONAL PROBABILITY

Definition:

Probability of B given that A has occured is represented by $P(B/A)$.

Probability of A given B has occured is represented by $P(A|B)$.

Where $P(B|A)$ & $P(A|B)$ both are called as "CONDITIONAL PROBABILITIES"

$$\boxed{P(B/A) = \frac{P(AB)}{P(A)}} \quad \& \quad \boxed{P(A|B) = \frac{P(AB)}{P(B)}}$$

Here $P(AB)$ is the joint probability of A & B. It has the Commutative Property:

$$P(AB) = P(BA).$$

$^{17}C_3$ ways.

$$N_A = {}^3C_1 \times {}^{17}C_3$$

$$= \frac{3 \times 2!}{2! \times 1} \times \frac{17 \times 16 \times 15 \times 14!}{14! \times 3!}$$

$$= 2040$$

$$P(A) = \frac{\text{Number of favourable ways } (N_A)}{\text{Total Possible ways}}$$

$$= \frac{2040}{4845}$$

$$P(A) = 0.42$$

AXIOMS (PROPERTIES) OF PROBABILITY.

1. $P(A) = P(S) = 1$

2. $0 \leq P(A) \leq 1$

3. For mutually exclusive events, the of event $A+B$ is
$P(A+B) = P(A) + P(B)$ [mutually exclusive)

4. $\boxed{P(\bar{A}) = 1 - P(A)}$ [Property 1]

5. Property 2
If $A_1, A_2, A_3 \ldots A_m$ are mutually exclusive events,
$P(A_1) + P(A_2) + \ldots\ldots + P(A_m) = 1$

(i) To obtain $P(B_1)$ [o/p of channel 1]?

$$P(B_1) = P(B_1/A_1)\,P(A_1) + P(B_1/A_0)\,P(A_0)$$

$$= 0.9 \times 0.7 + 0.01 \times 0.3$$

$$P(B_1) = 0.633$$

(ii) To obtain $P(A_1/B_1)$ [If a '1' is received then what is the probability that the input to the channel was 1.

$$P(A_1/B_1) = \frac{P(A_1)\,P(B_1/A_1)}{P(A_0)P(B_1/A_0) + P(A_1)P(B_1/A_1)}$$

$$= \frac{0.7 \times 0.9}{[0.3 \times 0.01 + 0.7 \times 0.9]}$$

$$\boxed{P(A_1/B_1) = 0.9952}$$

Continuous Random Variable.

A random variable 'x' takes on any value in a whole observation interval, x is called continuous random variable.

Ex:

Noise voltage generated by an electronic amplifier has a continuous amplitude. Therefore sample space 's' of noise voltage amplifier is continuous. Random variable x has continuous range of values. ie uncountable values.

## COMULATIVE DISTRIBUTION FUNCTION (CDF) :

Definition:

The cumulative distribution function (CDF) of a random variable 'X' is the probability that the random variable 'X' takes a value less than or equal to x. Cummulative distribution function (CDF) is given by $F_X(x)$.

$$\boxed{CDF : F_X(x) = P(X \leq x)}$$

## PROPERTIES

1. CDF is bounded between 0 & 1

$$0 \leq F_X(x) \leq 1$$

2. $F_X(-\infty) = 0$ & $F_X(\infty) = 1$

$= 1$ $\qquad x \cdot x_3$

$$F_X(x) = P(X = x_0) + P(X = x_1) + P(X = x_2) + P(X = x_3)$$

$P(X = x_0) = \{CCC\}$ [all correct bit]

$= P(C) \cdot P(C) \cdot P(C)$

$= \frac{3}{5} \times \frac{3}{5} \times \frac{3}{5} = \frac{27}{125}$

$P(X = x_1)$ means [one bit error]

$x_1 = \{CCE, CEC, ECC\}$

$p(X = x_1) = [p(C)\,p(C)\,p(E)]\,[p(C)\,p(E)\,p(C)]\,[p(E)\,p(C)\,p(C)]$

$= 3 \times [p(C) \cdot p(C)\,p(E)]$

$= 3 \times \left[\frac{3}{5} \times \frac{3}{5} \times \frac{2}{5}\right]$

$P(X = x_1) = \frac{54}{125}$

$P(X = x_2)$ means probability of random variable having two digits in error.

$x_2 = \{CEE, ECE, EEC\}$

$p(X = x_2) = [P(C) \cdot P(E) \cdot P(E)] \times 3$

$= \left[\frac{2}{5} \times \frac{2}{5} \times \frac{2}{5}\right] \times 3$

$P(X = x_3)$ means all three digits error.

$$P(X = x_3) = [P(E)\,P(E) \cdot P(E)]$$
$$= \left(\frac{2}{5}\right)\left(\frac{2}{5}\right)\left(\frac{2}{5}\right)$$

$$P(X = x_3) = \frac{8}{125}$$

Now CDF $F_X(x)$ can be calculated at all values of $x$.

(i) $F_X(x) = 0$ for $x < x_0$

$$F_X(x_0) = P(X \leq 0)$$
$$= P(X < x_0) + P(X = x_0)$$
$$= 0 + \frac{27}{125}$$

(ii) $F_X(x_1) = P(X \leq x_1)$
$$= P(X < x_0) + P(X = x_0) + P(X = x_1)$$
$$= 0 + \frac{27}{125} + \frac{54}{125} = \frac{81}{125}$$

(iii) $F_X(x_2) = P(X \leq x_2)$
$$= P(X < 0) + P(X = x_0) + P(X = x_1) + P(X = x_2)$$
$$= 0 + \frac{27}{125} + \frac{54}{125} + \frac{36}{125} = \frac{117}{125}$$

$F_X(x_1) \leq F_X(x_2)$ if $x_1 \leq x_2$

2 A three digit message is transmitted over a noisy channel having a probability of error as $P(E) = \frac{2}{5}$ per digit. Find out the corresponding CDF.

Soln

$$P(E) = \frac{2}{5} \text{ per digit}$$

Probability of correct digit is,

$$P(C) = 1 - P(E)$$

$$P(C) = 1 - \frac{2}{5} = \frac{3}{5} \text{ per digit.}$$

Let C denote the correct digit and E denote error digit. The sample space 'S' will be,

$$S = \{CCC, CCE, CEC, CEE, ECC, ECE, EEC, EEE\}$$

Let random variable X denote the number of errors in the received message. Then the four possible values of random variable will be,

$X = \{$no error, one digit error, two digits in error, all digits in error$\}$

$i = 0, 1, 2, 3$ to denote discrete values of random variable X.

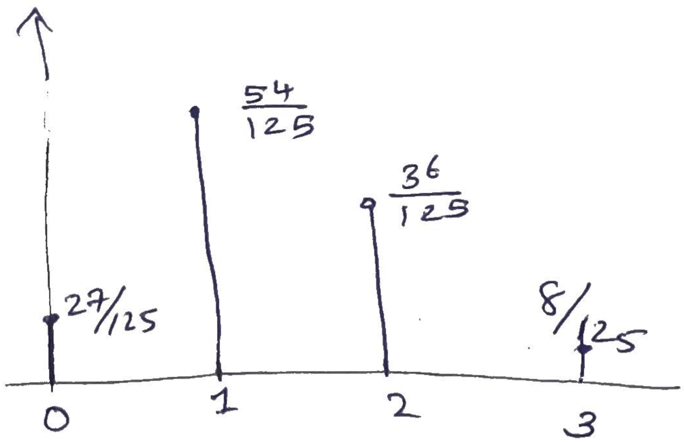
$$F_X(x) = 0 \quad \text{for } x < x_0$$

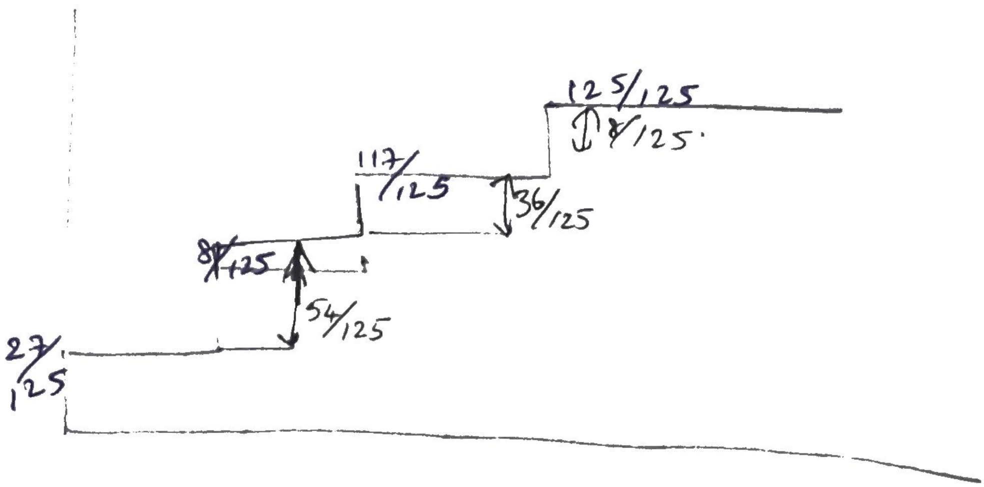
$$= \sum^{3} P(X = x_i) \; ; \; x_0 \leq x \leq x_3$$

$$= P(X < x_0) + P(X = x_0) + P(X = x_1) +$$
$$P(X = x_2) + P(X = x_3)$$

$$= 0 + \frac{27}{125} + \frac{54}{125} + \frac{36}{125} + \frac{8}{125}$$

$$= \frac{125}{125} = 1$$



Pb The Probability density function is given as $f_X(x) = a\, e^{-b|x|}$, where X is a random variable whose allowable values range from $x = -\infty$ to $x = +\infty$. Find
a) Relationship between 'a' and 'b' (b) CDF (c) Probability that Outcome lies between 1 and 2.

Solution:

To find the relationship between a and b

W.K.T the area under PDF Curve is equal to 1.

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

$$\int_{-\infty}^{\infty} a\, e^{-b|x|}\, dx = 1$$

$$\int_{-\infty}^{0} a\, e^{bx}\, dx + \int_{0}^{\infty} a\, e^{-bx}\, dx = 1$$

$$a\left[\frac{e^{bx}}{b}\right]_{-\infty}^{0} + a\left[\frac{e^{-bx}}{-b}\right]_{0}^{\infty} = 1$$

$$\frac{a}{b}\left\{e^{0} - e^{-\infty} - e^{-\infty} + e^{0}\right\} = 1$$

$e^{-\infty} = 0$

$e^{-\infty} = \cancel{\infty}$

$$\frac{a}{b}\{1+1\} = 1$$

$$\frac{a}{b} = \frac{1}{2}$$

$\cancel{a} = \cancel{2}$ This is the relationship

between a and b.

# PROBABILITY DENSITY FUNCTION (PDF)

**Definition:**

The derivative of Cummulative Distribution function (CDF) with respect to some dummy variable is called as Probability Density Function (PDF).

$$\boxed{\text{PDF}: \; f_X(x) = \frac{d}{dx} F_X(x)}$$

$x \rightarrow$ dummy variable.

## PROPERTIES OF PDF

1) Property 1

1) $f_X(x) \geq 0$ for all $x$.

2) Property 2

The area under the PDF curve is equal to 1. ie,

3) Property 3

CDF is obtained by integrating PDF.

$$F_X(x) = \int_{-\infty}^{x} f_X(x) \, dx.$$

4) Property 4

$$= \frac{a}{b}\left\{2 - e^{-bx}\right\}$$

$$= \frac{1}{2}\left\{2 - e^{-bx}\right\}$$

$$x \geq 0 = 1 - \frac{e^{-bx}}{2}$$

Thus the CDF is,

$$F_X(x) = \begin{cases} \frac{1}{2} e^{bx} & \text{for } x < 0 \\ 1 - \frac{1}{2} e^{-bx} & \text{for } x \geq 0 \end{cases}$$

To plot $F_X(x)$ versus $x$,

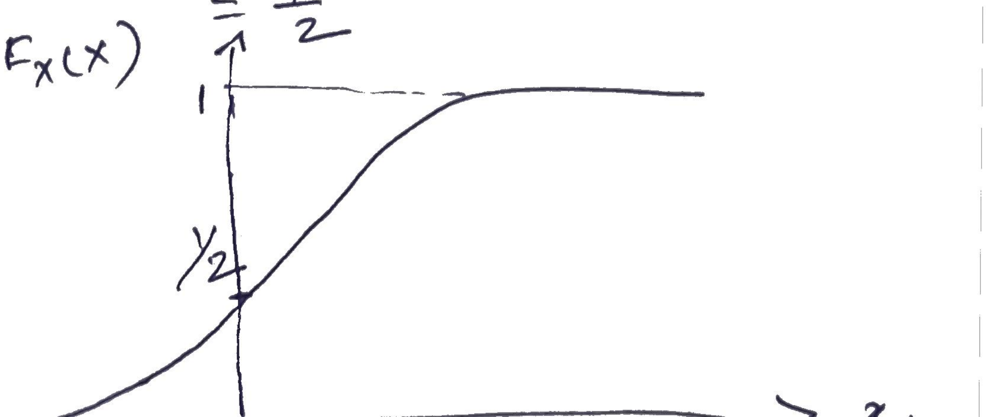
$$F_X(x)\Big|_{x=0} = 0 \quad \text{and} \quad F_X(x)\Big|_{x=\infty} = 1$$

also w.r.t

$$F_X(x)\Big|_{x=0} = 1 - \frac{1}{2} e^{-b \times 0}$$

$$= 1 - \frac{1}{2} e^{0}$$

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



b) To find CDF

We obtain CDF by integrating PDF

$$F_X(x) = \underbrace{\int_{-\infty}^{x} a e^{bx} dx}_{x<0} + \underbrace{\int_{x=0} a e^{-bx} dx}_{x=0} +$$

$$\underbrace{\int_{0^+}^{x} a e^{-bx} dx.}_{x>0}$$

(i) For $x < 0$ /

$$\int_{-\infty}^{x} a e^{bx} dx = \frac{a}{b}\left[e^{bx}\right]_{-\infty}^{x}$$

$$= \frac{a}{b}\left[e^{bx} - e^{-\infty}\right]$$

$$= \frac{a}{b}\left[e^{bx} - 0\right]$$

$$x<0 = \frac{1}{2} e^{bx}.$$

(ii) For $x \geq 0$

$$= \int_{x=0} a e^{bx} dx + \int_{0^+}^{x} a e^{-bx} dx$$

$$= \frac{a}{b}\left[e^{bx}\right]\Big/_{x=0} + \frac{a}{(-b)}\left[e^{-bx}\right]_{0^+}^{x}$$

$$= \frac{a}{b}\left[e^{0}\right] - \frac{a}{b}\left[e^{-bx}\right]_{0}$$

$$\left[e^{-bx} - e^{0+}\right]$$

# RANDOM VARIABLES:

A random variable, usually written X, is a variable whose possible values are numerical outcomes of a random phenomenon. Random variable consists of two types they are discrete and continuous type variable this defines discrete- or continuous-time random processes. Sample function values may take on discrete or continuous a value is defines discrete- or continuous Sample function values may take on discrete or continuous values. This defines discrete- or continuous-parameter random process.

## RANDOM PROCESSES VS. RANDOM VARIABLES:

For a random variable, the outcome of a random experiment is mapped onto variable, e.g., a number.

• For a random processes, the outcome of a random experiment is mapped onto a waveform that is a function of time.Suppose that we observe a random process X(t) at some time t1 to generate the servation X(t1) and that the number of possible waveforms is finite. If Xi(t1) is observed with probability Pi, the collection of numbers {Xi(t1)}, i =1, 2, . . . , n forms a random variable, denoted by X(t1), having the probability distribution Pi, i = 1, 2, . . . , n. E[ • ] = ensemble average operator.

## ✓ DISCRETE RANDOM VARIABLES:

A discrete random variable is one which may take on only a countable number of distinct values such as 0,1,2,3,4,........ Discrete random variables are usually (but not necessarily) counts. If a random variable can take only a finite number of distinct values, then it must be discrete. Examples of discrete random variables include the number of children in a family, the Friday night attendance at a cinema, the number of patients in a doctor's surgery, the number of defective light bulbs in a box of ten.

## PROBABILITY DISTRIBUTION:

The probability distribution of a discrete random variable is a list of probabilities associated with each of its possible values. It is also sometimes called the probability function or the probability mass function. Suppose a random variable X may take k different values, with the probability that X = $x_i$ defined to be P(X = $x_i$) = $p_i$. The probabilities $p_i$ must satisfy the following:

1: $0 < p_i < 1$ for each i

2: $p_1 + p_2 + ... + p_k = 1$.

All random variables (discrete and continuous) have a cumulative distribution function. It is a function giving the probability that the random variable X is less than or equal to x, for every value x. For a discrete random variable, the cumulative distribution function is found by summing up the probabilities.

# CENTRAL LIMIT THEOREM:

In probability theory, the central limit theorem (CLT) states that, given certain conditions, the arithmetic mean of a sufficiently large number of iterates of independent random variables, each with a well-defined expected value and well-defined variance, will be approximately normally distributed.

The Central Limit Theorem describes the characteristics of the "population of the means" which has been created from the means of an infinite number of random population samples of size (N), all of them drawn from a given "parent population". The Central Limit Theorem predicts that regardless of the distribution of the parent population:

[1] The mean of the population of means is always equal to the mean of the parent population from which the population samples were drawn.

[2] The standard deviation of the population of means is always equal to the standard deviation of the parent population divided by the square root of the sample size (N).

[3] The distribution of means will increasingly approximate a normal distribution as the size N of samples increases.

A consequence of Central Limit Theorem is that if we average measurements of a particular quantity, the distribution of our average tends toward a normal one.

In addition, if a measured variable is actually a combination of several other uncorrelated variables, all of them "contaminated" with a random error of any distribution, our measurements tend to be contaminated with a random error that is normally distributed as the number of these variables increases.Thus, the Central Limit Theorem explains the ubiquity of the famous bell-shaped "Normal distribution" (or "Gaussian distribution") in the measurements domain.

Examples:

- Uniform distribution
- Triangular distribution
- 1/X distribution
- Parabolic distribution
- CLT Summary
- more statistical fine-print

The uniform distribution on the left is obviously non-Normal. Call that the parent distribution.

To compute an average, Xbar, two samples are drawn, at random, from the parent distribution and averaged. Then another sample of two is drawn and another value of Xbar computed. This process is repeated, over and over, and averages of two are computed. The distribution of averages of two is shown on the left.

Repeatedly taking three from the parent distribution, and computing the averages, produce the probability density on the left.

# STATIONARY PROCESS:

In mathematics and statistics, a stationary process is a stochastic process whose joint probability distribution does not change when shifted in time. Consequently, parameters such as the mean and variance, if they are present, also do not change over time and do not follow any trends.

Stationary is used as a tool in time series analysis, where the raw data is often transformed to become stationary; for example, economic data are often seasonal and/or dependent on a non-stationary price level. An important type of non-stationary process that does not include a trend-like behaviour is the cyclostationary process.

Note that a "stationary process" is not the same thing as a "process with a stationary distribution". Indeed there are further possibilities for confusion with the use of "stationary" in the context of stochastic processes; for example a "time-homogeneous" Markov chain is sometimes said to have "stationary transition probabilities". Besides, all stationary Markov random processes are time-homogeneous.

- **Definition:**

Formally, let $\{X_t\}$ be a stochastic process and let $F_X(x_{t_1+\tau}, \ldots, x_{t_k+\tau})$ represent the cumulative distribution function of the joint distribution of $\{X_t\}$ at times $t_1 + \tau, \ldots, t_k + \tau$. Then, $\{X_t\}$ is said to be stationary if, for all $k$, for all $\tau$, and for all $t_1, \ldots, t_k$, $F_X(x_{t_1+\tau}, \ldots, x_{t_k+\tau}) = F_X(x_{t_1}, \ldots, x_{t_k})$.

Since $\tau$ does not affect $F_X(\cdot)$, $F_X$ is not a function of time.

- **Wide Sense Stationary:**

Weaker form of stationary commonly employed in signal processing is known as weak-sense stationary, wide-sense stationary (WSS), covariance stationary, or second-order stationary. WSS random processes only require that 1st moment and covariance do not vary with respect to time. Any strictly stationary process which has a mean and a covariance is also WSS.

So, a continuous-time random process x(t) which is WSS has the following restrictions on its mean function.

$$\mathbb{E}[x(t)] = m_x(t) = m_x(t+\tau) \quad \text{for all } \tau \in \mathbb{R}$$

and auto covariance function.

# CORRELATION:

In statistics, dependence is any statistical relationship between two random variables or two sets of data. Correlation refers to any of a broad class of statistical relationships involving dependence. Familiar examples of dependent phenomena include the correlation between the physical statures of parents and their offspring, and the correlation between the demand for a product and its price. Correlations are useful because they can indicate a predictive relationship that can be exploited in practice. For example, an electrical utility may produce less power on a mild day based on the correlation between electricity demand and weather. In this example there is a causal relationship, because extreme weather causes people to use more electricity for heating or cooling; however, statistical dependence is not sufficient to demonstrate the presence of such a causal relationship.

Formally, dependence refers to any situation in which random variables do not satisfy a mathematical condition of probabilistic independence. In loose usage, correlation can refer to any departure of two or more random variables from independence, but technically it refers to any of several more specialized types of relationship between mean values. There are several correlation coefficients, often denoted ρ or r, measuring the degree of correlation. The most common of these is the Pearson correlation coefficient, which is sensitive only to a linear relationship between two variables. Other correlation coefficients have been developed to be more robust than the Pearson correlation that is, more sensitive to nonlinear relationships. Mutual information can also be applied to measure dependence between two variables.

- ✓ **Pearson's correlation coefficient:**

He most familiar measure of dependence between two quantities is the Pearson product-moment correlation coefficient, or "Pearson's correlation coefficient", commonly called simply "the correlation coefficient". It is obtained by dividing the covariance of the two variables by the product of their standard deviations. Karl Pearson developed the coefficient from a similar but slightly different idea by Francis Galton.

The population correlation coefficient ρX,Y between two random variables X and Y with expected values μX and μY and standard deviations σX and σY is defined as:

$$\rho_{X,Y} = \mathrm{corr}(X,Y) = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y},$$

where E is the expected value operator, cov means covariance, and, corr a widely used alternative notation for the correlation coefficient.

The Pearson correlation is defined only if both of the standard deviations are finite and nonzero. It is a corollary of the Cauchy–Schwarz inequality that the correlation cannot exceed 1 in absolute value. The correlation coefficient is symmetric: corr(X,Y) = corr(Y,X).

The Pearson correlation is +1 in the case of a perfect direct (increasing) linear relationship (correlation), −1 in the case of a perfect decreasing (inverse) linear relationship (autocorrelation), and some value between −1 and 1 in all other cases, indicating the degree of linear dependence between the variables. As it approaches zero there is less of a relationship (closer to uncorrelated). The closer the coefficient is to either −1 or 1, the stronger the correlation between the variables.

If the variables are independent, Pearson's correlation coefficient is 0, but the converse is not true because the correlation coefficient detects only linear dependencies between two variables. For example, suppose the random variable X is symmetrically distributed about zero, and $Y = X^2$.

Then Y is completely determined by X, so that X and Y are perfectly dependent, but their correlation is zero; they are uncorrelated. However, in the special case when X and Y are jointly normal, uncorrelatedness is equivalent to independence.

If we have a series of n measurements of X and Y written as xi and yi where i = 1, 2, ..., n, then the sample correlation coefficient can be used to estimate the population Pearson correlation r between X and Y.

where x and y are the sample means of X and Y, and sx and sy are the sample standard deviations of X and Y.

This can also be written as:

$$r_{xy} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{(n-1)s_x s_y} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 - (\sum x_i)^2}\ \sqrt{n\sum y_i^2 - (\sum y_i)^2}}.$$

If x and y are results of measurements that contain measurement error, the realistic limits on the correlation coefficient are not −1 to +1 but a smaller range.

# COVARIANCE FUNCTIONS:

In probability theory and statistics, covariance is a measure of how much two variables change together, and the covariance function, or kernel, describes the spatial covariance of a random variable process or field. For a random field or stochastic process Z(x) on a domain D, a covariance function C(x, y) gives the covariance of the values of the random field at the two locations x and y:

$$C(x,y) := \operatorname{cov}(Z(x), Z(y)).$$

The same C(x, y) is called the auto covariance function in two instances: in time series (to denote exactly the same concept except that x and y refer to locations in time rather than in space), and in multivariate random fields (to refer to the covariance of a variable with itself, as opposed to the cross covariance between two different variables at different locations, Cov(Z($x_1$), Y($x_2$)))

- ✓ **Mean & Variance of covariance functions:**

For locations $x_1, x_2, \ldots, x_N \in D$ the variance of every linear combination

$$X = \sum_{i=1}^{N} w_i Z(x_i)$$

can be computed as

$$\operatorname{var}(X) = \sum_{i=1}^{N} \sum_{j=1}^{N} w_i C(x_i, x_j) w_j.$$

A function is a valid covariance function if and only if this variance is non-negative for all possible choices of N and weights $w_1, \ldots, w_N$. A function with this property is called positive definite.

# RGODIC PROCESS:

In the event that the distributions and statistics are not available we can avail ourselves of the time averages from the particular sample function. The mean of the sample function Xλo(t) is referred to as the sample mean of the process X(t) and is defined as

$$\langle \mu(X)T \rangle = \left(\frac{1}{T}\right) \int_{-T/2}^{T/2} X\lambda o(t)dt$$

This quantity is actually a random-variable by itself because its value depends on the parameter sample function over it was calculated. the sample variance of the random process is defined as

$$\langle \sigma^2(X)T \rangle = \left(\frac{1}{T}\right) \int_{-T/2}^{T/2} \left| |X\lambda o(t)| - \langle \mu x \rangle T \right|^2 dt$$

The time-averaged sample ACF is obtained via the relation is

$$\langle RXX \rangle T = \left(\frac{1}{T}\right) \int_{-T/2}^{T/2} x(t) * x(t - T) dt$$

These quantities are in general not the same as the ensemble averages described before. A random process X(t) is said to be ergodic in the mean, i.e., first-order ergodic if the mean of sample average asymptotically approaches the ensemble mean

$$\lim_{T \to \infty} E\{\langle \mu X \rangle T\} = \mu x(t)$$

$$\lim_{T \to \infty} \mathrm{var}\{\langle \mu X \rangle T\} = 0$$

In a similar sense a random process X(t) is said to be ergodic in the ACF, i.e, second-order ergodic if

$$\lim_{T \to \infty} E\{\langle RXX(\tau) \rangle\} = RXX(\tau)$$

$$\lim_{T \to \infty} \mathrm{var}\{\langle RXX(\tau) \rangle\} = 0$$

The concept of ergodicity is also significant from a measurement perspective because in Practical situations we do not have access to all the sample realizations of a random process. We therefore have to be content in these situations with the time-averages that we obtain from a single realization. Ergodic processes are signals for which measurements based on a single sample function are sufficient to determine the ensemble statistics. Random signal for which this property does not hold are referred to as non-ergodic processes. As before the Gaussian random signal is an exception where strict sense ergodicity implies wide sense ergodicity.

## GUASSIAN PROCESSS:

A random process $X(t)$ is a Gaussian process if for all n and all ($t_1$ ,$t_2$ ,…,$t_n$ ), the random variables have a jointly Gaussian density function. For Gaussian processes, knowledge of the mean and autocorrelation; i.e., $m_X(t)$ and $R_x(t_1, t_2)$ gives a complete statistical description of the process. If the Gaussian process X(t) is passed through an LTI system, then the output process Y(t) will also be a Gaussian process. For Gaussian processes, WSS and strict stationary are equivalent.

A Gaussian process is a stochastic process $X_t$, $t \in T$, for which any finite linear combination of samples has a joint Gaussian distribution. More accurately, any linear functional applied to the sample function$X_t$ will give a normally distributed result. Notation-wise, one can write $X \sim GP(m,K)$, meaning the random function X is distributed as a GP with mean function m and covariance function K.[1] When the input vector t is two- or multi-dimensional a Gaussian process might be also known as a Gaussian random field.

A sufficient condition for the ergodicity of the stationary zero-mean Gaussian process X(t) is that

$$\int_{-\infty}^{\infty} RX(\tau)d\tau < \infty.$$

**Jointly Gaussian processes:**

The random processes X(t) and Y(t) are jointly Gaussian if for all n, m and all ($t_1$ ,$t_2$ ,…,$t_n$ ), and ($\tau_1$ , $\tau_2$ ,…, $\tau_m$ ), the random vector ($X(t_1)$,$X(t_2)$,…,$X(t_n)$, $Y(\tau_1)$,$Y(\tau_2)$,…, $Y(\tau_m)$) is distributed according to an n+M dimensional jointly Gaussian distribution.

For jointly Gaussian processes, uncorrelatedness and independence are equivalent.

# LINEAR FILTERING OF RANDOM PROCESSES:

• A random process X(t) is applied as input to a linear time-invariant filter of impulse response h(t),

• It produces a random process Y (t) at the filter output as

$$X(t) \rightarrow\rightarrow\rightarrow\rightarrow\rightarrow h(t) \rightarrow\rightarrow\rightarrow \quad Y(t)$$

• Difficult to describe the probability distribution of the output random process Y (t), even when the probability distribution of the input random process X(t) is completely specified for $-\infty \le t \le +\infty$.

• Estimate characteristics like mean and autocorrelation of the output and try to analyse its behaviour.

• Mean The input to the above system X(t) is assumed stationary. The mean of the output random process Y (t) can be calculated

$$\begin{aligned} mY(t) = E[Y(t)] &= E[\int_{-\infty}^{\infty} h(\tau)X(t-\tau)\,d\tau] \\ &= \int_{-\infty}^{\infty} h(\tau)E[X(t-\tau)]\,d\tau \\ &= mX \int_{-\infty}^{\infty} h(\tau)\,d\tau \\ &= mXH(0) \end{aligned}$$

where H(0) is the zero frequency response of the system.

- ✓ **Autocorrelation:**

The autocorrelation function of the output random process Y (t). By definition, we have

RY (t, u) = E[Y (t)Y (u)]

where t and u denote the time instants at which the process is observed. We may therefore use the convolution integral to write

$$\begin{aligned} RY(t,u) &= E\,[\int_{-\infty}^{\infty} h(\tau 1)X(t-\tau 1)\,d\tau 1] \;\; \int_{\infty}^{\infty} h(\tau 2)X(t-\tau 2)\,d\tau 2\,] \\ &= \int_{-\infty}^{\infty} h(\tau 1)\,d\tau 1 \int_{-\infty}^{\infty} h(\tau 2)E\,[X(t-\tau 1)X(t-\tau 2)]\,d\tau 2 \end{aligned}$$

When the input X(t) is a wide-stationary random process, autocorrelation function of X(t) is only a function of the difference between the observation times t − τ1 and u − τ2.

Putting τ = t − u, we get

$$RY(\tau) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\tau 1)h(\tau 2)RX(\tau - \tau 1 + \tau 2)\, d\tau 1\, d\tau 2$$

$$RY(0) = E[Y^2(t)]$$

The mean square value of the output random process Y (t) is obtained by putting $\tau = 0$ in the above equation.

$$E[Y^2(t)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\tau 1)h(\tau 2)RX(\tau 2 - \tau 1)\, d\tau 1\, d\tau 2$$

The mean square value of the output of a stable linear time-invariant filter in response to a wide-sense stationary random process is equal to the integral over all frequencies.

of the power spectral density of the input random process multiplied by the squared magnitude of the transfer function of the filter.

## APPLICATION AND ITS USES:

• A Gaussian process can be used as a prior probability distribution over functions in Bayesian inference.

• Wiener process (aka Brownian motion) is the integral of a white noise Gaussian process. It is not stationary, but it has stationary increments.