**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 $\rho_{X,Y}$ between two random variables X and Y with expected values $\mu_X$ and $\mu_Y$ and standard deviations $\sigma_X$ and $\sigma_Y$ is defined as:

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

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 = X2.

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.

### Introduction

The power spectrum of a time series x(t) describes how the variance of the data x(t) is distributed over the frequency components into which x(t) may be decomposed. This distribution of the variance may be described either by a measure μ or by a statistical cumulative distribution function S(f) = the power contributed by frequencies from 0 upto f. Given a band of frequencies [a, b) the amount of variance contributed to x(t) by frequencies lying within the interval [a,b) is given by S(b) - S(a). Then S is called the spectral distribution function of x.

The spectral density at a frequency f gives the rate of variance contributed by frequencies in the immediate neighbourhood of f to the variance of x per unit frequency.

## 1 Auto Correlation of a Random Process



Let X(t1) and X(t2) be the two given random variables. Then auto correlation is

RXX (t1, t2) = E[X(t1) X(t2)] Mean Square Value

Putting t1 = t2 = t in (1)

RXX (t, t) = E[X(t) X(t)]

⇒ RXX (t, t) = E[X2(t)]

Which is called the mean square value of the random process.

Auto Correlation Function

Definition: Auto Correlation Function of the random process {X(t)} is

RXX = (τ) = E{(t) X(t+τ)}

Note: RXX (τ) = R(τ) = RX (τ)

## PROPERTY: 1

The mean square value of the Random process may be obtained from the auto correlation function. RXX(τ), by putting τ = 0.

is known as Average power of the random process {X(t)}.

## PROPERTY: 2

RXX(τ) is an even function of τ.

RXX (τ) = RXX (-τ)

## PROPERTY: 3

If the process X(t) contains a periodic component of the same period.

## PROPERTY: 4

If a random process {X(t)} has no periodic components, and E[X(t)] = X then

$$\lim_{|\tau|\to\infty} R_{XX}(\tau) = \bar{X}^2 \quad (\text{or}) \bar{X} = \sqrt{\lim_{|\tau|\to\infty} R_{XX}(\tau)}$$

i.e., when τ→∞, the auto correlation function represents the square of the mean of the random process.

## PROPERTY: 5

The auto correlation function of a random process cannot have an arbitrary shape.

## Correlation Coefficient

$$\rho_{XX}(t_1, t_2) = \frac{C_{XX}(t_1, t_2)}{\sqrt{\text{Var}\, X(t_1) \times \text{Var}\, X(t_2)}}$$

Where $C_{XX}(t_1, t_2)$ denotes the auto covariance.

## Cross Correlation

Cross correlation between the two random process {X(t)} and {Y(t)} is defined as RXY (t1, t2) = E[X(t1) Y(t2)] where X(t1) Y(t2) are random variables.

## Cross Covariance

Let {X(t)} and {Y(t)} be any two random process. Then the cross covariance is defined as

$$C_{XY}(t_1,t_2) = E\left\{\left[X(t_1)-E(Y(t_1))\right]\left[X(t_2-E[Y(t_2)])\right]\right\}$$

The relation between Mean Cross Correlation and cross covariance is as follows:

$$C_{XY}(t_1,t_2) = R_{XY}(t_1,t_2)-E\left[X(t_1)E\left[Y(t_2)\right]\right]$$

**Definition**

Two random process {X(t)} and {Y(t)} are said to be uncorrelated if

$C_{XY}(t_1,t_2)$ 0, $\forall\ t_1,t_2$

Hence from the above remark we have,

$R_{XY}(t_1, t_2) = E[X(t_1)\ Y(t_2)]$

## Cross Correlation Coefficient

$$\rho_{XY}(t_1,t_2) = \frac{c_{XY}(t_1,t_2)}{\sqrt{Var(X(t_1))Var(X(t_2))}}$$

## Cross Correlation And Its Properties

Let {X(t)} and {Y(t)} be two random. Then the cross correlation between them is also defined as

$$R_{XY}(t, t+\tau) = E\left[X(t)Y(t+\tau)\right] = R_{XY}(\tau)$$

**Property : 1**

$$RXY(\tau) = RYX(-\tau)$$

**Property : 2**

If {X(t)} and {Y(t)} are two random process then $\left|R_{XY}(\tau)\right| \le \sqrt{R_{XX}(0)R_{YY}(0)}$, where $R_{XX}(\tau)$ and $R_{YY}(\tau)$ are their respective auto correlation functions.

**Property : 3**

If {X(t)} and {Y(t)} are two random process then,

$$|R_{XY}(\tau)| \leq \tfrac{1}{2}[R_{XX}(0) + R_{YY}(0)]$$

# 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(x1), Y(x2)))

## ✓ Mean & Variance of covariance functions:

For locations x1, x2, …, xN ∈ 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 w1, …, wN. A function with this property is called positive definite.

## 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(x1), Y(x2)))

### ✓ Mean & Variance of covariance functions:

For locations x1, x2, …, xN ∈ 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 w1, …, wN. A function with this property is called positive definite.

Covariance signifies the direction of the linear relationship between the two variables. By direction we mean if the *variables* are directly proportional or inversely proportional to each other. (Increasing the value of one variable might have a positive or a negative impact on the value of the other variable).

The values of covariance can be any number between the two opposite infinities. Also, it's important to mention that covariance only measures how two variables change together, not the dependency of one variable on another one.

The value of covariance between 2 variables is achieved by taking the summation of the product of the differences from the means of the variables as follows:

$$Cov(x,y) = \frac{\sum (x_i - \bar{x}) * (y_i - \bar{y})}{N}$$

The upper and lower limits for the covariance depend on the variances of the variables involved. These variances, in turn, can vary with the scaling of the variables. Even a change in the units of measurement can change the covariance. Thus, covariance is only useful to find the direction of the relationship between two variables and not the magnitude. Below are the plots which help us understand how the covariance between two variables would look in different directions.

## Differences between Covariance and Correlation

Both the Covariance and Correlation metric evaluate two variables throughout the entire domain and not on a single value. The differences between them are summarized in a tabular form for quick reference. Let us look at Covariance vs Correlation.

| Covariance | Correlation |
|---|---|
| Covariance is a measure to indicate the extent to which two random variables change in tandem. | Correlation is a measure used to represent how strongly two random variables are related to each other. |
| Covariance is nothing but a measure of correlation. | Correlation refers to the scaled form of covariance. |
| Covariance indicates the direction of the linear relationship between variables. | Correlation on the other hand measures both the strength and direction of the linear relationship between two variables. |
| Covariance can vary between -∞ and +∞ | Correlation ranges between -1 and +1 |
| Covariance is affected by the change in scale. If all the values of one variable are multiplied by a constant and all the values of another variable are multiplied, by a similar or different constant, then the covariance is changed. | Correlation is not influenced by the change in scale. |
| Covariance assumes the units from the product of the units of the two variables. | Correlation is dimensionless, i.e. It's a unit-free measure of the relationship between variables. |

| | |
|---|---|
| Covariance of two dependent variables measures how much in real quantity (i.e. cm, kg, liters) on average they co-vary. | Correlation of two dependent variables measures the proportion of how much on average these variables vary w.r.t one another. |
| Covariance is zero in case of independent variables (if one variable moves and the other doesn't) because then the variables do not necessarily move together. | Independent movements do not contribute to the total correlation. Therefore, completely independent variables have a zero correlation. |

## ERGODIC 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

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

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

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

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

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

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 (t1 ,t2 ,…,tn ), the random variables have a jointly Gaussian density function. For Gaussian processes, knowledge of the mean and autocorrelation; i.e., mX (t) and Rx (t1 ,t2 ) 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 Xt, t ∈ T, for which any finite linear combination of samples has a joint Gaussian distribution. More accurately, any linear functional applied to the sample functionXt will give a normally distributed result. Notation-wise, one can write X ~ 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-dimensionala 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 (t1 ,t2 ,…,tn ), and (τ1 , τ2 ,…, τm ), the random vector (X(t1 ),X(t2 ),…,X(tn ), Y(τ1 ),Y( τ2 ),…, Y(τm )) 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

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

• 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 $-\infty \leq t \leq +\infty$.

- 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

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

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

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

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 $t - \tau 1$ and $u - \tau 2$.

Putting $\tau = t - u$, 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.

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.

## Mean or Mathematical Expectations

**Definition :** Let 'X' be a continuous random variable with probability density function f(x). Then the mathematical expectation of 'X' is denoted by E(X) and is given by

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

It is denoted by

$$\mu_r' = \int_{-\infty}^{\infty} x^r f(x) dx$$

Thus

$$\mu_1' = E(X) \qquad (\mu_1' \text{ about origin})$$

$$\mu_2' = E(X^2) \qquad (\mu_2' \text{ about origin})$$

$$\therefore \text{Mean} = \bar{X} = \mu_1' = E(X)$$

And

$$\text{Variance} = \mu_2' - \mu_2'^2$$

$$\text{Variance} = E(X^2) - [E(X)]^2 \qquad \text{(a)}$$

* $r^{th}$ moment (abut mean)

Now

$$E\{X - E(X)\}^r = \int_{-\infty}^{\infty} \{x - E(X)\}^r f(x) dx$$

$$= \int_{-\infty}^{\infty} \{x - \bar{X}\}^r f(x) dx$$

Thus

$$\mu_r = \int_{-\infty}^{\infty} \{x - \bar{X}\}^r f(x) dx \qquad \text{(b)}$$

Where $\mu_r = E[X - E(X)^r]$

This gives the $r^{th}$ moment about mean and it is denoted by $\mu_r$

Put r = 1 in (B) we get

$$\mu_r = \int_{-\infty}^{\infty} \{x - \overline{X}\} f(x)\,dx$$

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

$$= \overline{X} - \overline{X} \int_{-\infty}^{\infty} f(x)\,dx \qquad \left[\because \int_{-\infty}^{\infty} f(x)\,dx = 1\right]$$

$$= \overline{X} - \overline{X}$$

$$\mu_1 = 0$$

Put r = 2 in (B), we get

$$\mu_2 = \int_{-\infty}^{\infty} (x - \overline{X})^2 f(x)\,dx$$

$$\text{Variance} = \mu_2 = E[X - E(X)]^2$$

Which gives the variance interms of expectations.

**Note**

Let g(x) = K (Constant), then

$$E[g(X)] = E(K) = \int_{-\infty}^{\infty} K f(x)\,dx$$

$$= K \int_{-\infty}^{\infty} f(x)\,dx \qquad \left[\because \int_{-\infty}^{\infty} f(x)\,dx = 1\right]$$

$$= K \cdot 1 = K$$

Thus $E(K) = K \Rightarrow$ E[a constant] = constant.

## 3.4 EXPECTATIONS (Discrete R.V.'s)

Let 'X' be a discrete random variable with P.M.F p(x)

Then

$$E(X) = \sum_x x\, p(x)$$

For discrete random variables 'X'

$$E(X^r) = \sum_x x^r\, p(x) \qquad \text{(by def)}$$

If we denote

$$E(X^r) = \mu_r'$$

Then

$$\mu_r' = E[X^r] = \sum_x x^r p(x)$$

Put r = 1, we get

$$\text{Mean } \mu_r' = \sum x\, p(x)$$

Put r = 2, we get

$$\mu_2' = E[X^2] = \sum_x x^2 p(x)$$

$$\therefore \mu_2 = \mu_2' - \mu_1'^2 = E(X^2) - \{E(X)\}^2$$

The $r^{th}$ moment about mean

$$\mu_r' = E[\{X - E(X)\}^r]$$

$$= \sum_x (x - \bar{X})^r p(x), \qquad E(X) = \bar{X}$$

Put r = 2, we get

$$\text{Variance} = \mu_2 = \sum_x (x - \bar{X})^2 p(x)$$

## 3.5 ADDITION THEOREM (EXPECTATION)

**Theorem 1**

If X and Y are two continuous random variable with pdf fx(x) and fy(y) then

E(X+Y) = E(X) + E(Y)

## Multiplication Theorem Of Expectation

## Theorem 2

If X and Y are independent random variables,

Then E(XY) = E(X) . E(Y)

## Note :

If X1, X2, ……, Xn are 'n' independent random variables, then

E[X1, X2, ……, Xn] = E(X1), E(X2), ……, E(Xn)

## Theorem 3

If 'X' is a random variable with pdf f(x) and 'a' is a constant, then

(i) E[a G(x)] = a E[G(x)]

(ii) E[G(x)+a] = E[G(x)+a]

Where G(X) is a function of 'X' which is also a random variable.



## Theorem 4

If 'X' is a random variable with p.d.f. f(x) and 'a' and 'b' are constants, then E[ax + b] = a E(X) + b

## Correlation :

If we take a = 1 and b = –E(X) = – X , then we get

E(X- X ) = E(X) – E(X) = 0

## Note

$$E\left(\frac{1}{X}\right) \neq \frac{1}{E(X)}$$

$$E[\log(x)] \neq \log E(X)$$

$$E(X^2) \neq [E(X)]^2$$

## 3.7 Expectation Of A Linear Combination Of Random Variables

Let X1, X2, ……, Xn be any 'n' random variable and if a1, a2 , ……, an are constants, then E[a1X1 + a2X2 + ……+ anXn] = a1E(X1) + a2E(X2)+ ……+ anE(Xn)

### Result

If X is a random variable, then

Var (aX + b) = a2Var(X) 'a' and 'b' are constants.

### Covariance :

If X and Y are random variables, then covariance between them is defined as Cov(X, Y) = E{[X - E(X)] [Y - E(Y)]}

$$Cov(X, Y) = E(XY) - E(X) \cdot E(Y) \quad (A)$$

If X and Y are independent, then

$$E(XY) = E(X) E(Y) \quad (B)$$

Sub (B) in (A), we get Cov (X, Y) = 0



∴ If X and Y are independent, then

Cov (X, Y) = 0

Note

(i) Cov(aX, bY) = ab Cov(X, Y)

(ii) Cov(X+a, Y+b) = Cov(X, Y)

(iii) Cov(aX+b, cY+d) = ac Cov(X, Y)

(iv) Var (X1 + X2) = Var(X1) + Var(X2) + 2 Cov(X1, X2)

If X1, X2 are independent

Var (X1+ X2) = Var(X1) + Var(X2)

## EXPECTATION TABLE

| Discrete R.V's | Continuous R.V's |
|---|---|
| 1. $E(X) = \sum x\, p(x)$ | 1. $E(X) = \int_{-\infty}^{\infty} x f(x)\,dx$ |
| 2. $E(X^r) = \mu_r' = \sum_x x^r p(x)$ | 2. $E(X^r) = \mu_r' = \int_{-\infty}^{\infty} x^r f(x)\,dx$ |
| 3. Mean $= \mu_r' = \sum x\, p(x)$ | 3. Mean $= \mu_r' = \int_{-\infty}^{\infty} x\, f(x)\,dx$ |
| 4. $\mu_2' = \sum x^2 p(x)$ | 4. $\mu_2' = \int_{-\infty}^{\infty} x^2 f(x)\,dx$ |
| 5. Variance $= \mu_2' - \mu_1'^2 = E(X^2) - \{E(X)\}^2$ | 5. Variance $= \mu_2' - \mu_1'^2 = E(X^2) - \{E(X)\}^2$ |

# Power Spectral Density

## Introduction

(i) Fourier Transformation

(ii) Inverse Fourier Transform

(iii) Properties of Auto Correlation Function

(iv) Basic Trigonometric Formula

(v) Basic Integration

## 5.1 SPECIAL REPRESENTATION

Let x(t) be a deterministic signal. The Fourier transform of x(t) is defined as

$$F[x(t)] = x(w) = \int_{-\infty}^{\infty} x(t) e^{-i\omega t} dt$$

Here X(ω) is called "spectrum of x(t)".

Hence x(t) = Inverse Fourier Transform of X(ω)

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\omega) e^{i\omega t} d\omega .$$

### Definition

The average power P(T) of x(t) over the interval (-T, T) is given by

$$P_{XX} = \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} E[X^2(t)] dt$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \lim_{T\to\infty} \frac{E\left[|X_T(\omega)|^2\right]}{2T} d\omega \qquad (2)$$

### Definition

The average power PXX for the random process {X(t)} is given by

$$P_{XX} = \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} E\left[X^2(t)\right] dt$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \lim_{T\to\infty} \frac{E\left[\left|X_T(\omega)\right|^2\right]}{2T} d\omega \qquad (2)$$

# 6 POWER SPECTRAL DENSITY FUNCTION

**Definition**

If {X(t)} is a stationary process (either in the strict sense or wide sense) with auto correlation function RXX(τ), then the Fourier transform of RXX(τ) is called the power spectral density function of {X(t)} and is denoted by SXX(ω) or S(ω) or SX(ω).

SXX(ω)= Fourier Transform of RXX (τ)

$$S_{XX}(\omega) = \int_{-\infty}^{\infty} R_{XX}(\tau) e^{-i\omega\tau} d\tau$$

$$S_{XX}(f) = \int_{-\infty}^{\infty} R_{XX}(\tau) e^{-i2\pi f\tau} d\tau$$

To find $R_{XX}(\tau)$ if $S_{XX}(\omega)$ or $S_{XX}(f)$ is given

$$R_{XX}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{XX}(\omega) e^{i\omega\tau} d\omega \quad \text{[inverse Fourier transform of } S_{XX}(\omega)\text{]}$$

$$\text{(or) } R_{XX}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{XX}(f) e^{-i2\pi f\tau} d\tau$$

[inverse Fourier transform of $S_{XX}(f)$]

## 6.1 WIENER KHINCHINE RELATION

$$S_{XX}(\omega) = \int_{-\infty}^{\infty} R_{XX}(\tau) e^{-i\omega\tau} d\tau$$

$$S_{XX}(f) = \int_{-\infty}^{\infty} R_{XX}(\tau) e^{-i2\pi f\tau} d\tau$$

To find $R_{XX}(\tau)$ if $S_{XX}(\omega)$ or $S_{XX}(f)$ is given

$$R_{XX}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{XX}(\omega) e^{i\omega\tau} d\omega$$ [inverse Fourier transform of $S_{XX}(\omega)$]

$$\text{(or) } R_{XX}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{XX}(f) e^{-i2\pi f\tau} d\tau$$

[inverse Fourier transform of $S_{XX}(f)$]

## PROPERTIES OF POWER SPECTRAL DENSITY FUNCTION



**Property 1:**

The value of the spectral density function at zero frequency is equal to the total area under the group of the auto correlation function.

$$S_{XX}(f) = \int_{-\infty}^{\infty} R_{XX}(\tau) e^{-i2\pi fc} d\tau$$

Taking $f = 0$, we get

$$Sxx(0) = \int_{-\infty}^{\infty} R_{XX}(\tau) d\tau$$

**Example 1.** Check whether the following function are valid auto correlation function (i) 5 sin nπ (ii) $\frac{1}{1+9\tau^2}$

**Solution:**

(i) Given $R_{XX}(\tau) = 5 \text{ Sin } n\pi$

$R_{XX}(-\tau) = 5 \text{ Sin } n(-\pi) = -5 \text{ Sin } n\pi$

$R_{XX}(\tau) \neq R_{XX}(-\tau)$, the given function is not an auto correlation function.

(ii) Given $R_{XX}(\tau) = \frac{1}{1+9\tau^2}$

$$R_{XX}(-\tau) = \frac{1}{1+9(-\tau)^2} = R_{XX}(\tau)$$

$\therefore$ The given function is an auto correlation function.

**Example : 2**

Find the mean and variance of a stationary random process whose auto correlation function is given by

$$R_{XX}(\tau) = 18 + \frac{2}{6+\tau^2}$$

**Solution**

Given $R_{XX}(\tau) = 18 + \frac{2}{6+\tau^2}$

$$\bar{X}^2 = \lim_{|\tau|\to\infty} R_{XX}(\tau)$$

$$= \lim_{|\tau|\to\infty}\left[18 + \frac{2}{6+\tau^2}\right]$$

$$= 18 + \lim_{|\tau|\to\infty} \frac{2}{6+\tau^2}$$

$$= 18 + \frac{2}{6+}$$

$$= 18 + 0$$

$$= 18$$

$$\bar{X} = \sqrt{18}$$

$$E[X(t)] = \sqrt{18}$$

$$\text{Var } \{X(t)\} = E[X^2(t)] - \{E[X(t)]\}^2$$

We know that

$$E[X^2(t)] = R_{XX}(0)$$

$$= 18 + \frac{2}{6+0} = \frac{55}{3}$$

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

**Example : 3**

Express the autocorrelation function of the process {X'(t)} in terms of the auto correlation function of process {X(t)}

Solution

**Solution**

Consider, $R_{XX}'(t_1, t_2) = E\{X(t_1)X'(t_2)\}$

$$= E\left[X(t_1)\lim_{n\to 0}\left\{\frac{X(t_2+h)-X(t_2)}{h}\right\}\right]$$

$$= \lim_{h\to 0} E\left\{\frac{X(t_1)X(t_2+h)-X(t_1)X(t_2)}{h}\right\}$$

$$= \lim_{h\to 0}\left\{\frac{R_{XX}(t_1,t_2+h)-R_X(t_1,t_2)}{h}\right\}$$

$$\Rightarrow R_{XX}'(t_1, t_2) = \frac{\partial}{\partial t_2}R_{XX}(t_1,t_2) \qquad (1)$$

Similarly $R_{XX}'(t_1, t_2) = \frac{\partial}{\partial t_1}R_{XX}'(t,t_2)$

$$\Rightarrow R_X'_X(t_1, t_2) = \frac{\partial}{\partial t,\partial t_2}R_{XX}(t_1,t_2) \qquad \text{by (1)}$$

**Example :4**

Two random process {X(t)} and {Y(t)} are given by $X(t) = A\cos(\omega t+\theta)$, $Y(t) = A\sin(\omega t + \theta)$ where A and ω are constants and 'θ' is a uniform random variable over 0 to 2π. Find the cross correlation function.

**Solution**

By def. we have

**Solution**

By def. we have

$$R_{XY}(\tau) = \underline{R_{XY}}\ (t, t+\tau)$$

Now, $R_{XY}(t, t+\tau) = E[X(t).\ Y(t+\tau)]$

$= E\ [A \cos(\omega t + \theta).\ A \sin(\omega (t+\tau) + \theta)]$

$$= A^2 E\left[\sin\{\omega(t+\tau)+\theta\}\cos(\omega t+\theta)\right]$$

Since '$\theta$' is a uniformly distributed random variable we have

$$f(0) = \frac{1}{2\pi}, \quad 0 \le \theta \le 2\pi$$

Now $E\left[\sin\{\omega(t+\tau)+\theta\}\cos(\omega t+\theta)\right]$

$$= \int_{-\infty}^{\infty} \sin(\omega t+\omega\tau+\theta).\cos(wt+\theta) f(\theta)\,d\theta$$

$$= \int_{0}^{2\pi} \sin(\omega t+\omega^{t+\theta}).\cos(\omega t+\theta)\left(\frac{1}{2\pi}\right)d\theta$$

$$= \frac{1}{2\pi}\int_{0}^{2\pi} \sin(\omega t+\omega\tau+\theta)\cos(\omega t+\theta)\,d\theta$$

$$= \frac{1}{2\pi}\int_{0}^{2\pi} \frac{1}{2}\{\sin(\omega t+\omega\tau+\theta+\omega t+\theta)$$

$$+\sin\left[\omega t+\omega\tau+\theta-\omega t-\theta\right]\}d\theta$$

$$-\frac{1}{2\pi}\int_0^{2\pi}\frac{\sin\left[2\omega t+\omega\tau+2\theta\right]+\sin\left(\omega\tau\right)}{2}d\theta$$

$$=\frac{1}{4\pi}\left(-\frac{\cos\left(2\omega t+\omega\tau+2\theta\right)}{2}+\sin\omega\tau\left(\theta\right)\right)_0^{2\pi}$$

$$=\frac{1}{4\pi}\left[-\frac{\cos\left(2\omega t+\omega\tau\right)}{2}+\frac{\cos\left(2\omega t+\omega\tau+0\right)}{2}+\sin\omega\tau\left(2\pi-0\right)\right]$$

$$=\frac{1}{4\pi}\left[-\frac{\cos\left(2\omega t+\omega\tau\right)}{2}+\frac{\cos\left(2\omega t+\omega\tau\right)}{2}+2\pi\sin\omega\tau\right]$$

$$=\frac{1}{4\pi}\left[0+2\pi\sin\omega\tau\right]$$

$$=\frac{1}{2}\sin\omega\tau \qquad (3)$$

Substituting (3) in (1) we get

$$R_{XY}\left(t,t_\tau\right)=\frac{A^2}{2}\sin\omega\tau$$

# Random Processes

## Introduction

Random variable is a function of the possible outcomes of a experiment. But, it does not include the concept of time. In the real situations, we come across so many time varying functions which are random in nature. In electrical and electronics engineering, we studied about signals.

Generally, signals are classified into two types.

(i) Deterministic

(ii) Random

Here both deterministic and random signals are functions of time. Hence it is possible for us to determine the value of a signal at any given time. But this is not possible in the case of a random signal, since uncertainty of some element is always associated with it. The probability model used for characterizing a random signal is called a random process or stochastic process.



## Random Process Concept

A random process is a collection (ensemble) of real variable {X(s, t)} that are functions of a real variable t where $s \in S$, S is the sample space and $t \in T$. (T is an index set).

Remark

i) If t is fixed, then {X(s, t)} is a random variable.

ii) If S and t are fixed {X(s, t)} is a number.

iii) If S is fixed, {X(s, t)} is a signal time function.

## Notation

Here after we denote the random process {X(s, t)} by {X(t)} where the index set T is assumed to be continuous process is denoted by {X(n)} or {Xn}.

A comparison between random variable and random process

| Random Variable | Random Process |
|---|---|
| A function of the possible outcomes of an experiment is X(s) | A function of the possible outcomes of an experiment and also time i.e, X(s, t) |
| Outcome is mapped into a number x. | Outcomes are mapped into wave from which is a fun of time 't'. |

## Random Variable

A function of the possible outcomes of an experiment is X(s)

Outcome is mapped into a number x.

## Random Process

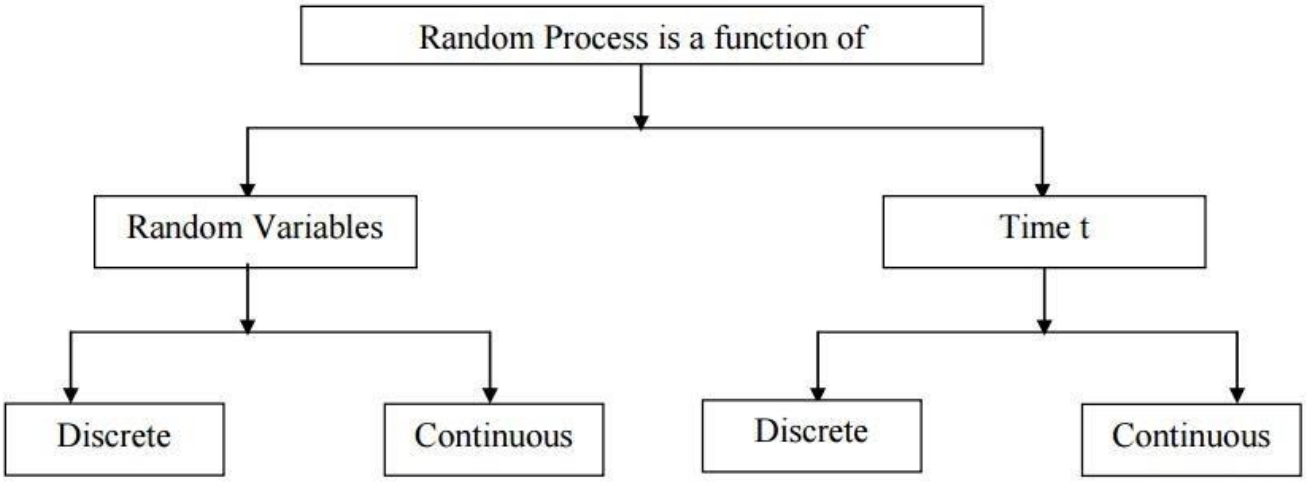
A function of the possible outcomes of an experiment and also time i.e, X(s, t)

Outcomes are mapped into wave from which is a fun of time 't'.

## Classification Of Random Processes



We can classify the random process according to the characteristics of time t and the random variable X = X(t) t & x have values in the ranges $-\infty < t < \infty$ and $-\infty < x < \infty$.



## 1 Continuous Random Process

If 'S' is continuous and t takes any value, then X(t) is a continuous random variable.

### Example

Let X(t) = Maximum temperature of a particular place in (0, t). Here 'S' is a continuous set and $t \geq 0$ (takes all values), {X(t)} is a continuous random process.

## 2 Discrete Random Process

If 'S' assumes only discrete values and t is continuous then we call such random process {X(t) as Discrete Random Process.

### Example

Let X(t) be the number of telephone calls received in the interval (0, t). Here, S = {1, 2, 3, …}

$$T = \{t, t \geq 0\}$$

$\therefore$ {X(t)} is a discrete random process.

## 3 Continuous Random Sequence

If 'S' is a continuous but time 't' takes only discrete is called discrete random sequence. Example: Let Xn denote the outcome of the nth toss of a fair die.

Here S = {1, 2, 3, 4, 5, 6} T = {1, 2, 3, …}

$\therefore$ (Xn, n = 1, 2, 3, …} is a discrete random sequence.



## Classification Of Random Processes Based On Its Sample Functions

### Non-Deterministic Process

A Process is called non-deterministic process if the future values of any sample function cannot be predicted exactly from observed values.

### Deterministic Process

A process is called deterministic if future value of any sample function can be predicted from past values.

**UNIT III**

**RANDOM PROCESS**

## 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 = xi defined to be P(X = xi) = pi. The probabilities pi 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.

## Distribution function of the random variable X or cumulative distribution of the random variable X

### Definition :

The distribution function of a random variable X defined in $(-\infty, \infty)$ is given by $F(x) = P(X \leq x) = P\{s : X(s) \leq x\}$

**Note**



Let the random variable X takes values x1, x2, ….., xn with probabilities P1, P2, ….., Pn and let x1< x2< ….. <xn

Then we have

$F(x) = P(X < x1) = 0, -\infty < x < x,$

$F(x) = P(X < x1) = 0, P(X < x1) + P(X = x1) = 0 + p1 = p1$

$F(x) = P(X < x2) = 0, P(X < x1) + P(X = x1) + P(X = x2) = p1 + p2$

$F(x) = P(X < xn) = P(X < x1) + P(X = x1) + ….. + P(X = xn)$

$= p1 + p2 + ……. + pn = 1$

## 2.2 Properties Of Distribution Functions

Property : 1 $P(a < X \leq b) = F(b) – F(a)$, where $F(x) = P(X \leq x)$

Property : 2 $P(a \leq X \leq b) = P(X = a) + F(b) – F(a)$

Property : 3 $P(a < X < b) = P(a < X \leq b) - P(X = b)$

$= F(b) - F(a) - P(X = b)$ by prob (1)

## 2.3 Probability Mass Function (Or) Probability Function

Let X be a one dimenstional discrete R.V. which takes the values x1, x2, …… To each possible outcome 'xi' we can associate a number pi.

i.e., $P(X = x_i) = P(x_i) = p_i$ called the probability of xi. The number $p_i = P(x_i)$ satisfies the following conditions.

(i) $p(x_i) \geq 0, \forall_i$ (ii) $\sum_{i=1}^{\infty} p(x_i) = 1$

The function p(x) satisfying the above two conditions is called the probability mass function (or) probability distribution of the R.V.X. The probability distribution {xi, pi} can be displayed in the form of table as shown below.



| $X = x_i$ | $x_1$ | $x_2$ | …… | $x_i$ |
|---|---|---|---|---|
| $P(X = x_i) = p_i$ | $p_1$ | $p_2$ | …… | $p_i$ |

## Notation

Let 'S' be a sample space. The set of all outcomes 'S' in S such that X(S) = x is denoted by writing X = x.

$P(X = x) = P\{S : X(s) = x\}$

|||ly $P(x \leq a) = P\{S : X() \in (-\infty, a)\}$

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

$P(X = a \text{ or } X = b) = P\{(X = a) \cup (X = b)\}$

$P(X = a \text{ and } X = b) = P\{(X = a) \cap (X = b)\}$ and so on.

Theorem :1 If X1 and X2 are random variable and K is a constant then KX1, X1 + X2, X1X2, K1X1 + K2X2, X1-X2 are also random variables.

**Theorem :2**

If 'X' is a random variable and f(•) is a continuous function, then f(X) is a random variable.

**Note**

If F(x) is the distribution function of one dimensional random variable then

I. $0 \le F(x) \le 1$

II. If $x < y$, then $F(x) \le F(y)$

III. $F(-\infty) = \lim_{x \to -\infty} F(x) = 0$

IV. $F(\infty) = \lim_{x \to \infty} F(x) = 1$

V. If 'X' is a discrete R.V. taking values $x_1, x_2, x_3$

Where $x_1 < x_2 < x_{i-1}\ x_i$ ..... then

$P(X = x_i) = F(x_i) - F(x_{i-1})$

## Continuous Random Variable



Definition : A R.V.'X' which takes all possible values in a given internal is called a continuous random variable.

Example : Age, height, weight are continuous R.V.'s.

## 3.1 Probability Density Function

Consider a continuous R.V. 'X' specified on a certain interval (a, b) (which can also be a infinite interval $(-\infty, \infty)$).

If there is a function y = f(x) such that

$$\lim_{\Delta x \to 0} \frac{P(x < X < x + \Delta x)}{\Delta x} = f(x)$$

Then this function f(x) is termed as the probability density function (or) simply density function of the R.V. 'X'.

It is also called the frequency function, distribution density or the probability density function.

The curve y = f(x) is called the probability curve of the distribution curve.

Remark

If f(x) is p.d.f of the R.V.X then the probability that a value of the R.V. X will fall in some interval (a, b) is equal to the definite integral of the function f(x) a to b.

$$P(a < x < b) = \int_a^b f(x)dx \quad \text{(or)}$$

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

### 3.2 Properties OF P.D.F

The p.d.f f(x) of a R.V.X has the following properties

1. In the case of discrete R.V. the probability at a point say at x = c is not zero. But in the case of a continuous R.V.X the probability at a point is always zero.

$$P(X = c) = \int_{-\infty}^{\infty} f(x)dx = [x]_c^C = C - C = 0$$

2. If x is a continuous R.V. then we have $p(a \le X \le b) = p(a \le X < b)$

$= p(a < X \text{ V } b)$

### IMPORTANT DEFINITIONS INTERMS OF P.D.F

If f(x) is the p.d.f of a random variable 'X' which is defined in the interval (a, b) then

| | | |
|---|---|---|
| i | Arithmetic mean | $\int_a^b x\, f(x)dx$ |
| ii | Harmonic mean | $\int_a^b \frac{1}{x} f(x)dx$ |
| iii | Geometric mean 'G' log G | $\int_a^b \log x\, f(x)dx$ |
| iv | Moments about origin | $\int_a^b x^r\, f(x)dx$ |
| v | Moments about any point A | $\int_a^b (x-A)^r\, f(x)dx$ |
| vi | Moment about mean $\mu_r$ | $\int_a^b (x-\text{mean})^r\, f(x)dx$ |
| vii | Variance $\mu_2$ | $\int_a^b (x-\text{mean})^2\, f(x)dx$ |
| viii | Mean deviation about the mean is M.D. | $\int_a^b \lvert x-\text{mean}\rvert\, f(x)dx$ |

## Continuous Distribution Function

Definition :

If f(x) is a p.d.f. of a continuous random variable 'X', then the function

$$F_X(x) = F(x) = P(X \le x) = \int_{-\infty}^{\infty} f(x)dx, \quad -\infty < x < \infty$$

is called the distribution function or cumulative distribution function of the random variable.

## Properties Of CDF Of A R.V. 'X'

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

(ii) $\underset{x\to-\infty}{Lt}\, F(x) = 0, \qquad \underset{x\to-\infty}{Lt}\, F(x) = 1$

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

(iv) $F'(x) = \frac{dF(x)}{dx} = f(x) \ge 0$

(v) $P(X = x_i) = F(x_i) - F(x_i - 1)$

## Moment Generating Function

**Definition :** The moment generating function (MGF) of a random variable 'X' (about origin) whose probability function f(x) is given by

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

$$= \begin{cases} \int_{x=-\infty}^{\infty} e^{tx} f(x)dx, \text{ for a continuous probably function} \\ \sum_{x=-\infty}^{\infty} e^{tx} p(x), \text{ for a discrete probably function} \end{cases}$$

Where t is real parameter and the integration or summation being extended to the entire range of x.

## Discrete Distributions

The important discrete distribution of a random variable 'X' are

1. Binomial Distribution
2. Poisson Distribution
3. Geometric Distribution



## 6.1 Binomial Distribution

**Definition:**

A random variable X is said to follow binomial distribution if its probability law is given by

P(x) = p(X = x successes) = nCx px qn-x Where x = 0, 1, 2, ……., n, p+q = 1

**Note**

Assumptions in Binomial distribution

i) There are only two possible outcomes for each trail (success or failure).

ii) The probability of a success is the same for each trail.

iii) There are 'n' trails, where 'n' is a constant.

iv) The 'n' trails are independent.

## Passion Distribution

### Definition :

A random variable X is said to follow if its probability law is given by

$$P(X = x) = p(x) = \frac{e^{-\lambda}\lambda^{x}}{x!}, \quad x = 0, 1, 2, \ldots, \infty$$

Poisson distribution is a limiting case of binomial distribution under the following conditions or assumptions.

1. The number of trails 'n' should e infinitely large i.e. $n \to \infty$.
2. The probability of successes 'p' for each trail is infinitely small.
3. $np = \lambda$ , should be finite where $\lambda$ is a constant.

### 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.

## 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.

## 1 STATIONARY PROCESS

A random process is said to be stationary if its mean, variance, moments etc are constant. Other processes are called non stationary.

### 1. 1st Order Distribution Function of {X(t)}

For a specific t, X(t) is a random variable as it was observed earlier.

$F(x, t) = P\{X(t) \le x\}$ is called the first order distribution of the process $\{X(t)\}$.

### 1st Order Density Function of {X(t)}

$$f(x,t) = \frac{\partial}{\partial x} F(x,t)$$ is called the first order density of {X, t}

### 2nd Order distribution function of {X(t)}

$F(x_1, x_2; t_1, t_2) = P\{X(t_1) \le x_1;\ X(t_2) \le x_2\}$ is the point distribution of the random variables X(t1) and X(t2) and is called the second order distribution of the process {X(t)}.

### 2nd order density function of {X(T)}

$$f(x_1, x_2; t_1, t_2) = \frac{\partial^2 F(x_1, x_2; t_1, t_2)}{\partial x, \partial x_2}$$ is called the second order density of {X(t)}

## 2 First - Order Stationary Process

### Definition

A random process is called stationary to order, one or first order stationary if its 1st order density function does not change with a shift in time origin.

In other words,

$f_X(x_1, t_1) = f_X(x_1, t_1 + C)$ must be true for any t1 and any real number C if {X(t1)} is to be a first order stationary process.

Let as show that a first order stationary process has a constant mean.

Consider a random process {X(t1)} at two different times t1 and t2.

$$\therefore E[X(t_1)] = \int_{-\infty}^{\infty} x f(x, t_1) dx$$

[f(x,t1) is the density form of the random process X(t1)]

$$\therefore E[X(t_2)] = \int_{-\infty}^{\infty} x f(x, t_2) dx$$

[f(x,t2) is the density form of the random process X(t2)]

Let $t_2 = t_1 + C$

$$\therefore E[X(t_2)] = \int_{-\infty}^{\infty} x f(x, t_1 + C) dx = \int_{-\infty}^{\infty} x f(x, t_1) dx$$

$$= E[X(t_1)]$$

Thus $E[X(t_2)] = E[X(t_1)]$

Mean process {X(t1)} = mean of the random process {X(t2)}.

**Definition 2:**

If the process is first order stationary, then Mean = E(X(t)] = constant



## 4 Second Order Stationary Process

A random process is said to be second order stationary, if the second order density function stationary.

f (x 1 , x 2 ; t 1 , t 2 ) = f (x 1 , x 2 ; t 1 + C, t 2 + C )∀ x 1 , x2 and C.

E (X12 ), E (X 22 ), E (X1 , X2 )denote change with time, where

X = X(t1); X2 = X(t2).

## 5 Strongly Stationary Process

A random process is called a strongly stationary process or Strict Sense Stationary Process (SSS Process) if all its finite dimensional distribution are invariance under translation of time 't'.

fX(x1, x2; t1, t2) = fX(x1, x2; t1+C, t2+C)

fX(x1, x2, x3; t1, t2, t3) = fX(x1, x2, x3; t1+C, t2+C, t3+C) In general

fX(x1, x2..xn; t1, t2…tn) = fX(x1, x2..xn; t1+C, t2+C..tn+C) for any t1 and any real number C.

## 6 Jointly - Stationary in the Strict Sense

{X(t)} and Y{(t)} are said to be jointly stationary in the strict sense, if the joint distribution of X(t) and Y(t) are invariant under translation of time.

**Definition Mean:**

$$\mu_X(t) = E[X(t_1)], \quad -\infty < t < \infty$$

$\mu[X(t)]$ is also called mean function or ensemble average of the random process.

## 7Auto Correlation of a Random Process

Let X(t1) and X(t2) be the two given numbers of the random process {X(t)}. The auto correlation is

$$R_{XX}(t_1, t_2) = E\{X(t_1) x X(t_2)\}$$

Mean Square Value

Putting $t_1 = t_2 = t$ in (1), we get

$R_{XX}(t,t) = E[X(t)\, X(t)]$

$\Rightarrow \quad R_{XX}(t,t) = E[X^2(t)]$ is the mean square value of the random process.

## 8 Auto Covariance of A Random Process

$$C_{XX}(t_1, t_2) = E\{[X(t_1) - E(X(t_1))][X(t_2) - E(X(t_2))]$$
$$= R_{XX}(t_1, t_2) - E[X(t_1)]E[X(t_2)]$$

**Correlation Coefficient**

The correlation coefficient of the random process {X(t)} is defined as

$$\rho_{XX}(t_1, t_2) = \frac{C_{XX}(t_1, t_2)}{\text{Var}\, X(t_1) x \text{Var}\, X(t_2)}$$

Where $C_{XX}(t_1, t_2)$ denotes the auto covariance.

Where CXX (t1, t2) denotes the auto covariance.

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.

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

## First Order Strictly Stationary Process

## Stationary Process (or) Strictly Stationary Process (or) Strict Sense Stationary Process [SSS Process]

A random process X(t) is said to be stationary in the strict sense, it its statistical characteristics do not change with time.

## Stationary Process:

Formula: $E[X(t)] = \text{Constant}$

$\gamma[X(t)] = \text{Constant}$

Consider the RP X(t) = Cos (w0t +θ) where θis uniformly distributed in the interval -π to π. Check whether X(t) is stationary or not? Find the first and Second moments of the process.

Given X(t) = cos (W0t + θ)

Where θ is uniformly distributed in (-π, π)

$$f(\theta) = \frac{1}{\pi - (-\pi)} - \frac{1}{2\pi}, \qquad -\pi < 0 < \pi$$

[from the def. of uniform distribution]

To prove (i) X(t) is SSS process
(ii) E[X(t)] = Constant
(iii) Var [X(t)] = Constant

$$E[X(t)] = \int_{-\infty}^{\infty} X(t) f(\theta) d\theta$$

$$= \int_{-\pi}^{\pi} \cos(w_0 t + \theta) . \frac{1}{2\pi} d\theta$$

$$= \frac{1}{2\pi} [\sin(w_0 t + \theta)]_{-\pi}^{\pi}$$

$$= \frac{1}{2\pi} [\sin w_0 t + \pi] + \sin[\pi - w_0 t]$$

$$= \frac{1}{2\pi} [-\sin w_0 t + \sin w_0 t] = 0$$

$$E[X^2(t)] = E[ws^2(w_0 t + \theta)]$$

$$= \frac{1}{2} E[1 + \cos(2w_0 t + 2\theta)]$$

$$E[1] = \int_{-\pi}^{\pi} 1/2\pi \, d\theta = 1$$

$$E[\cos(2w_0 t + 2\theta)] = \int_{-\pi}^{\pi} \cos(2w_0 t + 2\theta) . 1/2\pi$$

$$= 1/2\pi \left[ \sin \frac{(2\omega_0 t + 2\theta)}{2} \right]_{-\pi}^{\pi}$$

$$= 1/4\pi [0] = 0$$

$$\therefore \Rightarrow E[X^2(t)] = 1/2 (1) + 0 = Y_2$$

$$\therefore \Rightarrow E[X^2(t)] = 1/2 (1) + 0 = Y_2$$

$$\text{Var}[X(t)] = E[X^2(t)] - [E[X(t)]]^2$$

$$= 1/2 - 0$$

$$= 1/2 = \text{const}$$

$\therefore$ X(t) is a SSS Process./

The RP X(t): Acos (w 0t +θ is not stationary if A and w0 are constants and θ is

uniformly distributed random variable in (0, π). In X(t) = Acos (w 0t +θ

In 'θ'uniformly distributed

$$f(\theta)=\frac{1}{\pi-0}=\frac{1}{\pi} \qquad 0<\theta<\pi$$

$$E[X(t)]=\int_{-\infty}^{\infty}X(t)f(\theta)d\theta$$

$$=\int_{0}^{\pi}A\cos(w_0t+\theta)\,\frac{1}{\pi}d\theta$$

$$=\frac{A}{\pi}\left[\sin(w_0t+\theta)\right]_0^{\pi}$$

$$=-\frac{2A}{\pi}\sin w_0t\neq \text{const.}$$

$\therefore$ N(t) is not a stationary process.

∴ N(t) is not a stationary process.

## Second Order And Wide Sense Stationary Process

A Process is said to be second order stationary, if the second order density function statistics.

$$f(x_1,x_2:t_1,t_2)=f(x_1,x_2:t+\delta,t_2+\delta),\ \forall x_1\,x_2 \text{ and } \delta$$

If a random process X(t) is WSS then it must also be covariance stationary. In X(t) is WSS
i) E[X(t)] = μ= a const.

(ii) R(t1, t2) = a fy of (t 1 - t2) The auto covariance function is given by

The auto covariance fn is gn by

$$C(t_1,t_2)=R(t_1,t_2)-E[X(t_1)X(t_2)]$$

$$=R(t_1-t_2)-E[X(t_1)]E[X(t_2)]$$

$$=R(t_1-t_2)-\mu(\mu)$$

$$=R(t_1-t_2)-\mu^2$$

Which depends only on the time difference. Hence X(t) is covariance stationary.
If X(t) is a wide sense stationary process with auto correlation R ()τ =Ae−α(),
determine the second order moment of the random variable X(8) - X(5).

Given $R(\tau) = Ae^{-\alpha(\tau)}$

$R(t_1, t_2) = Ae^{-\alpha(t_1 - t_2)}$

$E\left[X^2(t)\right] = R(t,t) = A$

$E\left[X^2(8)\right] = A$

$E\left[X^2(5)\right] = A$

$E\left[X(8)X(15)\right] = R|8,5| = Ae^{-\alpha}(8,5)$

$= Ae^{-3\alpha}$

The second moment of X(8) - X(5) is given by

$$E\left[X(8) - X(15)\right]^2 = E\left[X^2(8)\right]^2 + E\left[X^2(5)\right] - 2E\left[X(8)X(5)\right]$$

The second moment of X(8) - X(5) is given by

$= A + A - 2Ae\text{-}3\alpha$

$= 2A(1–e–3\alpha)$