Module 2

Review of Random Variable and Random Process

Syllabus (9hrs)

- **Review of random variables – both discrete and continuous. CDF and PDF, statistical averages**. (Only definitions, computations and significance)

- **Entropy, differential entropy. Differential entropy of a Gaussian RV. Conditional entropy, mutual information**.

- **Stochastic processes, Stationarity. Conditions for WSS and SSS. Autocorrelation and power spectral density. LTI systems with WSS as input.**

Basics of Probability

Random experiment

- It is an **experiment** the outcome of which cannot be predicted precisely.
- All possible identifiable outcomes of a random experiment constitute its **sample space S**.
- An **event** is any subset of possible outcomes.

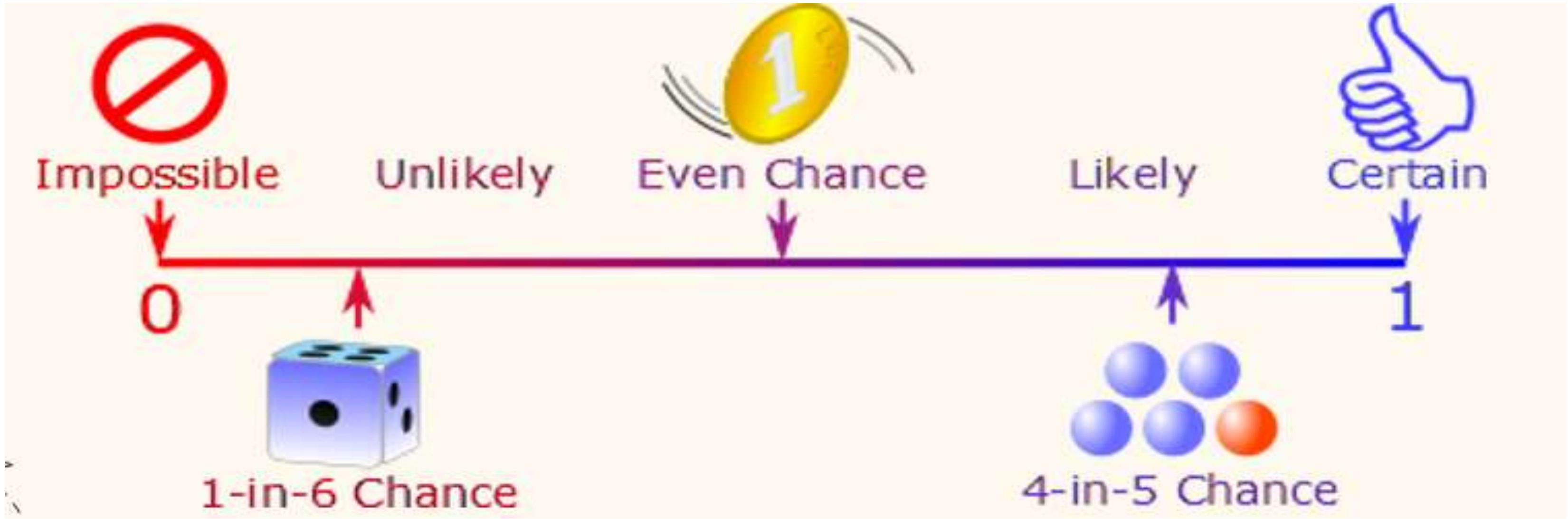
- Example – For tossing a coin S = { H T }

  – For rolling a die, S = { 1, 2, …, 6 }

Probability

- It is a branch of mathematics that deals with the occurrence of a random event.
- The value is expressed from zero to one.
- Probability has been introduced to predict how likely events are to happen.

Probability of event to happen P(E) = Number of favourable outcomes/Total Number of outcomes

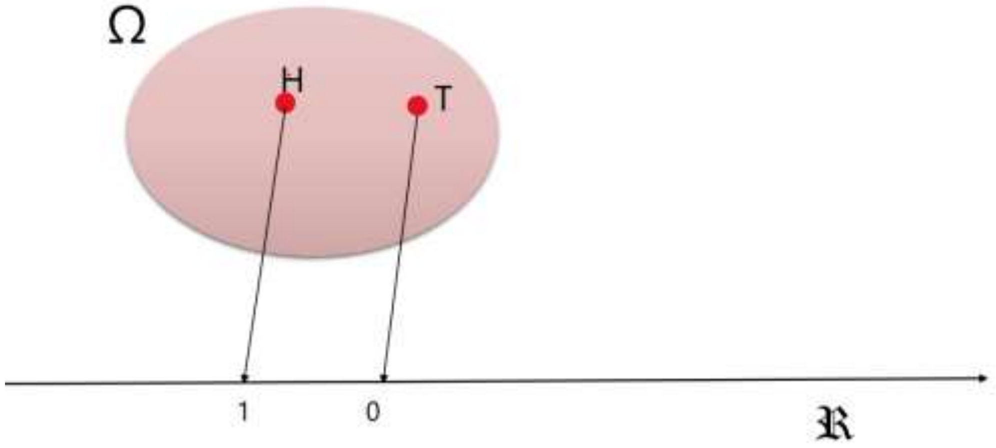


Random Variable

- A random variable is a rule that **assigns a numerical value to each outcome in a sample space**.
- Is a function that performs mapping between two sets.

    Two Sets – Sample space and Real line

- Random variables are denoted by capital letters X,Y, etc. ; individual values of the random variable X are X(ω)



Types of Random Variable

- Discrete Random Variable
- Continuous Random Variable

1. **Discrete Random Variable**

A discrete random variable can take only a **finite number** of distinct values such as 0, 1, 2, 3, 4, … and so on

2. Continuous Random Variable

If the random variable X can assume an **infinite and uncountable set of values**, it is said to be continuous random variable.

Mathematical tools to study Random Variable

1. Cumulative Distribution Function (CDF)
2. Probability Density Function (pdf)
3. Probability Mass Function (pmf)

1. Cumulative Distribution Function (CDF)

The **Cumulative Distribution Function (CDF)**, of a real-valued random variable X, evaluated at x, is the probability function that X will take a value less than or equal to x.

- _The CDF defined for a random variable and is given as

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

Properties

- Every CDF is non-decreasing and right continuous.
- The probability that a random variable takes on a value less than the smallest possible value is zero.

$$F_x(-\infty) = 0$$

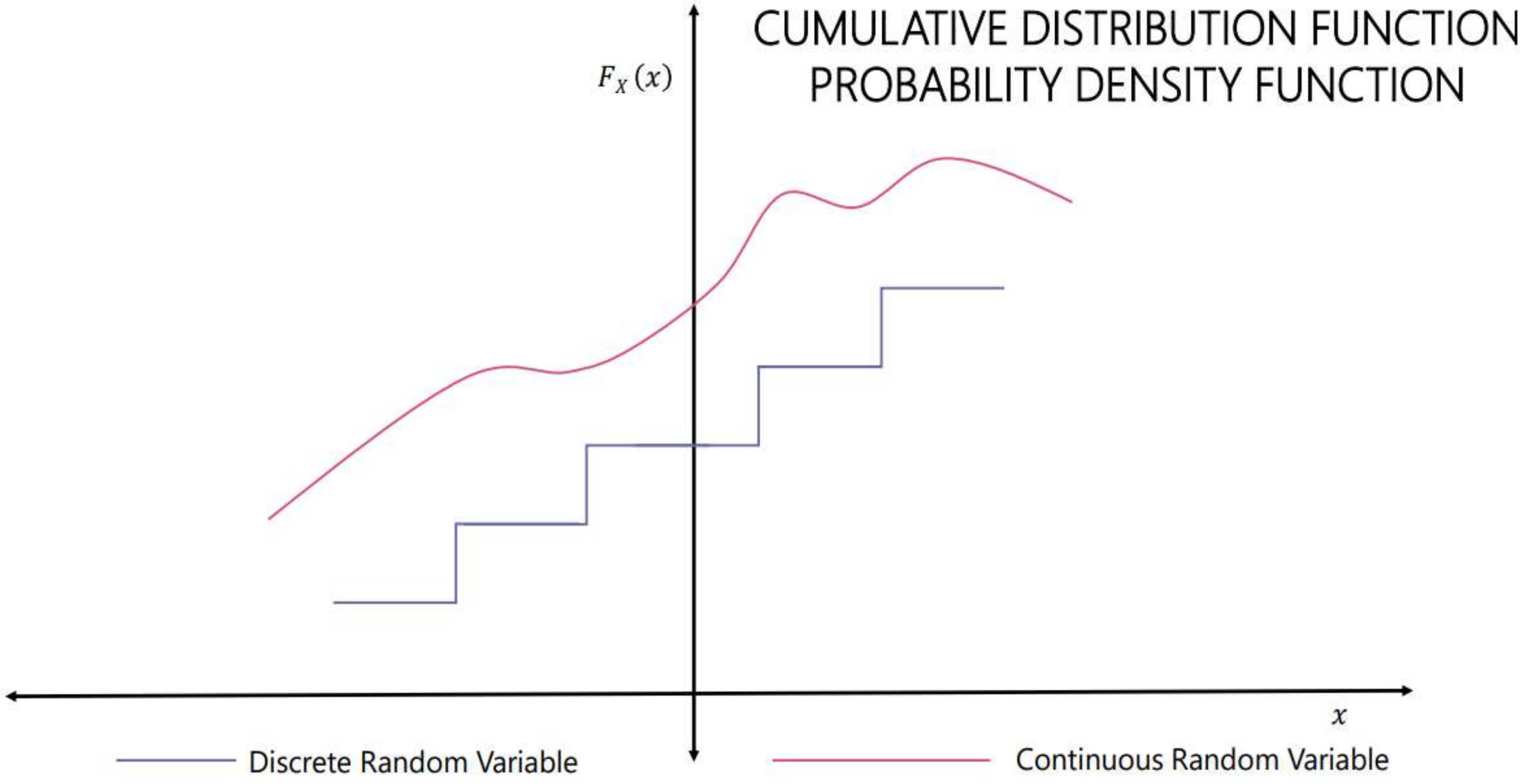
- The probability that a random variable takes on a value less than or equal to the largest possible value is one.

$$F_x(\infty) = 1$$

- The output of the CDF always ranges between 0 and 1.

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

- $F_x(x1) \leq F_x(x2)$, **if x1 <= x2**
- $P(x1 < X \leq x2) = F_x(x1)$ **-** $F_x(x2)$
- **P(X > x1) = 1 -** $F_x(x1)$


CUMULATIVE DISTRIBUTION FUNCTION
PROBABILITY DENSITY FUNCTION
$F_X(x)$
x
Discrete Random Variable
Continuous Random Variable

2. Probability Density Function (pdf)

- Derivative of Cumulative Distribution Function (CDF).
- For a **continuous RV** X, pdf is given by

$$\boldsymbol{f(x) = \frac{\mathrm{d}}{\mathrm{d}x} F(x)}$$

Properties

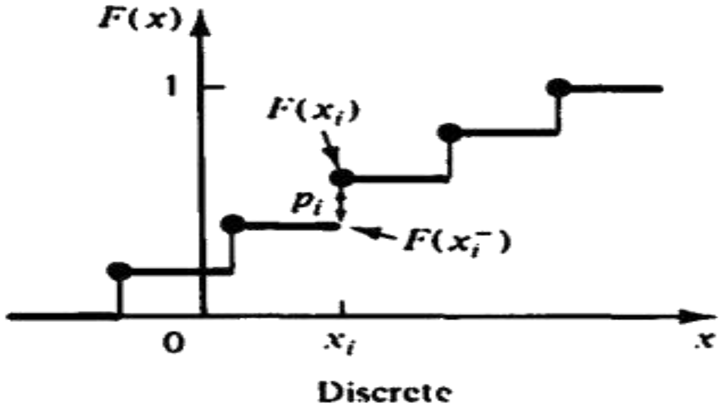
- $\int_{-\infty}^{\infty} \boldsymbol{f}(x)\, \boldsymbol{dx} = \mathbf{1}$
- CDF can be obtained by integrating the probability density function.
- Probability of the event within the range (x1<x<= x2) is given by area under the pdf curve within the same range.

$$\boldsymbol{P(x1 < X \leq x2)} = \int_{x_1}^{x_2} \boldsymbol{f}(x)\, \mathbf{dx}$$

- $\boldsymbol{P(X \leq x1)} = \boldsymbol{F_x(x1)} = \int_{-\infty}^{x_1} \boldsymbol{f}(x)\, \mathbf{dx}$
- $\boldsymbol{P(X > x1)} = \boldsymbol{F_x(x1)} = \int_{x_1}^{\infty} \boldsymbol{f(x)\ dx}$

3. Probability Mass Function (pmf)

- For a discrete RV X, let the discontinuities in the CDF be at points x_k
- This means that X takes the values x_k with the probability, $P\{X = x_k\} = p_k = F(x_k) - F(x_k^-)$
- The function $p_X(x) = P\{X = x\}$ is called the **probability mass function** (pmf) or the **point density function.**



QUESTIONS

1. In an experiment, a trial consists of three successive tosses of coin. If we define random variable as number of heads appearing in a trial, plot pmf and cdf.

2 In a game a six faced die is thrown. If 1 or 2 comes the player gets Rs 30, if 3 or 4 the player gets Rs 10, if 5 comes he loses Rs. 30 and in the event of 6 he loses Rs. 100. Plot the CDF and PDF of gain or loss

Expectation Of A Function

Consider a Random Variable X with pdf f(x).

Let y= ax + b, where a & b are constants.

then

$$\mathbf{E[y] = E[ax + b]} = \int (\boldsymbol{ax} + \boldsymbol{b})\, \boldsymbol{f}(\boldsymbol{x})\boldsymbol{dx}$$

Similarly

$$\mathbf{E}[\boldsymbol{y}^{\mathbf{2}}] = \mathbf{E}[(\boldsymbol{ax} + \boldsymbol{b})^{\mathbf{2}}] = \int (\boldsymbol{ax} + \boldsymbol{b})^{\mathbf{2}}\, \boldsymbol{f}(\boldsymbol{x})\boldsymbol{dx}$$

Statistical averages

1. MEAN

- Gives the DC Component of a signal.
- The mean (Expected Value) of a rv X, denoted by μ_X or $E(X)$ is defined by,

$$\mu_X = m_1 = E(X) = \begin{cases} \sum_k x_k p_X(x_k) & \text{for discrete RV} \\ \int_{-\infty}^{\infty} x f_X(x) dx & \text{for continuous RV} \end{cases}$$

Properties of Mean

1. $E[cX] = cE[\mathrm{X}]$, If c is any constant,
2. $E(X+Y) = E(X) + E(Y)$, if X and Y are any rvs
3. $E[XY] = E[X]E[Y]$, if X and Y are independent rvs
4. E[X+c] = E[X] + c , If c is any constant

2. MEAN SQUARE VALUE

- Gives the Total Power of a signal.

$$m_2 = E[X^2] = \begin{cases} \sum_k x_k^2 p_X(x_k) & \text{for discrete RV} \\ \int_{-\infty}^{\infty} x^2 f_X(x)dx & \text{for continuous RV} \end{cases}$$

3.VARIANCE

- Gives the ac power of a signal.

$$\sigma_X^2 = Var(X) = m_2 - m_1^2 = E(X^2) - \{E(X)\}^2$$

4. STANDARD DEVIATION

- Gives the rms value of ac component.

$$\sigma_x = \sqrt{Var(X)}$$

5. COVARIANCE

- Measure of relationship between two random variables.
- Positive covariance indicates that two variables tend to move in the same direction.
- Negative covariance indicates that two variables tend to move in inverse direction.

$$Cov(X,Y) = E[(X - \mu_X)(Y - \mu_y)]$$

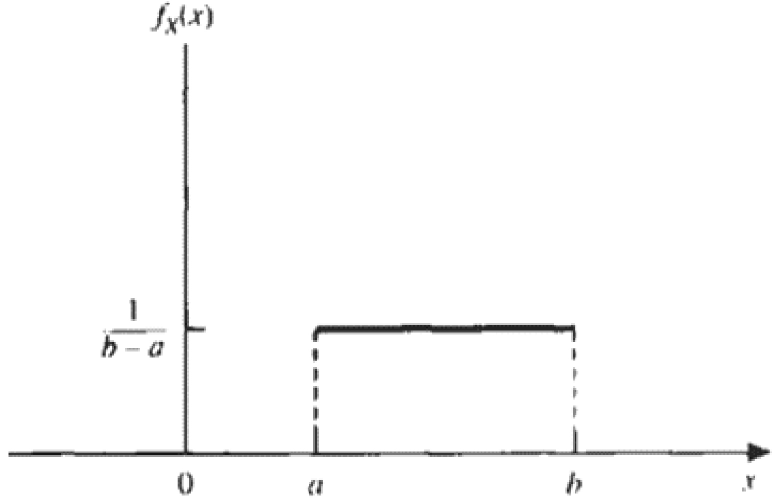
$$Cov(X,Y) = \mathrm{E[XY]} - \mu_X \mu_y$$

Probability Distributions

1. Uniform Distribution

- A continuous rv X is called a uniform rv over (a, b) if its pdf is given by,

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



- Mean & Variance of Uniform Distribution

$$\text{Mean: } \mu_X = E(X) = \frac{a+b}{2}$$

$$\text{Variance: } \sigma_X^2 = \frac{(b-a)^2}{12}$$

Question

- A continuous RV is uniformly distributed in the interval 0 to 10. Sketch the pdf and determine the following probabilities.

i. $P(X \leq 2)$

ii. $P(X > 9)$

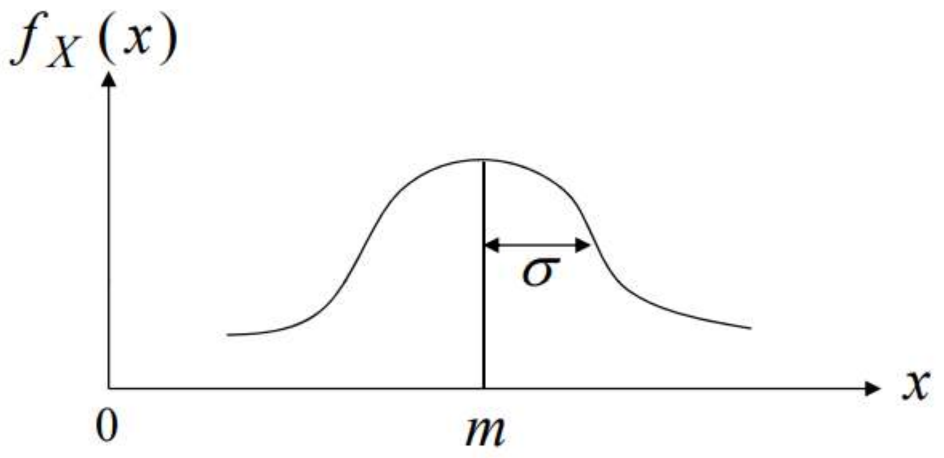
iii. $P(3 \leq X < 7)$

2. Normal (Gaussian) Distribution

- A continuous rv X is said to follow a normal distribution if it has the pdf,

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

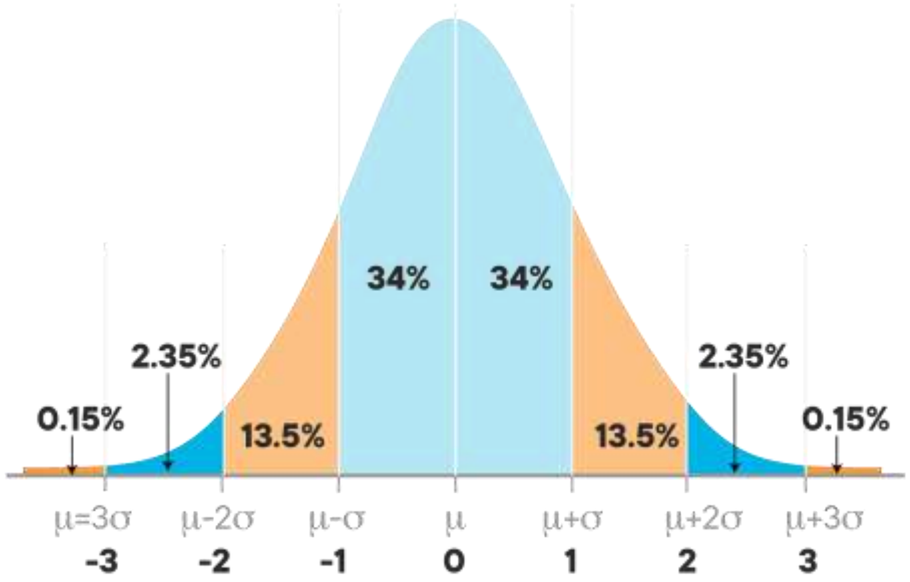
- x is the variable
- μ is the mean
- σ is the standard deviation



- Mean & Variance of the Normal Distribution, $N(\mu, \sigma^2)$

$$\text{Mean: } \mu_X = E(X) = \mu$$

$$\text{Variance: } \sigma_X^2 = \sigma^2$$



- The mean helps to determine the line of symmetry of a graph,
- The standard deviation helps to know how far the data are spread out.
- If the standard deviation is smaller, the data are somewhat close to each other and the graph becomes narrower. If the standard deviation is larger, the data are dispersed more, and the graph becomes wider.

PROPERTIES

- In a normal distribution, the mean, median and mode are equal.(i.e., **Mean = Median= Mode**).
- The **total area** under the curve should be **equal to 1**.
- The normally distributed curve should be **symmetric at the centre**.
- There should be **exactly half** of the values are to the right of the centre and exactly half of the values are to the left of the centre.
- The normal distribution should be defined by the mean and standard deviation.
- The normal distribution curve must have only one peak. (i.e., **Unimodal**)
- The curve approaches the x-axis, but it never touches, and it extends farther away from the mean.

Central limit theorem

- It is one of the most remarkable results in probability theory
- Let $X_1, X_2, \ldots, X_n$ be a sequence of independent, identically distributed rv's each with mean μ_i and variance σ_i^2
- Let $X = X_1 + X_2 + \cdots X_n$
- Then X forms a rv with mean $\mu = \mu_1 + \mu_2 + \cdots \mu_n$ and variance $\sigma^2 = \sigma_1^2 + \cdots + \sigma_n^2$
- CLT states that, under certain general conditions, as $n \to \infty$, the distribution of X approaches normal distribution with mean μ and variance σ^2
- In other words,

$$\frac{X-\mu}{\sigma} \sim N(0,1)$$

- The central limit theorem allows us to approximate a sum or average of i.i.d random variables by a normal random variable. This is extremely useful because it is usually easy to do computations with the normal distribution

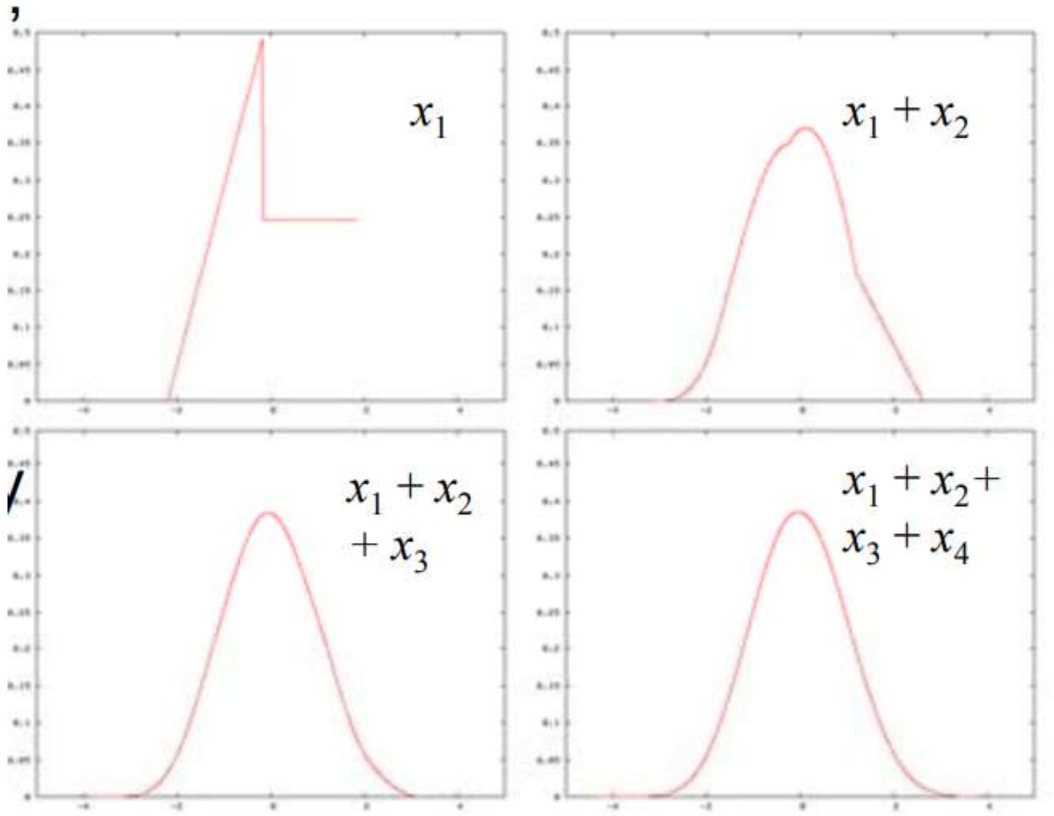


Illustration of convergence to Gaussian distribution

Entropy, differential entropy. Differential entropy of a Gaussian RV. Conditional entropy, mutual information.

INFORMATION

- Information: **any new knowledge** about something.
- $$Information \propto \frac{1}{probability\ of\ occurence}$$
- Messages containing knowledge of a high probability of occurrence ʹNot very informative Not very informative
- Messages containing knowledge of low probability of occurrence ʹMore informative.
- Suppose X - be a random experiment

  Xi – be a random event in that experiment

$$\mathbf{I[X{=}Xi]} = -\log_2 P[\mathrm{X} = X_i]\ \text{bits}$$

$$= -\log_{10} P[\mathrm{X} = X_i]\ \text{decit / Hartely}$$

$$= -\log_e P[\mathrm{X} = X_i]\ \text{nat}$$

Properties

- I[X=Xi] = 0 when P[X=Xi] = 1 : a symbol that is certain to occur contains information.

- I[X=Xi] > 0 when $0 \leq$ P[X=Xi] < 1 : the information measure is monotonic and non-negative.

- I[X=X_1] > I[X= X_2] : when P[X=X_1] < P[X= X_2]

- I[(X=X_1)(X= X_2)] $= -\log_2 P$ [(X=X_1)(X= X_2)] = I[X=X_1] + I[X= X_2] : information is additive for statistically independent events.

ENTROPY

- Average information transmitted per symbol by the source.
- The entropy of a random variable is the average level of "information", "surprise", or "uncertainty" inherent to the variable's possible outcomes.
- The entropy of a discrete random variable X is a function of its Probability Mass Function and is defined by
- Entropy = probability x information
- $$H(X) = -\sum_{i=1}^{m} P[X = X_i] \log_2 P[X = X_i] \text{ bits/symbol}$$
- It is the amount of uncertainty before we receive it before we receive it.
- It tells us how many bits of information per symbol we expect to get on the average

Case 1: Let probability of one symbol out of m symbol is 1.

- So probability of rest of the symbols = 0.

then **H(X) = 0** bits/symbol

Case 2: Let all the symbols are having same probability (Equiprobable)

- P[X= X_1] = P[X= X_2] = P[X= X_3] = …………….= P[X= X_m] $= \frac{1}{m}$

then **H(X) = $\mathbf{log_2 m}$** bits/symbol

- RANGE OF ENTROPY : **$0 \leq H(X) \leq \log_2 m$** bits/symbol

INFORMATION RATE (R)

- Average information per second.
- If the time rate at which source X emits symbols is r, the information rate R of the source is given by

$$\mathbf{R = r\,H(X)} \quad \text{bits/sec}$$

R – Information rate

H(X) – Entropy

r – symbol rate or rate at which symbols are generated

QUESTION

1) A source generates three symbols with probabilities 0.25, 0.25 and 0.50. Assuming independent generation of symbols, Calculate the average bit rate.

2) A continuous signal is bandlimited to 5kHz. The signal is quantized in 8 levels of a PCM with probabilities 0.25, 0.2, 0.2, 0.1, 0.1, 0.05 & 0.05. Calculate the entropy & the rate of information.