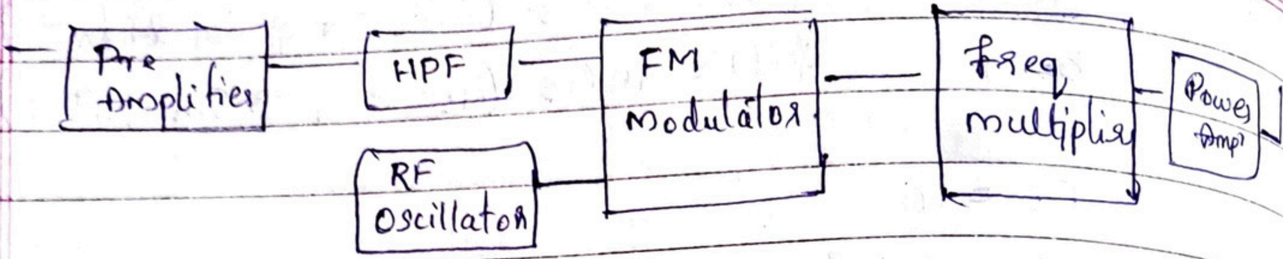


MODULE - 2.

Random Variables :

Random Variables are fundamental concepts ind probability & Statistics used to represent uncertain outcomes in numerical terms.

There a two types :- Discrete & Continous.

1. Discrete - Takes countable values
   0, 1, 2 ....
   represented by Probability mass fun.

2. Continous - Takes uncountable ~~i~~ infinite values.
   measured, not counted.
   represented by Probability density fn.

Page No.
Date / /

Cumulative Distributive Function – that a variable x takes a value less than or equal to x.

Example :- Suppose Random Variable X = the result of rolling a fair six sided die.
possible values – 1, 2, 3, 4, 5, 6
probability of each – 1/6

~~Cumulative is de~~ CDF is defined as $\boxed{F(x) = P(X \leq x)}$

Probability Density function – $\boxed{f(x) = \frac{d}{dx} F(x)}$

eg:- $F(x) = x^2$ for $0 \leq x \leq 1$
then $f(x) = \frac{d}{dx} x^2 = 2x$

Key characterstics of a pdf :-

1. Used for continous random variable (not discrete)
2. The PDF itself doesnot give a probability directly at a single point because for continous variables $P(X = x) = 0$. Instead it gives the density & the area under the curve gives the probability over the

Page No.
Date

3. total area under the PDF curve = 1

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

4. Probability for the interval (b/w two values a & b)

$$P(a \leq x \leq b) = \int_a^b f(x)\, dx.$$

(1) $f(x) = \begin{cases} 2x & 0 \leq x \leq 1 \\ 0 & \text{otherwise.} \end{cases}$

find $P(0.2 \leq X \leq 0.5)$

$$P = \int_{0.2}^{0.5} 2x\, dx = 0.25 - 0.04 = 0.21$$

Statistical Averages :- Includes mean, variance, moments etc.

Entropy

A measure of the uncertainity or randomness in a random variable.

→ Differential entropy – The continous analog of entropy, used for continous random variables.

→ Conditional entropy – The amount of uncertainity remaining in one random variable given another.

→ Mutual Information – Measures the amount of information one random variables contains another.

Page No.

Date

1. A game consists of tossing a coin three times & notify the outcome. player wins if all tosses give the same results. ie, three heads or three tails. Otherwise, he losses the game. Calculate the probability of winning the game.

pbb (possible outcomes) = {HHH, HHT, HTT, THT, TTT, THH, HTH, TTH)

pbb (winning the game) = {HHH, TTT}.

$= \frac{2}{8}$

Uniform Distribution

A uniform distribution is a type of probability all possible outcomes are equaly likely [equal pbb]

– Continous Case

In this case the probability density is constant over an interval. ie, if a random number is equally likely to be anywere b/w zero & 10, each value has the same probability density.

– Discrete Case

Each outcome has the probability, eg :- rolling a fair six sided die. [each case 1 -- 6] has a probability of 1/6

Page No.
Date / /

Probability Rules.

- Discrete uniform distribution

$$P(X=x) = \frac{1}{n} \qquad n = \text{no. of outcomes.}$$

- For continous uniform distribution.

$$f(x) = \frac{1}{b-a} \qquad a \le x \le b$$

Differential entrophy of a continous random variable with PDF with $f(x)$

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

for a uniform distribution we know that PDF $f(x) = \frac{1}{b-a}$ for all $a \le x \le b$

So the above integral simplifies to

$$h(X) = \ln(b-a) \quad \text{nats - unit.}$$

in terms of bits

$$h(X) = \log_2 (b-a)$$

Page No.
Date

2. let x be a continous random variable with following PDF

$$f(x) = \begin{cases} ke^{-4x} & x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

find the value of the constant k.

$$f(x) = ke^{-4x} \quad x \geq 0$$

$$= \int_0^\infty k e^{-4x} = 1$$

$$= k \int_0^\infty e^{-4x} = 1$$

$$k \left(\frac{e^{-4x}}{-4} \right)_0^\infty = 1$$

$$\frac{-k}{4} [0-1] = 1$$

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

$$\boxed{k = 4}$$

3. A Continous random variable x is uniformly distributed in the interval [0, 8]. find the differential entrophy h(x)?

Page No.
Date / /

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

uniform distribution $f(x) = \frac{1}{b-a}$ for $a \le x \le b$

$$h(x) = \ln(b-a)$$

there $a = 0 \quad b = 8$

$$h(x) = \ln(8-0)$$
$$= \underline{\underline{2.0794}} \text{ nats}.$$

In terms of bits $h(x) = \log_2 (b-a)$

$$= \log_2 8 = 3 \text{ bits}.$$

Entropy of a discrete random variable with possible values x_i and probabilities $P(x_i)$ is gn by $H(X) = -\sum_{i=1}^{i=n} P(x_i) \log_2 P(x_i)$ bits.

1. Find the entropy of binary memoryless source emitting two equally probable messages.

There are two messages (binary source) so $n = 2$
The msg are equally probable.
$\therefore P(x_1) = P(x_2) = 1/2$

Page No.
Date

entropy $H(X) = -\sum_{i=1}^{i=n} P(x_i) \log_2 P(x_i)$

$= -\sum_{i=1}^{2} P(x_i) \log_2 P(x_i)$

$= -\left[P(x_1)\log_2 P(x_1) + P(x_2)\log_2 P(x_2)\right]$

$= -\left[\frac{1}{2}\log_2 \frac{1}{2} + \frac{1}{2}\log_2 \frac{1}{2}\right]$ bit.

$= 1$ bit

2) Calculate the entropy of ~~th~~rolling a fair die.

$n = 6$

$P(x_i) = P(x_1) = P(x_2) \ldots P(x_6) = 1/6$

$H(X) = -\sum_{i=1}^{6} P(x_i)\log_2 P(x_i)$

$= -\left[P(x_1)\log_2 P(x_1) + P(x_2)\log_2 P(x_2) + \ldots P(x_6)\log_2 P(x_6)\right]$

$= -\left[\frac{1}{6}\log_2(1/6) + \ldots\ldots \frac{1}{6}\log_2(1/6)\right]$

$= -\left[-\frac{1}{6}\log_2(1/6) + \ldots\ldots + \frac{1}{6}\log_2(1/6)\right]$

$= +\left[6 \times \frac{1}{6}\log_2 6\right] = \log_2 6 = 2.584$

Page No.
Date / /

3) Find the mean & variance of a uniformly distributed random variable x, distributed in the range $\{1, 3\}$

Mean of uniformly distributed random is gn by $\boxed{\mu = \frac{a+b}{2}}$

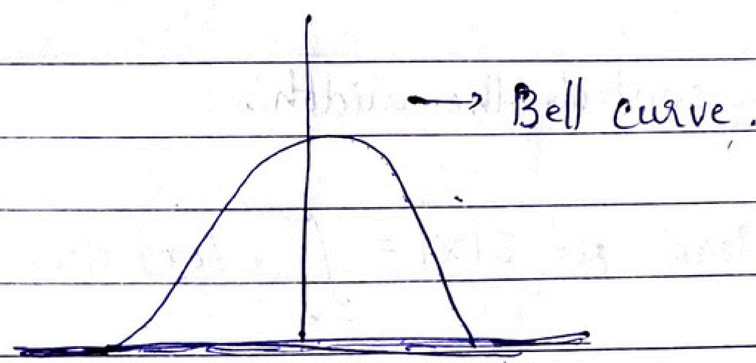
$a = 1, \; b = 3$

$$\mu = 2$$

$$\text{variance}, \; \sigma^2 = \frac{(b-a)^2}{12} = \frac{1}{3}$$

Gaussian Random Variable.

It is a continous random variable whose probability distribution follows the bell-shaped normal distribution curve.



→ Perfectly Symmetrical.

→ Concentrated around the peak & decrease on either side

→ peak represent the most probable event.

Page No.

Date

Standard Deviation.

Measures that describes the amount of variation or dispersion of a set of values around its mean. It indicates how much the data points deviate from the central value. ~~A low Standard.~~

Gaussian random variable – its defined by two parameters

Mean (μ) –

Variance (σ^2)

Probability density function (PDF)

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

σ – controls the width.

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

Variance $\sigma^2 = Var(X) = E[(X-\mu)^2]$

$$= \int_{\infty}^{\infty} (x-\mu)^2 f_X(x)\, dx$$

Page No.
Date / /

$$Var(X) = E|X^2| - (E|X|)^2 \qquad \begin{array}{l} E|X| = \mu \\ E|X^2| = \mu^2 + \sigma^2 \end{array}$$

1. Derive the differential entropy of a Gaussian random variable X with mean μ & variance σ^2.

PDF eq^n $f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right)$

differential entropy is
$$h(X) = -\int_{-\infty}^{\infty} f_X(x) \ln f_X(x)\, dx.$$

Taking log of PDF
$$\ln f_X(x) = \frac{1}{2} \ln(2\pi\sigma^2) - \frac{(x-\mu)^2}{2\sigma^2}$$

Multiply $f_X(x)$ & integrate.
$$h(x) = -\int f_X(x) \left[-\frac{1}{2} \ln(2\pi\sigma^2) - \frac{(x-\mu)^2}{2\sigma^2}\right] dx$$

$$= \frac{1}{2} \ln(2\pi\sigma^2) \int f_X(x)\, dx +$$

$$\frac{1}{2\sigma^2} \int (x-\mu)^2 f_X(x)\, dx.$$