## Random Variable (R.V)

In probability and statistics, a random variable, random quantity, aleatory variable, or stochastic variable is described informally as a variable whose values depend on outcomes of a random phenomenon. The random variable can be classified into two types

- Discrete Random Variable and Continuous Random Variable

**A discrete random variable** has a countable number of possible values. The probability of each value of a discrete random variable is between 0 and 1, and the sum of all the probabilities is equal to 1. Examples of Discrete Random Variables

Recall the case of dice, each face is numbered as 1,2,…,6.

if the dice is fair, then:-

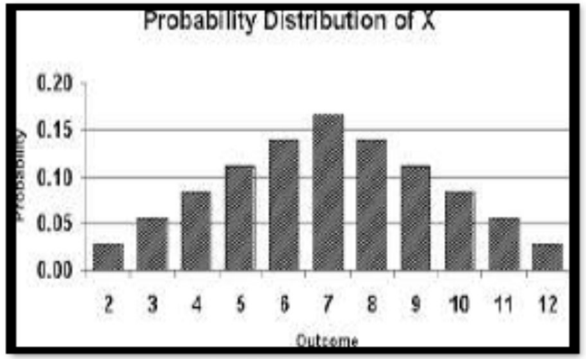
P(1) = P(2) = P(3) = ….P(6) = 1/6. Also $\sum_{i=1}^{6} P(X_i) = 1$.

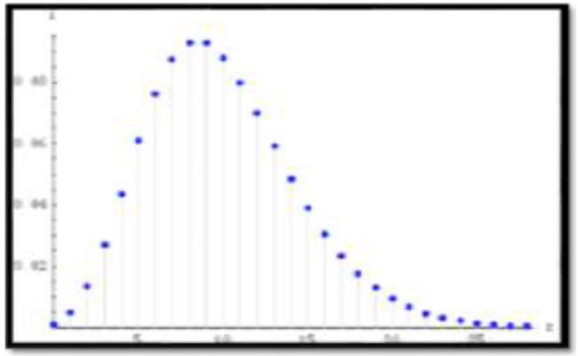
$$\bar{X} = mean\ of\ RV = \sum_{i=1}^{N} X_i\ P(X_i)$$

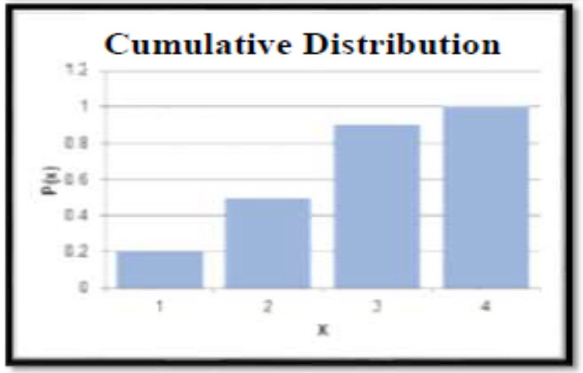
$$\overline{X^2} = mean\ square\ of\ RV = \sum_{i=1}^{N} X_i^2\ P(X_i)$$

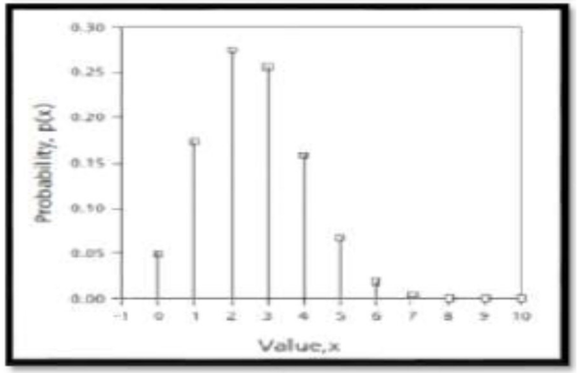
$$\sigma^2 = Variance\ of\ RV = \overline{X^2} - (\bar{X})^2$$

With a standard deviation of $\sigma$.



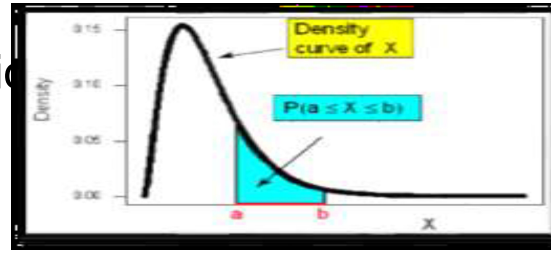






**Homework**: find the expected value (mean) $\bar{X}$, mean square $\overline{X^2}$ , standard deviation $\sigma$ and variance $\sigma^2$ of a fair dice ?

## -Continuous Random Variable



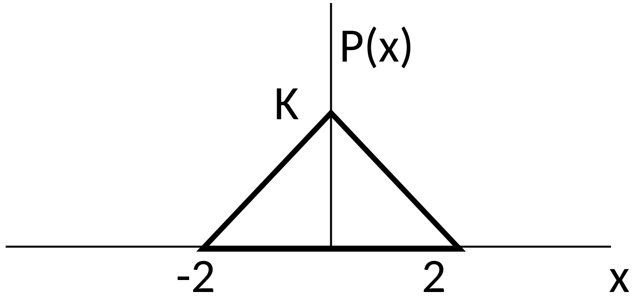

Here X can be all real values not discrete then we call P(X)=PDF=Probability Density functi
PDF gives the probability that X lies between any two points $X_1$ & $X_2$.

$$\textbf{P( X2 > X > X1)} = \int_{X1}^{X2} P(X)\,dx$$

Note that 1. $\int_{-\infty}^{\infty} P(X) = 1 \ldots$ 2. $\overline{X} = \int_{-\infty}^{\infty} X.\,P(X)\ \ dx \ \ldots$ 3. $\overline{X^2} = \int_{-\infty}^{\infty} X^2.\,P(X)\ \ dx \ \ldots$ 4. $\sigma^2 = \overline{X^2} - (\overline{X})^2$

If X is a random voltage signal then $\overline{X}$ ● D.C value of X, $\overline{X^2}$ ● total power (normalized on X) & $\sigma^2$ ● A.C. power of X. Also we can define the commutative distribution function (CDF) of X as $F(x) = \int_{-\infty}^{x} P(X)\ \ dx$. The advantage of the CDF is that it can be defined for any kind of random variable (discrete, continuous, and mixed)

***Example*** *If X a continuous R.V. having the following PDF.*
*Find: a.* ***Constant k ; b. P(X>1) , c.*** $\overline{X}, \overline{X^2}$ and $\sigma^2$




**Solution:**

a. $\int_{-\infty}^{\infty} P(X)dx = 1 \ = \int_{-2}^{2} P(X)dx = \frac{1}{2}.K.(4)$ K=$\frac{1}{2}$

b. $P(X>1) = \int_{1}^{2} P(X)dx = \int_{1}^{2} \frac{1}{2} - \frac{X}{4} = \frac{X}{2} - \frac{X^2}{8}\Big|_{1}^{2} = \frac{1}{8}$

c. $\overline{X} = \int_{-2}^{0} X.(\frac{1}{2} + \frac{X}{4})dx + \int_{0}^{2} X.(\frac{1}{2} - \frac{X}{4})dx = 0$

$\overline{X^2} = 2.\int_{0}^{2} X^2.(\frac{1}{2} - \frac{X}{4})dx = \frac{2}{3}$

$\sigma^2 = \overline{X^2} - (\overline{X})^2 = \frac{2}{3} - 0 = \frac{2}{3}$

**Homework:** ***If*** $P(X) = \frac{a}{2} e^{-a|x|}$ ***, find*** $\overline{X}$ , $\overline{X^2}$ ***and*** $\sigma^2$

The purpose of a communication system is to carry information-bearing baseband signal generated by an information source from one point to another over a communication channel, with high efficiency and reliability. Figure 1 illustrates the functional diagram and the basic elements of a digital communication system.

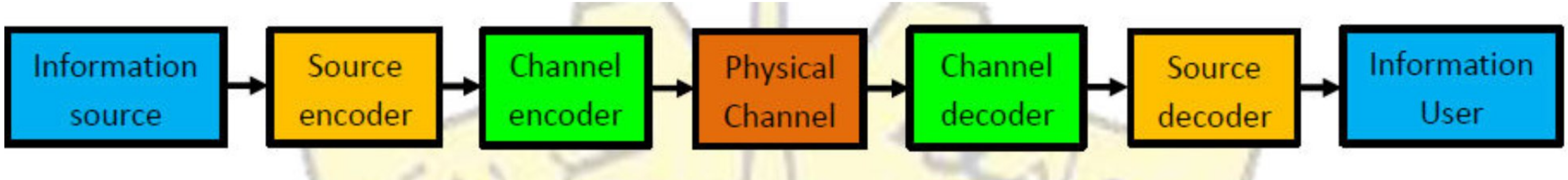



Figure 1. Communication system

- ❖ The information source may be either an analogue signal, such as an audio or video signal or a digital signal, such as the output of the computer that is discrete in time and has a finite number of the computer characters knows as information sequence.
- ❖ source encoding is a process of efficiently converting the output of either an analogue or digital source into a sequence of binary digits. It is also called data compression.
- ❖ Information theory provides a quantitative measure of the information contained in message signal and allows us to
determine the capacity of a communication system to transfer this information from source to destination. Through the use of
coding, redundancy can be reduced from message signal so that channels can be used with improved efficiency.

## Self Information:

Suppose that the source of information produces finite set of messages $X_1$, $X_2$, $X_n$ with prob. $P(X_1)$, $P(X_2)$,…$P(X_n)$ such that

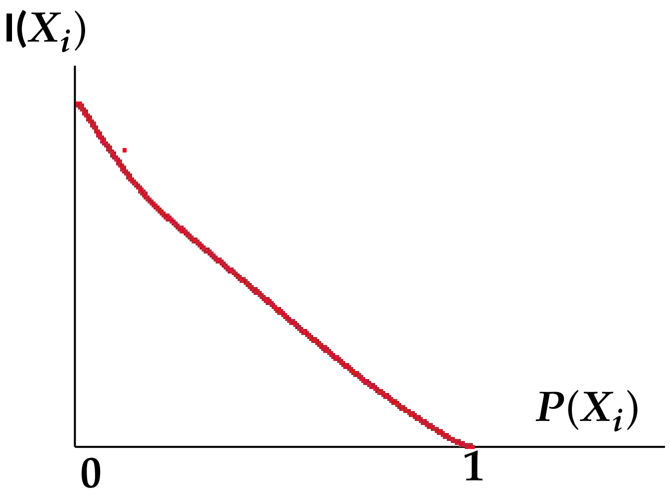
$$\sum_{i=1}^{N} P(X_i) = 1.$$

The amount of information gained from knowing that the source produces the messages $X_i$ as follows:

1. Information is zero if $P(X_i)=1$.
2. Information increases as $P(X_i)$ decreases.
3. information is a positive quantity.

The function that relates $P(X_i)$ with information of $X_i$ called: $I(X_i)$ = self information of $X_i$.

$$I(X_i) = -\log_a P(X_i)$$




The unit of $I(X_i)$ depends on **a**:

1. If a=2, $I(X_i)$ has the unit of bits.
2. if a=e=2.718, $I(X_i)$ has the unit of nat.
3. if a=10, $I(X_i)$ has the unit of hartly.

Note that:

$$\log_a P = \frac{Ln\,(P)}{Ln\,(a)}$$

***Example :*** A fair dice is thrown, find the amount of information gained if you are told that 4 will appear.

***Solution :***

Fair dice = $P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = \frac{1}{6}$

$$I(4) = -\log_2 P(4) = -\log_2 \frac{1}{6} = \log_2 6$$

$$I(4) = \frac{Ln\ (6)}{Ln\ (2)} = 2.5844\ bits$$

**Example*:*** Find the amount of information containing in a black and white TV picture if each picture has $2*10^5$ dots (pixels) and each pixel has 8 equal prob. Level of brightness.

**Solution**

Information / pixel = $-\log_2 P(\text{level}) = -\log_2 \frac{1}{8} =$ 3 bits

Information / picture = $3*2*10^5$ = 600 K bits

**Source Entropy**

If I(Xi), i=1,2,…n are different for a source producing unequal probability symbols, then the statical average of I(Xi) will give the average amount of uncertainty associated with the source X, this average is called **source entropy** and denoted by H(X) and measured by **bit per symbol.** This is given by

$$H(x) = \sum_{i=1}^{n} P(X_i)\ I(X_i) = -\sum_{i=1}^{n} P(X_i)\ \log_2 P(X_i) \qquad (bits/symbol)$$

**Example** : Find the entropy of source producing the symbols.   P(X)=[0.25  0.1  0.15  0.5]

**Solution**:

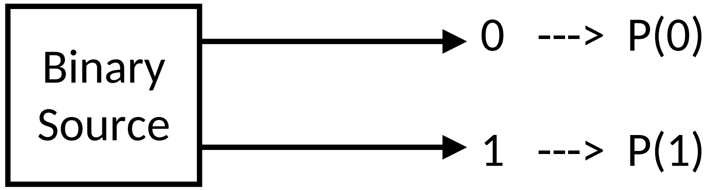
$$H(x) = -\sum_{i=1}^{n} P(X_i)\ \log_2 P(X_i) = \frac{1}{Ln\ (2)}\ [0.25\ Ln\ (0.25) + 0.1\ Ln\ (0.1) + 0.15\ Ln\ (0.15) + 0.5\ Ln\ (0.5)$$

$H(x) =$ ***1.7427 bits/symbols***

**Example:** Find and plot the entropy of a binary source.

**Solution:**

Binary Source P(0)+P(1)=1 => P(1)=1-P(0)



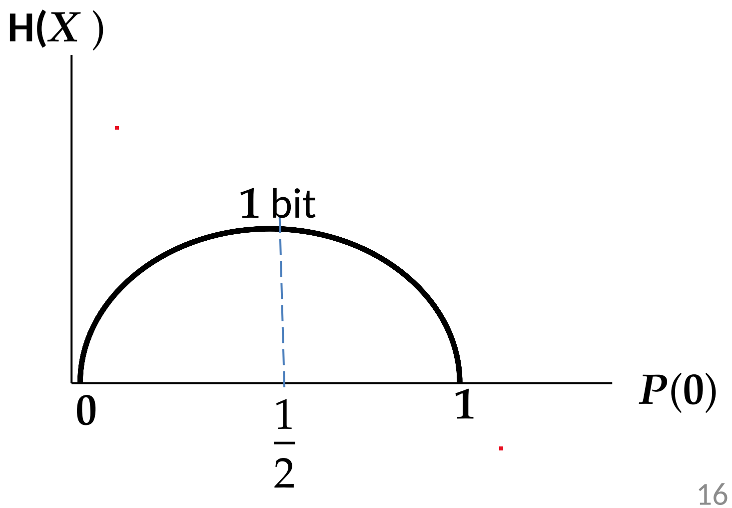

$$H(x) = -\sum_{i=1}^{n} P(X_i)\ \log_2 P(X_i) = [P(0)\ \log_2 P(0) + (1-P(0))\log_2 (1-P(0))$$

Note that:

$H(x) = log_2(n)$ if the ***n*** symbols $X_1$, $X_2$, …. $X_n$ are equal probability and $P(X_i) = \frac{1}{n}$

$H(x) =$ 0, if one of the symbols has prob. = 1

$H(x) = \max = 1 \text{ if } \mathrm{P}(0) = \mathrm{P}(1) = \frac{1}{2}$



**Source Entropy Rate R(X) :**

This is average rate amount of information produced per second.

$$R(X) = H(X) * \text{rate of producing symbols}$$

**Rate of producing symbols =** $\frac{1}{\tilde{\tau}}$**, Then**

$$R(X) = \frac{H(X)}{\tilde{\tau}}$$

$$\tilde{\tau} = \sum_{i=1}^{n} \tau_i \, P(X_i)$$

$\tilde{\tau}$=Average time duration of symbols and $\tau_i$= time duration of Xi

**Example**: A source produces dots "•" & dashes "-" with probability P(dot) =0.65, if time duration of a dot is 200 ms and that for a dash is 800 ms. Find the average source entropy ***R(X)***.?

**Solution:**

P(dot) = 0.65 **------>** P(dash)=1-P(dot) =1-0.65 =0.35

τ-dot = 200 ms , τ dash = 800 ms ....

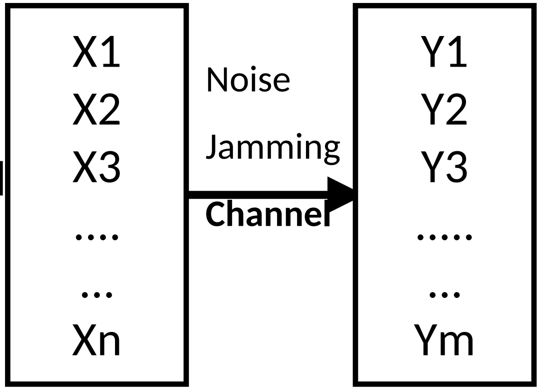
$$\tilde{\tau} = \sum_{i=1}^{2} \tau_i \, P(X_i) = [200*0.65 + 800*0.35] = 410 \; ms$$

***H(X)***= - [ 0.65 log2 (0.65) + 0.35 log2 (0.35)] = 0.934 bit/symbol*

$$R(X) = \frac{H(X)}{\tilde{\tau}} = \frac{0.934}{410\times 10^{-3}} = 2.278 \text{ bit/sec}$$

**Mutual Information :**

Consider the set of symbols $X_1$, $X_2$...... $X_n$ can be produced by source. The receiver may receive $Y_1$, $Y_2$...... $Y_m$. if the noise and jamming are zero the set X = set Y and (n=m), however, due to noise and jamming, there will be conditional probability P(Y/X).




Theoretically if the noise and jamming is zero, then Set X= Set Y and m=n. However, due to noise and jamming , there will be a conditional

**Definition:**

$P(X_i)$ is called a priori prob of the symbol $X_i$ which is the prob of selecting $X_i$ for transmission.

$P(X_i/Y_j)$ is known a posteriori prob of $X_i$ after the reception of $Y_j$.

The amount of information that $Y_j$ provides about $X_i$ is called **"Mutual Information"** between $X_i$ & $Y_j$ . This is given by :

$$\mathbf{I}(x_i\ ,\ y_j) = Log_2 \text{ (a posteriori prob.)/(a priori prob.)} = Log_2 \frac{P(X_i/Y_j)}{P(X_i)}\ bits$$

Also

$$\mathbf{I}(y_j, x_i) = Log_2 \frac{P(Y_j/X_i)}{P(Y_j)} = \mathbf{I}(x_i\ ,\ y_j)$$