Entropy.

Fundamental part of any communication system is information. Information or message produced by the Information source is received at the destination. The information that we passing to the destination is not deterministic. They are random in nature. The term Entropy indicate the measure of randomness of the information. For this we have to give a mathematical model for the information by relating probability with information.

see the examples.

Sun rises from the east : $P = 1$ $I = 0$

It is a sure event

There is no new information.

India got independence in 1947 $P = 1$ $I = 0$

Historical event

You will get two mails in next 2 days

Here the information depends on probability $P \uparrow$ $I \downarrow$

Dog will bark $P \uparrow$ $I \downarrow$

PM will come to your home Tomorrow $P \downarrow$ $I \uparrow$

It will snow in Kerala in April $P \downarrow\downarrow$ $I \uparrow\uparrow$

You will not get a holiday on Sunday $P\downarrow\downarrow\downarrow$ $I\uparrow\uparrow\uparrow$

So

$$I_{xi} \propto \frac{1}{P(x_i)}$$

$$I = \log \frac{1}{P(x_i)}$$

The term Entropy denotes the degree of randomness. It is used to find out the degree of randomness of the message. A message is described by various probabilities of it's symbols.
Some symbols are highly probable. So these message symbols are less informative.
Some symbols are less probable. so they are highly informative.
Entropy is defined as the average of all the information associated with all of the symbols.
Consider we have a message $m = \{x_1, x_2, x_3 \cdots x_n\}$ with it's probabilities $p = \{P_1, P_2, P_3 \cdots P_n\}$
the information of each message symbol is given by

$$I_1 = \log_2 \frac{1}{P(x_1)} \quad I_2 = \log_2 \frac{1}{P(x_2)} \quad \cdots \quad I_n = \log_2 \frac{1}{P(x_n)}$$

Entropy $H(X) = \sum_{i=1}^{n} P(x_i) \log_2 \frac{1}{P(x_i)}$ b/s

$$H(X) = -\sum P_{x_i} \log_2 P(x_i) \text{ bits/symbol.}$$

Properties of Entropy:

* Entropy of a sure event is zero.

Proof: if we are sure an event will not happen

Then $P = 0$

$$H = \sum_{i=1}^{n} P_{x_i} \log \frac{1}{P_{x_i}}$$

when $P_{x_i} = 0 \Rightarrow H = 0$

An event will happen

Then $P = 1$

$$H = \sum_{i=1}^{n} \underset{\downarrow}{P_{x_i}} \log \frac{1}{P_{x_i} \rightarrow 1}$$

1 $\qquad \log 1 = 0$

$H = 0$

* when $P_{x_i} = \frac{1}{n}$ for all the n symbols : it means all the symbols are equally probable

$$H = \log_2 n$$

Proof: It is given that all the symbols in a message having probability

$$P_{x_i} = \frac{1}{n}$$

$$H = \sum_{i=1}^{n} P_{x_i} \log_2 \frac{1}{P_{x_i}}$$

$$= \sum_{i=1}^{n} \frac{1}{n} \log_2 \frac{1}{1/n}$$

if we have summation of n times $\frac{1}{n}$ that is equal to 1

$\therefore H = \log_2 n$ for equi probable symbols.

This is upper bound of Entropy i.e; H_{max}.

* If all the symbols are equiprobable Entropy is maximum.

Differential entropy

the Entropy

$$H(X) = \sum_{i=1}^{n} P_{x_i} \log \frac{1}{P_{x_i}}$$ This is of a discrete

random variable $X = (x_i, x_2 \dots x_n)$

In case of Continuous random variable (information generated by source is continuous in nature) the X will take infinite no: of values. So the probability distribution is taken by the PDF ($f_X(x)$) instead of P_{x_i}.

the differential Entropy is defined by

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