Lecture 6: Information Theory

Information Theory basics, definition, Uncertainity & Property

Information: It is the intelligence/ideas/message in Information Theory

Message: - Electrical signal
- Speech / voice
- Picture / image
- video
- text

In Communication System, Information is transmitted from Source to destination

$$\boxed{\text{Information Source}} \rightarrow \boxed{\text{Encoder}} \rightarrow \boxed{\text{Tx}} \rightarrow \boxed{\text{Channel}} \rightarrow \boxed{\text{Rx}} \rightarrow \boxed{\text{Decoder}} \rightarrow \boxed{\text{Information}}$$

source — destination

Uncertainty

- $X = \{x_0, x_1, \ldots, x_n\}$
- $P = \{P_0, P_1, \ldots, P_n\}$
- total probability $P = \sum_{i=1}^{n} P_i$

$n_0 \rightarrow P_0 = 0$, $\rightarrow P_0 = 1$ } No uncertatinity

$n_0 \rightarrow P_0 = 0.9$, $\rightarrow P_0 = 0.1$

- As probability of message decreases then uncertainty increasing

Message of Information

- It is a information content of a message
- Consider an information source emitting independent message $M = \{m_1, m_2, \ldots, m_n\}$ with probabilities of occurence is $P = \{P_1, P_2, \ldots, P_n\}$
- Here, $P_1 + P_2 + \ldots + P_n = 1$
- Amount of Information is given by,

$$I_k = \log_2\left(\frac{1}{P_k}\right) = \frac{\log(1/P_k)}{\log 2} \quad \boxed{\text{bits}}$$

Properties of Information

- More uncertainty about message then information is more

$P_1 = 1/4$, $P_2 = 1/2$

U_1 U_2

$\rightarrow U_1 > U_2$

To prove above property

- $I_1 = \log_2\left(\frac{1}{P_1}\right)$
$= \log_2\left(\frac{1}{1/4}\right)$
$= \log_2 4$
$= 2\log_2 2$
$= 2$ bits

$I_2 = \log_2\left(\frac{1}{P_2}\right)$
$= \log_2\left(\frac{1}{(1/2)}\right)$
$= \log_2(2)$
$= 1$ bit

Then, $I_1 > I_2$

- Receiver knows message being transmitted than information is zero

- $P = 1$
- $I = \log_2\left(\frac{1}{P}\right)$
$= \log_2 1$
$= 0$ bit

- I_1 is the information of message m_1 and I_2 of m_2, then Combine information of m_1 & m_2 = $I_1 + I_2$

$\rightarrow I_1 = \log_2\left(\frac{1}{P_1}\right)$ $\rightarrow$ Since messages m_1 & m_2 are independent,

$\rightarrow I_2 = \log_2\left(\frac{1}{P_2}\right)$ So combined probability = $P_1 . P_2$

$\rightarrow I = \log_2\left(\frac{1}{P}\right) = \log_2\left(\frac{1}{P_1 P_2}\right) = \log_2\left(\frac{1}{P_1}\right) + \log_2\left(\frac{1}{P_2}\right)$

$\rightarrow I = I_1 + I_2$

- If there are $M = 2^N$ equally likely messages, then amount of Information carries by each message will be = N bits

- Probability of each message = $\frac{1}{M}$
- $I = \log_2\left(\frac{1}{P}\right)$ $\rightarrow m = 2^N$
$= \log_2(m)$ $\rightarrow I = \log_2 2^N$
$= N\log_2 2$
$I = N$ bits

Examples on Information in information Theory

① calculate amount of information for

a) $P_1 = 1/4$, b) $P_2 = 3/4$

$P_1 = 1/4$

$$I_1 = \log_2\left(\frac{1}{P_1}\right) = \log_2(4) = 2\log_2(2) = 2 \text{ bits}$$

$P_2 = 3/4$

$$I_2 = \log_2\left(\frac{1}{P_2}\right) = \log_2\left(\frac{4}{3}\right) = \frac{\log(4/3)}{\log 2} = 0.415 \text{ bits}$$

② A card is selected at random from a desk of playing cards. If you have been told that it is red in colour.

1) How much information you have receive.

2) How much more information you needed to completely specify the card.

1) ~~Total~~ cards = 52

Total Red cards = 26

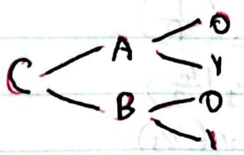
$$P = \frac{26}{52} = \frac{1}{2}$$

$$I = \log_2\left(\frac{1}{P}\right) = \log_2(2) = 1 \text{ bit}$$

2) Probability $P = \frac{1}{26}$

$$I = \log_2\left(\frac{1}{P}\right) = \log_2(26) = \frac{\log 26}{\log 2} = 4.7 \text{ bits}$$

③ Consider discrete memoryless Source 'C' that o/p two bits at a time. This Source Comprises two binary Sources 'A' and 'B', whose o/p are equally likely to occur and each Source contributing one bit. Suppose that the sources within the Source 'C' are independent. What is the information content of each o/p from Source 'C'?



→ total Samples of C

$$C = \{00, 01, 10, 11\}$$

$$\rightarrow P_C = 1/4$$

→ $P_A = 1/2$ $\quad P_B = 1/2$

→ Combined Probability of A & B

$$P_C = P_A P_B$$
$$= \frac{1}{2} \times \frac{1}{2}$$
$$= \frac{1}{4}$$

→ Information of C

$$I_C = \log_2\left(\frac{1}{P_C}\right)$$
$$= \log_2(4)$$
$$= 2\log_2(2)$$
$$\boxed{I_C = 2 \text{ bits}}$$

Entropy basics, definition & Property

Destination - It is average information of Symbols

- If we have $M = \{x_1, x_2, \ldots, x_n\}$ messages with Probability $P = \{P_1, P_2, \ldots, P_n\}$

- Then information of messages

$I_1 = \log_2\left(\frac{1}{P_1}\right)$, $I_2 = \log_2\left(\frac{1}{P_2}\right), \ldots\ldots,$ $I_n = \log_2\left(\frac{1}{P_n}\right)$

So entropy $H = \dfrac{\text{Total Information}}{\text{NO of messages}}$

$$= \frac{I_1 + I_2 + \ldots + I_n}{n}$$

$$\boxed{H = \sum_{i=1} P_i \log_2\left(\frac{1}{P_i}\right)}$$ bits/symbol

Properties

1) Entropy is zero, if the event is sure

- $P = 0$

$$H = \sum_{k=1}^{m} P_K \log_2\left(\frac{1}{P_K}\right)$$

$$\boxed{H = 0}$$

- $P = 1$

$$H = \sum_{k=1}^{m} P_K \log_2\left(\frac{1}{P_K}\right)$$

$$= 1 \log_2\left(\frac{1}{1}\right)$$

$$\boxed{H = 0}$$

2) When $P_K = 1/m$ for all 'm' Symbols, then Symbols are equally. So $H = \log_2 m$

$\rightarrow P_K = 1/m$

$\rightarrow H = \sum_{K=1}^{m} P_K \log_2\left(\frac{1}{P_K}\right)$

$$= \boxed{\sum_{K=1}^{m} \left(\frac{1}{m}\right)} \log_2(m)$$

$\rightarrow \boxed{H = \log_2 m} = H_{max}$

upper bound of Entropy

Entropy Source Efficiency, Redundancy & Information rate

- Source Efficiency

$$\eta = \frac{H}{H_{max}} \qquad \Big| \; H_{max} = \log_2 m$$

where H = Calculated Entropy of Source

H_{max} = max. Entropy

- Redundancy of Source

$$R_e = 1 - \eta$$

- Information Rate

$$R = r\,H = \left[\frac{\text{messages}}{\text{sec}}\right]\left[\frac{\text{bits}}{\text{message}}\right] = \frac{\text{bits}}{\text{sec}}$$

where, r = rate at which messages are generated [messages / sec]

H = entropy $\left[\frac{\text{bits}}{\text{messages}}\right]$

1) For a discrete memoryless source, there are three symbols with $P_1 = \alpha$ and $P_2 = P_3$. Find the entropy of the source

→ $P_1 = \alpha$

→ $P_2 = P_3$

→ $P_1 + P_2 + P_3 = 1$

$\alpha + P_2 + P_2 = 1$

$\alpha + 2P_2 = 1$

$P_2 = P_3 = \frac{1-\alpha}{2}$

$$H = \sum_{k=1}^{3} P_k \log_2\left(\frac{1}{P_k}\right)$$

$$H = \alpha \log_2\left(\frac{1}{\alpha}\right) + 2\left(\frac{1-\alpha}{2}\right)\log_2\left(\frac{2}{1-\alpha}\right)$$

$$= \alpha \log_2\left(\frac{1}{\alpha}\right) + (1-\alpha)\log_2\left(\frac{2}{1-\alpha}\right)$$

1) Show that the Entropy of the source with following Probability distribution is $\left[2 - \frac{2+n}{2^n}\right]$

S	S_1	S_2	S_3	....	S_n
P	$1/2$	$1/4$	$1/8$	....	$1/2^n$

$\rightarrow$ $$H = \sum_{k=1}^{n} P_k \log_2\left(\frac{1}{P_k}\right)$$

$$= P_1 \log_2\left(\frac{1}{P_1}\right) + P_2 \log_2\left(\frac{1}{P_2}\right) + P_3 \log\left(\frac{1}{P_3}\right) + \ldots P_n \log_2\left(\frac{1}{P_n}\right)$$

$$= \frac{1}{2}\log_2(2) + \frac{1}{4}\log_2(4) + \frac{1}{8}\log(8) + \ldots + \left(\frac{1}{2^n}\right)\log_2(2^n)$$

$$= \frac{1}{2} + \frac{2}{4} + \frac{3}{8} + \ldots\ldots + \frac{n}{2^n}$$

$$= 2 - \frac{2+n}{2^n}$$

2) The source emits three messages with probability $P_1 = 0.7$, $P_2 = 0.2$ and $P_3 = 0.1$.

Calculate

1) Source Entropy
2) maximum Entropy
3) Source Efficiency
4) Redundancy

1) $$H = \sum_{k=1}^{3} P_k \log_2\left(\frac{1}{P_k}\right)$$

$$= 0.7 \log_2\left(\frac{1}{0.7}\right) + 0.2 \log_2\left(\frac{1}{0.2}\right) + 0.1 \log_2\left(\frac{1}{0.1}\right)$$

$$= 1.156 \text{ bits/message}$$

2) $-H_{max} = \log_2 m$

$$= \log_2 3 = \frac{\log 3}{\log 2} = 1.585 \frac{\text{bits}}{\text{message}}$$

3) $$\eta = \frac{H}{H_{max}} = \frac{1.1568}{1.585} = 0.73.$$

4) Redundancy $= 1-\eta = \cancel{2}\, 1-0.73 = 0.27$

4) A discrete source emits one of six symbols once every milisec. The symbol probability are $1/2$, $1/4$, $1/8$, $1/16$, $1/32$ & $1/32$. Find the source entropy and information rate

$$r = \frac{1}{T} = \frac{1}{10^{-3}} = 10^3 \text{ messages/sec}$$

$$H = \sum_{k=1}^{r} P_k \log_2\left(\frac{1}{P_k}\right)$$

$$= \frac{1}{2}\log_2(2) + \frac{1}{4}\log_2(4) + \frac{1}{8}\log_2(8) + \frac{1}{16}\log_2(16) + 2\times\frac{1}{32}\cancel{x}\log_2 32$$

$$= \frac{1}{2} + \frac{2}{4} + \frac{3}{8} + \frac{4}{16} + 2\times\frac{5}{32}$$

$$= 1.9375 \text{ bits/sec}$$

$$R = rH$$
$$= 10^3 \times 1.9375$$
$$= \cancel{1.9375}\ 1937.5 \text{ bits/sec}$$

Shannon- Fano encoding Algorithm

Example - Find the codewords occuring in the probability $\{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}\}$ for symbols S_1, S_2, S_3 & S_4. Find efficiency and redundancy of code.

1. The messages are first written in the order of decreasing probability.
2. Then divide the messages set into two most equiprobable Subset X and Y
3. The message of 1st set X is given bit 0 and message in the 2nd Subset is given bit 1.
4. The Procedure is now applied for each set seperately till end.
5. Finally we get the code word for respective symbol.
6. Calculation

→ efficiency $(\eta) = \frac{H}{\hat{H}_{max}}$

where, $H = \text{Entropy} = \sum_{i=1}^{n} P_i \log_2\left(\frac{1}{P_i}\right)$

$\hat{H}_{max} = \sum_{i=1}^{n} P_i n_i$

→ Redundancy

$R_e = 1 - \eta$

Symbol	Prob				Codeword	length (n)
S_1	$1/2$	X — 0			0	1
S_2	$1/4$	Y	0		10	2
S_3	$1/8$	1	1	0	110	3
S_4	$1/8$			1	111	3

- Efficiency $\eta = \frac{H}{H_{max}}$

Where H = entropy $= \sum_{i=1}^{4} P_i \log_2 P_i$

$= \frac{1}{2}\log_2 2 + \frac{1}{4}\log_2 4 + \left(\frac{1}{8}\log_2 8\right)$

$= \frac{1}{2} + \frac{1}{2} + \frac{6}{8}$

$= 1.75$ bits/symbol

$H_{max} = \sum_{i=1}^{4} P_i n_i$

$= \frac{1}{2}\times 1 + \frac{1}{4}\times 2 + \frac{1}{8}\times 3\times 2$

$= 1.75$ bits/symbol

$\therefore \eta = \frac{H}{H_{max}} = \frac{1.75}{1.75} = 1$

efficiency is 100%

- Redundancy $R_e = 1 - \eta$

$= 1 - 1 = 0$

Example

Find the codeword for the probabilitis $1/4, 1/4, 1/8, 1/8, 1/16, 1/16, 1/16, 1/16$ for symbols $S_1, S_2, \ldots, S_8$. Find the code efficiency and ~~Pacfe~~ Redundancy

Symbol	Prob	Codeword	Length
S_1	$1/4$	00	2
S_2	$1/4$	01	2
S_3	$1/8$	100	3
S_4	$1/8$	101	3
S_5	$1/16$	1100	4
S_6	$1/16$	1101	4
S_7	$1/16$	1110	4
S_8	$1/16$	1111	4

$$- H = \sum P_i \log_2\left(\frac{1}{P_i}\right)$$

$$= 2\left(\frac{1}{4}\log_2(4)\right) + 2\left(\frac{1}{8}\log_2(8)\right) + 4\left(\frac{1}{16}\log_2(16)\right)$$

$$= 1 + \frac{3}{4} + 1$$

$$= 2.75 \text{ bits/Symbol}$$

$$- \hat{H}_{max} = \sum P_i n_i$$

$$= 2\left(\frac{1}{4}\times 2\right) + 2\left(\frac{1}{8}\times 3\right) + 4\left(\frac{1}{16}\times 4\right)$$

$$= 1 + 3/4 + 1$$

$$= 2.75 \text{ bits/symbol}$$

$$- \text{Efficiency } \eta = \frac{H}{\hat{H}} = \frac{2.75}{2.75} = 1$$

$$- \text{Redundancy } R_e = 1 - \eta = 1 - 1 = 0$$