POORNIMA
COLLEGE OF ENGINEERING

DETAILED LECTURE NOTES

Information measure for continuos Random Variables:

The defination of mutual information for discreate random variable can be extendend to continuos random variable.

for $x, y \rightarrow$ random variable

$p(x, y) \rightarrow$ Probablity density function

$p(x)$ & $p(y) \rightarrow$ marginal pdf

Average mutual Information

It provide between two continuos random variable X and Y

$$I(X;Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(x)\, p(y|x) \log \frac{(p(y|x) \cdot p(x))}{p(x) \cdot p(y)}\, dx\, dy.$$

for continuos variable we write x and y in the two terms.

→ Avg mutual information can be carried over from discreate random variable but the consept and physical interpretation would be not.

→ A continuos Random variable is actually infinite and it requires infinite no of bit to represent.

a continuos Rv. It can be solved by diffrential entropy.

$$h(x) = -\int_{-\infty}^{\infty} p(x) \log p(x) \, dx$$

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

Properties of Diffrential Variable:

→ $h(x+c) = h(x)$ translation does not affect the diffrential entropy, where c is the constant.

→ $h(ax) = h(x) + \log |a|$

Relative Entropy: It measure a distance between distribution. In other mode, it provide how similar are two distributions.

It is called kullback leibler (KL) distance between two probablity mass function $p(x)$ and $q(x)$

$$D(p||q) = \sum_{x \in X} p(x) \log \left[\frac{p(x)}{q(x)}\right]$$

POORNIMA
COLLEGE OF ENGINEERING

DETAILED LECTURE NOTES

Source Coding Theoram:

The conversion ~~of the output~~ to a Discreate memoryless Source (DMS) into a sequence of binary symbol is called source coding.

→ Application of source coding is to minimize the average bit rate required for representation of the source by reducing redundancy of the information source.

A code is a ~~set~~ of vectors called codeword.

Average Code length:

For a DMS ~~set~~ of alphabet $\{x_1, x_2 \cdots x_m\}$ with corresponding probablity $\{p_1, p_2 \cdots p_m\}$ and code length $\{l_1, l_2 \cdots l_m\}$.

$$L = \sum_{i=1}^{m} p_i l_i$$

L is the average code length L per source symbol is thus represented by upper term.

Source Code Theoram:

It will help to explore efficient representation of symbols generated by a source.

Theoram Suppose a DMS output a symbol for every t seconds. Each symbol is selected from the finite source of symbol $i = 1, 2, \ldots L$. and occur with the probalities p of $(x_i) = 1, 2 \ldots L$

The entropy of this DMS is bit per source symbol

$$H(x) = \sum_{i=1}^{L} p(x_i) \log_2 \frac{1}{p(x_i)} \leq \log_2 L$$

$\downarrow$

Average Code Word length

Example Let $p(x_1) = 0.5$

$p(x_2) = 0.5$

$$H(x) = -\sum_{i=1}^{2} p(x_i) \log_2 p(x_i)$$

$$-0.5 \log_2 0.5 - 0.5 \log_2 0.5$$

$$= 0.5 + 0.5 = 1$$

$$\log_2 L = \log_2 2 = 1$$

$$\boxed{H(x) = \log_2 L}$$

Scanned by CamScanner

Example 2

$p(x_1) = 0.5$

$p(x_2) = 0.25$

$p(x_3) = 0.25$

$$H(x) = -0.5\log_2 0.5 - 0.25\log_2 0.25 - 0.25\log_2 0.25$$

$$= -0.5\log_2 2^{-1} - 0.25\log_2 2^{-2} - 0.25\log_2 2^{-2}$$

$$= 0.5 + 0.5 + 0.5 = 1.5$$

AND $\log_2 L = \log_2 3 = \frac{\log_{10} 3}{\log_{10} 2} = \frac{0.4771}{0.301} = 1.58$

$$\boxed{\log_2 L = 1.58}$$

Fixed Length Code

DMS output a symbol selected from a finite set of symbol x_i, $i = 1, 2 \ldots L$.

The number of binary digit R required for unique coding when L is the power of 2 =

$$R = \log_2 L$$

If L is the power of 2 than it should be considered by above equation.

$$R = \log_2 L$$

or

$$R = \lceil \log_2 L \rceil \text{ — Celing}$$

or

If L is not power of 2 → $\lfloor \log_2 L \rfloor + 1$ — floor

if L is not the power of two it should be round off at 2

for Ex $\lceil 1.6 \rceil$ it should be round off at 2

Variable Length Code: When the source symbol are not equally probable, we label unequal no of bit for an efficient coding.

Example

Fixed Length Code Example

Letter	Codeword	Letter	Codeword
A	000	E	100
B	001	F	101
C	010	G	110
D	011	H	111

Variable Length Code Example

Letter	Codeword	Letter	Codeword
A	00	E	101
B	010	F	110
C	011	G	1110
D	100	H	1111

Suppose the transmitted signal A BAD CAD

Fixed Length Code 000 001 000 011 010 000 001 Total bit - 21

Variable Length Code 00 010 00 100 011 00 010 Total bit - 18