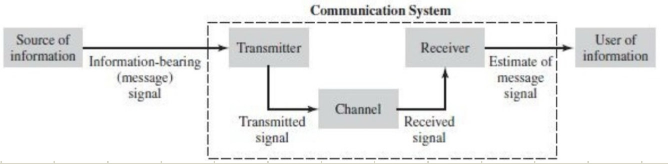
Digital Communication

UNIT-1 INFORMATION THEORY

Communication

→ Transfer of message from source to destination
- Analog
- Digital

Basic elements of Communication System

→ 

→ Transmitter : Converts message signal from source to channel in a suitable form of transmission

Channel : Transports message signal & delivers to receiver at some other location in space which can distort the signal due to channel imperfections

Receiver : Provides an estimate of original message signal from corrupted received signal

→ There are 2 basic modes of communication :

i) Broadcasting - Single powerful transmitter & many receivers which is relatively inexpensive.
(unidirectional flow of information bearing signal) 

ii) Point-to-Point - Communication between single transmitter & receiver
(Bi directional flow of information bearing signal) 

→ To support transmissions of multiple users from diverse locations, we can use multiple access techniques

Multiple Access Techniques

→ A technique whereby many subscribers / local stations can share the use of a communication channel at same time despite there being individual transmissions originating from widely different locations
(Permits communication resources which is shared by many users seeking to communicate with each other)

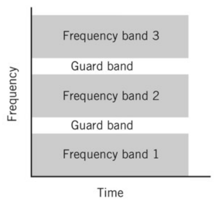
→ Examples of multiple access refers to satellite or radio channel

Multiple - Access	Multiplexing
→ Remote sharing of a communication channel (Satellite/Radio channel by users in dispersed locations)	→ Sharing of a channel by users confined to a local site (sharing telephone channel)
→ Requirements can change dynamically	→ Requirements are ordinally fixed

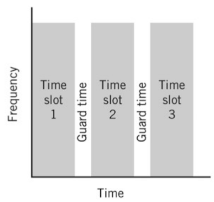
→ 4 Basic types of multiple access ex,

i) Frequency-division multiple access
ii) Time-division multiple access
iii) Code-division multiple access
iv) Space-division multiple access

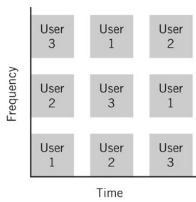
Allocate communication resources of channel through use of disjointedness



FDMA
Disjoint subbands allocated to users on continuous time basis with guard band to reduce interference



TDMA
(user is allocated full spectral freq but certain time slot with buffer zones called guard time



FHMA
Hopping from 1 user to another. Completely Random

Communication Network

→ Consists of interconnection of many nodes made of intelligent processors (microcomputers)

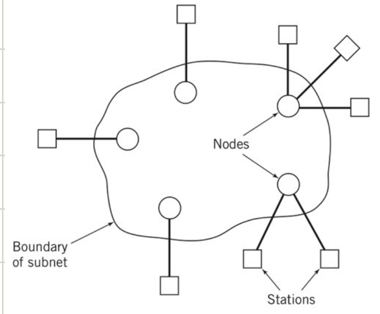
→ The primary purpose of nodes is routing data through the network

→ Network is a shared resource for stations Devices wishing to communicate

→ 2 Ways to route data :

i) Circuit Switching ⟶ Dedicated Route

ii) Packet Switching ⟶ Dynamic Route



Circuit Switching

→ A communication link is shared b/w different session using that link on a fixed allocation basis

→ It is usually controlled by centralized hierarchical control mechanism with Knowledge of network's organisation

→ To establish circuit-switched connection, an available path through the network is seized & then dedicated to the exclusive use of 2 stations wishing to communicate

→ During connection time, bandwidth & resources allocated are "owned" by 2 stations, until disconnected Circuit thus represents an efficient use of resources only to the extent that allocated bandwidth is properly utilized

→ Suitable for voice conversations as they usually are for longer time (1-2 mins) & it takes lesser time to setup (0.1 - 0.5s)

Packet Switching

→ Sharing is done on a demand basis & therefore has an advantage over circuit switching so when a link has traffic to send, link maybe fully utilized

→ The basic principle is to "store and send", any message larger than specified size is divided before transmission, into segments not exceeding specified size called packets.

→ Original message is reassembled at destination on a packet-to-packet basis

→ Network maybe viewed as distributed pool of network resources whose capacity is shared dynamically by a community of competing users willing to communicate

→ Packet switching is suited to computer-communication environment in which bursts of data are exchanged occasionally

Data Network

→ It is a network in which stations are all made up of computers and terminals

→ The design may proceed in an orderly way by looking at network in layered architecture

Refers to a process/ device inside computer designed to perform a specific function

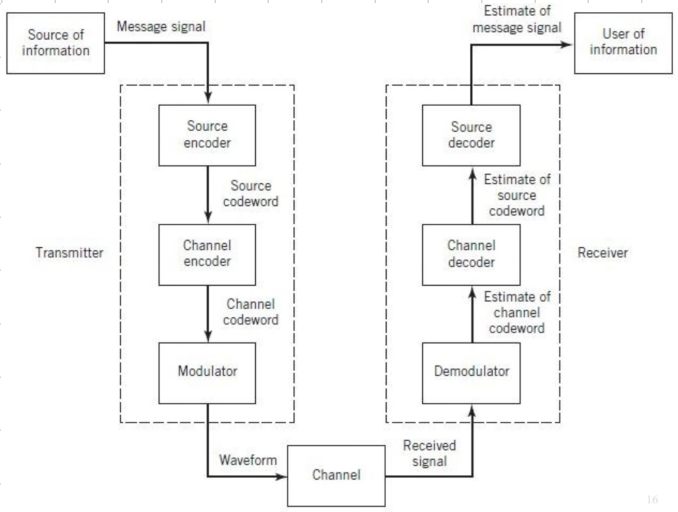
→ At system level, user views this layer as a "black box" described in terms of inputs, outputs, & their relationship

Digital Communication

→ For every action in transmitter, there is a corresponding opposite action at receiver

Functional blocks include :

- Source encoder - decoder
- Channel encoder - decoder
- Modulator encoder - decoder



→ Source encoder removes redudant information from message signal & responsible for efficient use of channel →Gives "source codeword"

→ Channel encoder adds redudancy to overcome channel imperfections & produces new sequence "channel codeword"

→ Channel codeword is longer than source codeword by virtue of controlled redundancy built into its construc

→ Modulator converts digital symbols into analog symbols suitable for transmission over the channel called "wave

→ At receiver, channel output is processed in reverse order of transmitter, reconstructing recognizable version of original message signal.

→ From this, it may look complex but actually easy to build

System is also robust offering great tolerance to physical effects

Information Theory

→ The purpose of communication system is to facilitate the transmission of signals generated by a source of information over a communication channel

→ Information Theory deals with mathematical modeling & analysis of a communication system rather than with physical sources & physical channels

→ It provides answer for 2 questions –

i) What is the irreducible complexity below which signal can't be compressed ?

A. Entropy is the probabilistic behaviour of a source of information

ii) What is the ultimate transmission rate of reliable communication over a noisy channel ?

A. Capacity is the intrinsic ability of a channel to convey information

Information

→ Assume a probabilistic experiment involving observation of output emitted by discrete source

This source output is modelled as stochiatic process denoted by discrete random variable S

where, $S = \{ s_0, s_1, \ldots, s_{K-1} \}$

with probabilities, $P(S = s_K) = p_K \quad (K = 0, 1, \ldots, K-1)$

which satisfies normalization property, $\sum_{K=0}^{K-1} p_K = 1$, $p_K \geq 0$

→ The idea of information is closely related to that of uncertainity or surprise

→ Consider $S = s_K$ which describes emission of symbol s_K by source with probability p_K,

i) If $p_K = 1$ & $p_i = 0$, There is no surprise ⇒ No information

ii) If S occurs with different probabilities & p_K is low, there is more surprise ⇒ information is higher

So from this, $s_K \propto \frac{1}{p_K}$

Amount of information gained after observing $S = s_K$ which occurs with probability p_K as log function,

$$I(s_K) = \log_2\left(\frac{1}{p_K}\right) = -\log_2(p_K)$$

→ Properties :

i) If $p_K = 1 \Rightarrow I(s_K) = 0$

ii) If $0 \leq p_K \leq 1 \Rightarrow I(s_K) \geq 0$

iii) If $p_K < p_i \Rightarrow I(s_K) > I(s_i)$

iv) If s_K & s_l are statistically independant ⇒ $I(s_K, s_l) = I(s_K) + I(s_l)$

→ When $p_K = 0.5$, $I(s_K) = 1$ bit, So we can say,

One bit is the amount of information we gain when one of 2 possible & equally likely events occur

Entropy

→ Entropy of discrete random variable representing output of source of information, is the measure of average information content per source symbol

→ Also, it is the expectation of $I(s_K)$ over all probable values taken by random variable S,

$$H(S) = E(I(s_K))$$

Entropy

$$= \sum_{K=0}^{K-1} p_K I(s_K)$$

$$= \sum_{K=0}^{K-1} p_K \log_2\left(\frac{1}{p_K}\right)$$

Properties

→ Entropy is bounded by $0 \leq H(S) \leq \log_2 K$

by This, we can make 2 statements

i) $H(S) = 0$ iff probability $p_K = 1$ for some K & remaining possibilities are set to zero (No uncertainity)

ii) $H(S) = \log K$ iff $p_K = \frac{1}{K}$ for all K (Maximum uncertainity)

To prove these statements :

i) Since $p_k \leq 1$,

$p_k \log_2\left(\frac{1}{p_k}\right)$ is always non-negative , so $H(s) \geq 0$ $\left(p_k \log_2\left(\frac{1}{p_k}\right) \geq 0 \Rightarrow H(s) \geq 0 \right)$

For lower bound, product must be 0 , Which is possible only if $p_k = 0$ or 1

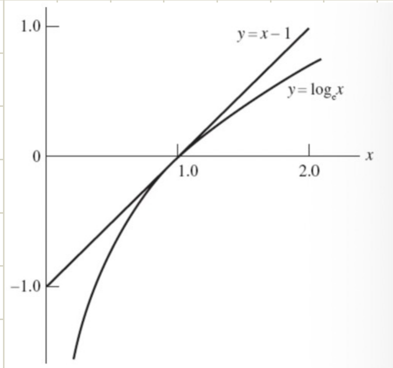
ii) To prove upper bound,

we make use of a property $\Rightarrow \log_e x \leq x - 1$

Now consider first any 2 different probability distributions,

denoted by $p_0, p_1, p_2 \cdots p_{K-1}$ & $q_0, q_1, q_2, \ldots, q_{K-1}$

on the alphabet $S = \{s_0, s_1, s_2, \ldots, s_{K-1}\}$ of a discrete source



Relative entropy of these 2 distributions

$$\Rightarrow D(p \| q) = \sum_{k=0}^{K-1} p_k \log_2\left(\frac{p_k}{q_k}\right) \qquad \left[\log_2 x = \frac{\ln x}{\ln 2} \;\&\; -\ln x \leq -(x-1) \right]$$

$$= -\sum_{k=0}^{K-1} p_k \log_2\left(\frac{q_k}{p_k}\right) \geq \frac{1}{\ln 2} \sum_{k=0}^{K-1} p_k \left(\frac{q_k}{p_k} - 1\right) = \frac{1}{\log_e 2} \sum_{k=0}^{K-1} (q_k - p_k) = 0$$

So, $D(p\|q) \geq 0$

↳ It is only zero when 2 distributions are identical

Relative Entropy - Measure of avg. information per source symbol emitted by one source relative to another source which emits the same set of symbols

$$D(p\|q) = \sum_{k=0}^{K-1} p_k \log_2\left(\frac{p_k}{q_k}\right)$$

Suppose $q_k = \frac{1}{K}$ for all k,

$$D(p\|q) = \sum_{k=0}^{K-1} p_k \log_2 p_k + \log_2 K \sum_{k=0}^{K-1} p_k$$

$$= -H(s) + \log_2 K$$

Then $D(p\|q) \geq 0$

$-H(s) + \log_2 K \geq 0$

$H(s) \leq \log_2 K \Rightarrow$ Upper bound proved !

Extended Source

→ If source emits independant symbols with entropy $H(s)$, then entropy of n^{th} order is

$H(S^n) = nH(s)$

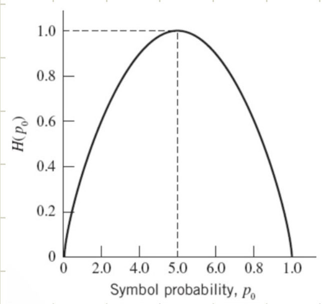
→ We use because sometimes avg. codeword length L can't exactly match $H(s)$, but grouping symbols can reduce the gap

Q. Consider Bernoulli Random Variable $S = \{0, 1\}$ occurring with probabilities p_0 and $1-$

Find entropy.

A. $H(s) = -p_0 \log_2 p_0 - p_1 \log_2 p_1$

$= -p_0 \log_2 p_0 - (1-p_0) \log_2 (1-p_0)$ bits



Q. Example: Suppose a discrete memoryless source emits symbols having the distribution $p_0 = P(S = s_0) = 1/4$, $p_1 = P(S = s_1) = 1/4$ and $p_2 = P(S = s_2) = 1/2$

Find the entropy of the discrete random variable S representing the symbols emitted by the source

A. $H(S) = \sum_{k=0}^{2} P_k \log_2\left(\frac{1}{P_k}\right) = \frac{1}{4}\log_2 4 + \frac{1}{4}\log_2 4 + \frac{1}{2}\log_2 2 = 1.5$ bits

Consider the 2nd order extension of the source. Find the entropy of the extended source, denoted by $H(S^{(2)})$

Let the symbols emitted by the extended source be represented as $\sigma_0, \sigma_1, \ldots, \sigma_8$

★ $\left(H(S^{(n)}) = nH(S)\right)$

Symbols of $S^{(2)}$	σ_0	σ_1	σ_2	σ_3	σ_4	σ_5	σ_6	σ_7	σ_8
Symbol block	s_0s_0	s_0s_1	s_0s_2	s_1s_0	s_1s_1	s_1s_2	s_2s_0	s_2s_1	s_2s_2
$q_i = P(\sigma_i)$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$

A. $H(S^{(2)}) = \sum_{i=0}^{8} q_i \log_2\left(\frac{1}{q_i}\right) = 3$ bits

$H(S^{(2)}) = 2H(S) = 3\text{bits} \Rightarrow H(S) = 1.5$ bits

A source emits one of four possible symbols during each signaling interval. The symbols occur with the probabilities $p_0 = 0.4$, $p_1 = 0.3$, $p_2 = 0.2$ and $p_3 = 0.1$ which sum to unity as they should. Find the amount of information gained by observing the source emitting each of these symbols.

A. $H(S) = -\sum_{k=0}^{3} P_k \log_2(P_k)$

$= -0.4\log_2 0.4 - 0.3\log_2 0.3 - 0.2\log_2 0.2 - 0.1\log_2 0.1 = 1.846$ bits

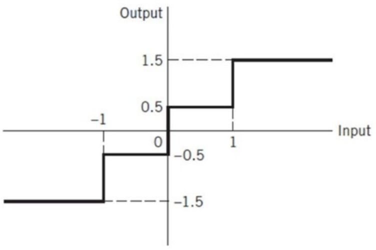
A source emits one of four symbols s_0, s_1, s_2 and s_3 with probabilities 1/3, 1/6, 1/4 and 1/4 respectively. The successive symbols emitted by the source are statistically independent. Calculate the entropy of the source.

A. $H(S) = -\sum_{k=0}^{3} P_k \log_2(P_k)$

$= -\frac{1}{3}\log_2\left(\frac{1}{3}\right) - \frac{1}{6}\log_2\left(\frac{1}{6}\right) - \frac{1}{4}\log_2\left(\frac{1}{4}\right) - \frac{1}{4}\log_2\left(\frac{1}{4}\right) = 1.959$ bits

The sample function of a Gaussian process of zero mean and unit variance is uniformly sampled and then applied to a uniform quantizer having the input–output amplitude characteristic shown in the figure below. Calculate the entropy of the quantizer output.

A.



$H(S) = -\sum_{k=0}^{3} P(y_k) \log_2 P(y_k)$

$P(y_0) = P(y_3) = P(1 < \text{input} < \infty) = \frac{1}{\sqrt{2\pi}} \int_1^\infty \exp\left(\frac{-x^2}{2}\right) dx = Q(1) = 0.15866$

$P(y_1) = P(y_2) = P(0 < \text{input} < 1) = \frac{1}{\sqrt{2\pi}} \int_0^1 \exp\left(-\frac{x^2}{2}\right) dx = Q(0) - Q(1) = 0.5 - 0.15866 = 0.34134$

$H(S) = 1.9$ bits

- Consider a discrete memoryless source with source alphabet $S = \{s_0, s_1, s_2\}$ and source statistics {0.7, 0.15, 0.15}.

a) Calculate the entropy of the source.

b) Calculate the entropy of the second-order extension of the source.

A. $H(S) = -\sum_{k=0}^{2} p_k \log_2(p_k) = -0.7\log_2 0.7 - 2\left(0.15\log_2 0.15\right) = 1.181$ bits

$H(S^{(2)}) = 2 \times H(S) = 2.362$ bits

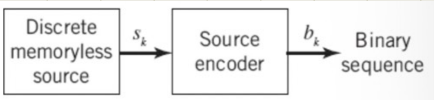
Source Coding Theorem

→ The main issue in communication theory is representation of data generated by discrete source

We use source encoding to solve this issue using a device called source encoder

→ If some symbols are known to be more probable than others, so we exploit this feature,

We assign short codewords to frequent source symbols & long codewords to rare source symbols → Variable-length code



→ The encoder must satisfy 2 requirements :

i) Codewords produced by encoder are in binary form

ii) Source code is uniquely decodable, so original source sequence can be reconstructed

→ Average codeword length $\bar{L}$ of source encoder ⇒ $\bar{L} = \sum_{k=0}^{K-1} p_k l_k$

It denotes average no. of bits per source symbol used in source encoding process

→ Let L_{min} be min. value of L

Then we can coding efficiency as $\eta = \frac{L_{min}}{\bar{L}}$

Since $\bar{L} \geq L_{min}$, $\eta \leq 1$. Which means source encoder is said to be efficient when η approaches 1

Shannon's first theorem : source-coding theorem

→ Given a discrete memoryless source whose output is denoted by random variable S, then entropy $H(S)$ imposes the following bound on $\bar{L}$ ⇒ $\bar{L} \geq H(S)$

→ From this $L_{min} = H(S)$

and we rewrite $\eta = \frac{H(S)}{\bar{L}}$

Lossless data compression Algorithms

→ Usually signals in their natural form contain significant amount of redundant information which is performed by an operation known as lossless data compression.

→ The code from such operation is efficient in terms of $\bar{L}$ & original data can be reconstructed without any loss of data because entropy establishes fundamental limit.

→ We use a technique known as prefix coding

Prefix Coding

→ Code in which no codeword is prefix of any other codeword

Q. Find the right code among I, II & III

Symbol source	Probability of occurrence	Code I	Code II	Code III
s_0	0.5	0	0	0
s_1	0.25	1	10	01
s_2	0.125	00	110	011
s_3	0.125	11	111	0111

A. Code I → Can't be codeword because 0 is codeword for s0 & 00 is codeword for s2

Code II → Can be codeword because it is easily distinguishable

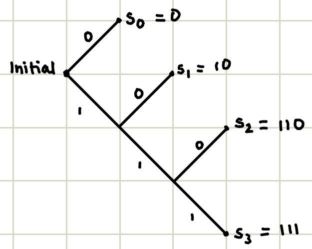
Code III → Can't be codeword because all of them start with 0

Decoding of Prefix Code

Q. Find the decision tree & decode 1011111000

Symbol source	Code II
s_0	0
s_1	10
s_2	110
s_3	111

A.



Code = $\underbrace{10}_{s_1}\underbrace{111}_{s_3}\underbrace{110}_{s_2}\underbrace{0}_{s_0}\underbrace{0}_{s_0}$

Kraft Inequality

→ According to Kraft inequality, codeword lengths must always satisfy the inequality $\sum_{k=0}^{K-1} 2^{-l_k} \leq 1$

→ Kraft inequality is a necessary but not sufficient condition for source code to be prefix code

For ex, In prev q, Code I doesn't satisfy Kraft inequality, and code II & III does but only code II is right

→ Given a discrete memoryless source of entropy $H(S)$, $\bar{L}$ is bounded by $H(S) \leq \bar{L} \leq H(S) + 1$

→ LHS of $H(S) \leq \bar{L} \leq H(S) + 1$ is satisfied under the condition that symbol s_k is emitted by source with probability $p_k = 2^{-l_k}$

l_k : length of codeword assigned to source symbol s_k

So,

$$\sum_{k=0}^{K-1} 2^{-l_k} = \sum_{k=0}^{K-1} p_k = 1$$

For dyadic distribution, $\bar{L} = \sum_{k=0}^{K-1} p_k l_k = \sum_{k=0}^{K-1} \frac{l_k}{2^{l_k}}$

and, $H(S) = -\sum_{k=0}^{K-1} p_k \log_2 p_k$

$$= \sum_{k=0}^{K-1} \left(\frac{1}{2^{l_k}}\right) \log_2 (2^{l_k}) = \sum_{k=0}^{K-1} \frac{l_k}{2^{l_k}} = \bar{L}$$

Hence $H(S) = \bar{L}$

→ For an extended source of order n,

$H(S^{(n)}) \leq \bar{L}_n \leq H(S^{(n)}) + 1$

$nH(S) \leq \bar{L}_n \leq nH(S) + 1$

Let $n \to \infty$, $\lim_{n \to \infty} \frac{\bar{L}_n}{n} = H(S)$

Therefore, by making the order n of an extended prefix source encoder large enough, we can can make the code faithfully represent the discrete memory

Huffman Coding

→ The construction of a simple algorithm that computes an optimal prefix code for a given distribution. (Optimal → Code has shortest expected length)

→ The essence of the algorithm used to synthesize the Huffman code is replace the prescribed set of source statistics of a discrete memoryless source with a simpler one.

→ Procedure :

i) Source symbols are listed in order of decreasing probability. The 2 source symbols of lowest probability are assigned 0 and 1. This step is referred as splitting stage

ii) The 2 source symbols are combined into new source symbol with probability equal to sum of the 2 original probabilities (list is reduced by 1 byte).
The probability of new symbol is placed in the list in accordance to its value.

iii) The procedure is repeated until we are left with final list of source statistics of only two for which symbols 0 and 1 are assigned

Q. Consider 5 symbols as given. Construct Huffman tree & find Huffman Code. Verify Kraft Inequality

Symbols of S	s_0	s_1	s_2	s_3	s_4
Probability p_i	0.4	0.2	0.2	0.1	0.1

A. $H(S) = -\sum_{k=0}^{4} p_k \log_2 p_k = -0.4\log_2 0.4 - 2(0.2\log_2 0.2) - 2(0.1 \log_2 0.1)$

$= 2.121$ bits

Symbol	Stage 1	Stage 2	Stage 3	Stage 4
s_0	0.4	0.4	0.4	0.6 0
s_1	0.2	0.2	0.4 0	0.4 1
s_2	0.2	0.2 0	0.2 1	
s_3	0.1 0	0.2 1		
s_4	0.1 1			

Symbol	Prob.	Code
s_0	0.4	00
s_1	0.2	10
s_2	0.2	11
s_3	0.1	010
s_4	0.1	011

Average code length $\Rightarrow \bar{L} = \sum_{k=0}^{4} p_k l_k = 0.4(2) + 0.2(2) + 0.2(2) + 0.1(3) + 0.1(3) = 2.2$ bits

Code efficiency $\Rightarrow \eta = \frac{H(S)}{\bar{L}} = \frac{2.121}{2.2} = 0.964$

Q. Find Huffman codes & Code efficiency

Symbol	s_0	s_1	s_2	s_3	s_4	s_5	s_6
Probability	0.25	0.25	0.125	0.125	0.125	0.0625	0.0625

A. $H(S) = -\sum_{k=0}^{6} p_k \log_2 p_k = -2(0.25\log_2 0.25) - 3(0.125\log_2 0.125) - 2(0.0625 \log_2 0.0625)$

$= 2.625$ bits

Symbol	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	Stage 6
s_0	0.25	0.25	0.25	0.25	0.5	0.5 0
s_1	0.25	0.25	0.25	0.25	0.25 0	0.5 1
s_2	0.125	0.125	0.25	0.25 0	0.25 1	
s_3	0.125	0.125	0.125 0	0.25 1		
s_4	0.125	0.125 0	0.125 1			
s_5	0.0625 0	0.125 1				
s_6	0.0625 1					

Symbol	Prob.	Code
s_0	0.25	10
s_1	0.25	11
s_2	0.125	001
s_3	0.125	010
s_4	0.125	011
s_5	0.0625	0000
s_6	0.0625	0001

$\bar{L} = \sum_{k=0}^{6} p_k l_k = 2(0.25) + 2(0.25) + 3(0.125) + 3(0.125) + 3(0.125) + 4(0.0625) + 4(0.0625)$

$= 2.625$

$\eta = \frac{H(S)}{\bar{L}} = 1$

Q. Consider discrete memoryless source with alphabet $\{s_0, s_1, s_2\}$ & statistics $\{0.7, 0.15, 0.15\}$. Find Huffman Code

Symbol	Stage 1	Stage 2
s_0	0.7	0.7 0
s_1	0.15 0	0.3 1
s_2	0.15 1	

Symbol	Prob.	Code
s_0	0.7	0
s_1	0.15	10
s_2	0.15	11

$H(S) = -0.7\log_2 0.7 - 2(0.15\log_2 0.15)$

$= 1.183$ bits

$\bar{L} = 0.7(1) + 0.15(2) + 0.15(2) = 1.3$ bits

$\eta = \frac{H(S)}{\bar{L}} = 0.9087$

Q. For prev q, if it was extended to order 2, Find Huffman Codes

Symbol	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	Stage 6	Stage 7	Stage 8
$\sigma_0 = s_0 s_0$	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.51 0
$\sigma_1 = s_0 s_1$	0.105	0.105	0.105	0.105	0.195	0.21	0.3 0	0.49 1
$\sigma_2 = s_0 s_2$	0.105	0.105	0.105	0.105	0.105	0.195 0	0.21 1	
$\sigma_3 = s_1 s_0$	0.105	0.105	0.105	0.105	0.105 0	0.105 1		
$\sigma_4 = s_2 s_0$	0.105	0.105	0.105	0.105 0	0.105 1			
$\sigma_5 = s_1 s_1$	0.0225	0.045	0.045 0	0.09 1				
$\sigma_6 = s_1 s_2$	0.0225	0.0225 0	0.045 1					
$\sigma_7 = s_2 s_1$	0.0225 0	0.0225 1						
$\sigma_8 = s_2 s_2$	0.0225 1							

Symbol	Prob.	Code
σ_0	0.49	1
σ_1	0.105	001
σ_2	0.105	010
σ_3	0.105	011
σ_4	0.105	0000
σ_5	0.0225	000100
σ_6	0.0225	000101
σ_7	0.0225	000110
σ_8	0.0225	000111

$$\bar{L} = \sum_{k=0}^{8} p_k l_k = 0.49(1) + (0.105(3)) \times 3 + 0.105(4) + (0.0225(6)) \times 4 = 2.395 \text{ bits}$$

$$H(S^{(2)}) = 2 \times H(S) = 2 \times 1.183 = 2.366$$

$$\eta = \frac{H(S^{(2)})}{\bar{L}} = 0.9869$$

Discrete Memoryless Channel

→ It is a statistical model with an input X & output Y that is noisy version of X

X, Y → Random Variables

→ Channel is discrete if X & Y have finite sizes

→ Channel is Memoryless when current output symbol depends only on current input symbol & not previous or future symbols.

→ So, assume, $X = \{x_0, x_1, x_2, \dots, x_{J-1}\}$

$Y = \{y_0, y_1, y_2, \dots, y_{K-1}\}$

The carnality of X & Y is defined as the number of elements in the alphabet

So, $p(y_k | x_j) = P(Y = y_k | X = x_j) \quad \in j, k$

and by probability theory $0 \leq p(y_k | x_j) \quad \in j, k$

& $\sum_{k} p(y_k | x_j) = 1$ for fixed j

→ A convenient way of representing the various transition probabilities is in the form of a matrix

$$\vec{P} = \begin{bmatrix} p(y_0|x_0) & p(y_1|x_0) & \cdots & p(y_{K-1}|x_0) \\ p(y_0|x_1) & p(y_1|x_1) & & p(y_{K-1}|x_1) \\ \vdots & \vdots & \cdots & \vdots \\ p(y_0|x_{J-1}) & p(y_1|x_{J-1}) & \cdots & p(y_{K-1}|x_{J-1}) \end{bmatrix}_{J \times K}$$

Channel / Stochiastic Matrix

fixed channel i/p

fixed channel o/p

& sum of all elements in matrix = 1

→ The joint probability distribution of the random variables X and Y given as

$$p(x_j, y_k) = P(X = x_j, Y = y_k) = P(Y = y_k | X = x_j) P(X = x_j) = p(y_k | x_j) p(x_j)$$

→ The marginal probability distribution of output random variable Y is obtained by averaging out the dependance of $p(x_j, y_k)$ on x_j

$$p(y_k) = P(Y = y_k)$$

$$= \sum_{j=0}^{J-1} P(Y = y_k | X = x_j) P(X = x_j)$$

$$= \sum_{j=0}^{J-1} p(y_k | x_j) p(x_j) \quad \in k = 0, 1, \ldots K-1$$

The probabilities of $p(x_j)$ for $j = 0, 1, \ldots, J-1$ are known as prior probabilities

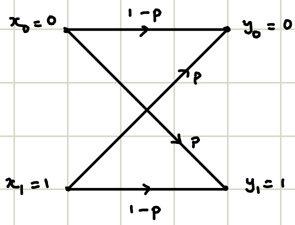
Binary Symmetric Channel

→ It is a special case of the discrete memoryless channel where $J = K = 2$.

Channel matrix is 2×2 matrix

→ Channel has 2 input symbols ($x_0 = 0$, $x_1 = 1$)

& 2 output symbols ($y_0 = 0$, $y_1 = 1$)



$$\vec{P} = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$$

Conditional Entropy

→ Entropy is measure of prior uncertainity about channel input X

Let $H(X | Y = y_k)$ denote conditional entropy,

Then, $H(X | Y = y_k) = \sum_{j=0}^{J-1} p(x_j | y_k) \log_2 \left(\frac{1}{p(x_j | y_k)} \right)$

So expectation on entropy $H(X | Y = y_k)$ is denoted by $H(X | Y)$

→ So, conditional entropy $H(X|Y)$ is the average amount of uncertainity remaining about channel input after channel output has been observed

$$H(X|Y) = \sum_{k=0}^{K-1} p(y_k) H(X | Y = y_k) = \sum_{k=0}^{K-1} \sum_{j=0}^{J-1} p(x_j | y_k) p(y_k) \log_2 \left(\frac{1}{p(x_j | y_k)} \right)$$

$$H(X|Y) = -\sum_{k=0}^{K-1} \sum_{j=0}^{J-1} p(x_j, y_k) \log_2 \left(p(x_j | y_k) \right)$$

Mutual Information

→ Using entropies $H(X)$ and $H(X|Y)$, we define mutual information $I(X;Y) = H(X) - H(X|Y)$

$$I(X;Y) = H(X) - H(X|Y)$$

(or) $I(Y;X) = H(Y) - H(Y|X)$

→ It is the measure of uncertainity about channel input which is resolved by observing channel output

(or)

It is the measure of uncertainity about channel output which is resolved by sending channel input

$\rightarrow \quad H(Y) = -\sum_{k=0}^{K-1} p(y_k) \log_2 (p(y_k))$

$$H(Y|X=x_j) = -\sum_{k=0}^{K-1} p(y_k|x_j) \log_2 (p(y_k|x_j))$$

$$= -\sum_{k=0}^{K-1} p(y_k|x_j)\, H(Y|X=x_j)$$

$$= -\sum_{j=0}^{J-1} \sum_{k=0}^{K-1} p(y_k|x_j)\, p(x_j) \log_2 p(y_k|x_j)$$

$$= -\sum_{j=0}^{J-1} \sum_{k=0}^{K-1} p(x_j|y_k) \log_2 p(y_k|x_j)$$

$\rightarrow$ Properties of Mutual Information

i) Symmetry $\Rightarrow \; I(X;Y) = I(Y;X)$

Proof $\Rightarrow \quad I(Y;X) = H(Y) - H(Y|X)$

$$= -\sum_{k=0}^{K-1} p(y_k) \log_2 p(y_k) + \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} p(x_j, y_k) \log_2 (p(x_j|y_k))$$

$$= -\sum_{k=0}^{K-1} p(y_k) \log_2 p(y_k) + \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} p(x_j, y_k) \log_2 \frac{p(x_j|y_k)\, p(y_k)}{p(x_j)} \qquad \left(\text{Bayes' Rule : } p(x_j)p(y_k|x_j) = p(x_j|y_k)p(y_k)\right)$$

$$= -\sum_{k=0}^{K-1} p(y_k) \log_2 p(y_k) - \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} p(x_j, y_k) \log_2 p(x_j) + \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} p(x_j, y_k) \log_2 p(x_j|y_k)$$

$$+ \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} p(x_j, y_k) \log_2 p(y_k)$$

$$= \cancel{H(Y)} + H(X) - H(X|Y) - \cancel{H(Y)}$$

$$= H(X) - H(X|Y) = I(X;Y)$$

ii) Non-negative $\Rightarrow \; I(X;Y) \geq 0$

$$I(X;Y) = H(X) - H(X|Y) = \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} p(x_j, y_k) \log_2 \frac{p(x_j|y_k)}{p(x_j)}$$

$$= \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} p(x_j, y_k) \log_2 \frac{p(x_j, y_k)}{p(x_j)p(y_k)} \qquad \left(\text{Bayes' Rule: } p(x_j|y_k) = \frac{p(x_j, y_k)}{p(x_j)}\right)$$

So it can be interpreted as relative entropy b/w 2 sources one following $p(x_j, y_k)$ & other following $p(x_j)p(y_k)$

Since relative entropy is non-negative, $I(X;Y)$ is non-negative

Proof $\Rightarrow \quad H(X,Y) = \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} p(x_j, y_k) \log_2 \left(\frac{1}{p(x_j, y_k)}\right)$

$$= \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} p(x_j, y_k) \log_2 \left(\frac{p(x_j)p(y_k)}{p(x_j, y_k)}\right) + \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} p(x_j, y_k) \log_2 \left(\frac{1}{p(x_j)p(y_k)}\right)$$

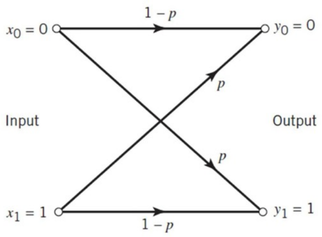
$-I(X;Y)$

$$\sum_{j=0}^{J-1} \sum_{k=0}^{K-1} p(x_j, y_k) \log_2 \left(\frac{1}{p(x_j)}\right) + \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} p(x_j, y_k) \log_2 \left(\frac{1}{p(y_k)}\right)$$

From this, $\quad H(X,Y) = -I(X;Y) + H(X) + H(Y) \qquad = H(X) + H(Y)$

$$I(X;Y) = H(X) + H(Y) - H(X,Y)$$

Q. Consider the transition probability diagram of a binary symmetric channel shown below. The input binary symbols 0 and 1 occur with equal probability. Find the probabilities of the binary symbols 0 and 1 appearing at the channel output.



A. Given, they have equal probabilities, $p(x_0) = p(x_1) = 0.5$

For binary symmetric channel $p(y_0|x_0) = p(y_1|x_1) = 1-p$

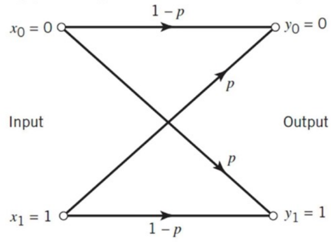
$p(y_0|x_1) = p(y_1|x_0) = p$

Channel output probabilities : $p(y_k) = \sum_{j=0}^{1} p(y_k|x_j)\, p(x_j)$ for $k = 0, 1$

$p(y_0) = p(y_0|x_0)\, p(x_0) + p(y_0|x_1)\, p(x_1) = \frac{1-p}{2} + \frac{p}{2} = 0.5$

$p(y_1) = p(y_1|x_0)\, p(x_0) + p(y_1|x_1)\, p(x_1) = \frac{p}{2} + \frac{1-p}{2} = 0.5$

Q. Consider the transition probability diagram of a binary symmetric channel shown below. The input binary symbols 0 and 1 occur with probability ¼ and ¾ respectively. Find the probabilities of the binary symbols 0 and 1 appearing at the channel output.



A. $p(x_0) = \frac{1}{4}$ $p(x_1) = \frac{3}{4}$

$p(y_0|x_0) = p(y_1|x_1) = 1-p$

$p(y_0|x_1) = p(y_1|x_0) = p$

Channel o/p probabilities are given by

$p(y_0) = \sum_{j=0}^{1} p(y_0|x_j)\, p(x_j) = p(y_0|x_0)\, p(x_0) + p(y_0|x_1)\, p(x_1)$

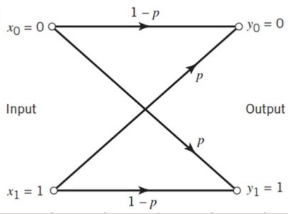
$= \frac{1-p}{4} + \frac{3p}{4} = \frac{1+2p}{4}$

$p(y_1) = \sum_{j=0}^{1} p(y_1|x_j)\, p(x_j) = p(y_1|x_0)\, p(x_0) + p(y_1|x_1)\, p(x_1)$

$= \frac{p}{4} + \frac{3(1-p)}{4} = \frac{3-2p}{4}$

Q.
- Let the source probabilities $p(x_0) = 1/4$ and $p(x_1) = 3/4$.

  Suppose the discrete memoryless channel, depicted below, has the conditional probability of error $p = 1/4$. Find the following:
- Entropies $H(X)$, $H(Y)$, $H(X|Y)$ and $H(Y|X)$
- Mutual information of the channel $I(X;Y)$
- Mutual information of the channel $I(Y;X)$
- Find the joint entropy $H(X,Y)$



A. $p(x_0) = 1/4$ $p(x_1) = 3/4$ $p(y_0) = \frac{1+2p}{4} = 0.375$ $p(y_1) = \frac{3-2p}{4} = 0.625$

$p(y_0|x_1) = p(y_1|x_0) = p = 1/4$ & $p(y_0|x_0) = p(y_1|x_1) = 1-p = 3/4$

$H(X) = \sum_{j=0}^{1} p(x_j) \log_2 \frac{1}{p(x_j)} = -0.25\log_2(0.25) - 0.75\log_2(0.75) = 0.811$ bits/symbol

$H(Y) = \sum_{k=0}^{1} p(y_k) \log_2 \frac{1}{p(y_k)} = -0.375\log_2(0.375) - 0.625\log_2(0.625) = 0.954$ bits/symbol

$H(Y|X) = -\sum_{j=0}^{1}\sum_{k=0}^{1} p(x_j, y_k) \log_2 (p(y_k|x_j))$

$= -\sum_{j=0}^{1}\sum_{k=0}^{1} p(y_k|x_j)\, p(x_j) \log_2 p(y_k|x_j)$

$= 0.811$ bits/symbol

$H(X|Y) = -\sum_{j=0}^{1}\sum_{k=0}^{1} p(x_j, y_k) \log_2 p(x_j|y_k)$

$= -\sum_{j=0}^{1}\sum_{k=0}^{1} p(y_k|x_j)\, p(x_j) \log_2 p(x_j|y_k)$

$= -\sum_{j=0}^{1}\sum_{k=0}^{1} p(y_k|x_j)\, p(x_j) \log_2\left(\frac{p(y_k|x_j)\, p(x_j)}{p(y_k)}\right)$

$= 0.668$ bits/symbol

$I(X;Y) = H(X) - H(X|Y) = 0.811 - 0.668 = 0.143$ bits/symbol

$I(Y;X) = H(Y) - H(Y|X) = 0.954 - 0.811 = 0.143$ bits/symbol

Now $I(X;Y) = H(X) + H(Y) - H(X,Y)$

$H(X,Y) = H(X) + H(Y) - I(X;Y) = 0.811 + 0.954 - 0.143 = 1.622$ bits/symbol

Channel Capacity

→ The intrinsic ability of the channel to convey information

→ Consider a discrete memoryless channel with transition probabilities $p(x_j|y_k)$ where x_j & y_k denote the channel input & channel output respectively

Then mutual information is $I(X;Y) = \sum_{k=0}^{K-1} \sum_{j=0}^{J-1} p(x_j, y_k) \log_2\left(\frac{p(y_k|x_j)}{p(y_k)}\right)$

where $p(x_j|y_k) = p(y_k|x_j)\, p(x_j)$

& $p(y_k) = \sum_{j=0}^{J-1} p(y_k|x_j)\, p(x_j)$

From these 3 eqs, $I(X;Y) = \sum_{k=0}^{K-1} \sum_{j=0}^{J-1} p(y_k|x_j)\, p(x_j) \log_2\left(\frac{p(y_k|x_j)}{\sum_{j=0}^{J-1} p(y_k|x_j)\, p(x_j)}\right)$

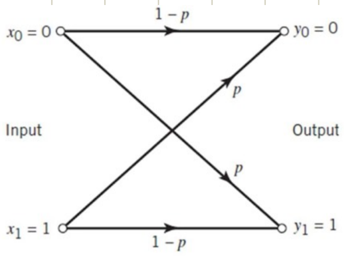
→ Given the transition probabilities,

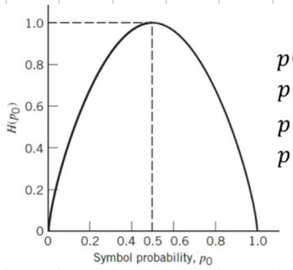
Then Channel Capacity is given by $C = \max I(X;Y)$

subject to : $p(x_j) \geq 0$ & $\sum_{j=0}^{J-1} p(x_j) = 1$

→ So, Channel Capacity of a discrete memoryless channel is defined as maximum mutual information $I(X;Y)$ in any single use of the channel (i.e, signaling interval) where maximization is over all possible input probability distributions $\{p(x_j)\}$ on X

ex: Channel Capacity of binary symmetric channel



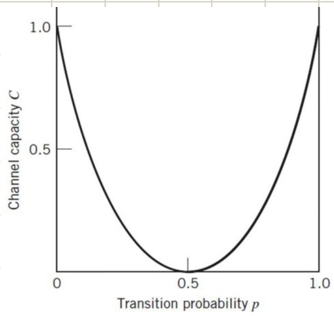


$p(y_0|x_0) = 1 - p$
$p(y_1|x_1) = 1 - p$
$p(y_0|x_1) = p$
$p(y_1|x_0) = p$

Entropy of source emitting symbols with Bernoulli distribution achieved $H(X) = 1$ bit when $p(x_0) = p(x_1) = 0.5$

$C = I(X;Y)\big|_{p(x_0)=p(x_1)=0.5} = 1 + p\log_2 p + (1-p)\log_2(1-p)$

$= 1 - H(p)$



→ Channel capacity is convex function of p

→ Channel capacity attains max value of 1 bit when channel is noise free

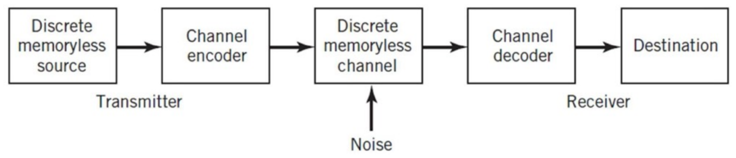
→ channel capacity attains min value of 0 bits when probability of error is ½

Channel Encoder & Channel Decoder

→ Encoder maps input data sequence to channel input sequence

→ Decoder maps channel output sequence to output data sequence

→ They ensure impact of channel noise on digital communication system is minimized



→ Encoder introduces redundancy in a controlled manner so it is possible to reconstruct original sequence as accurately as possible

→ In block codes, message is divided into k-bit blocks & each k-bit block is mapped to a n-bit code $(n > k)$

$r = \frac{k}{n}$ (r : code rate)

$\varepsilon = n - r$ (ε : redundancy introduced to achieve low avg. probability of symbol error)

→ If $\frac{H(S)}{T_s} \leq \frac{C}{T_c}$,

avg. information rate → Channel Capacity per unit time

There is a coding scheme using which source output can be retransmitted over a noisy channel & reconstructed with an arbitrarily small probability of error

where,

$H(S)$: Source Entropy

T_s : Signaling rate of source

C : Channel Capacity

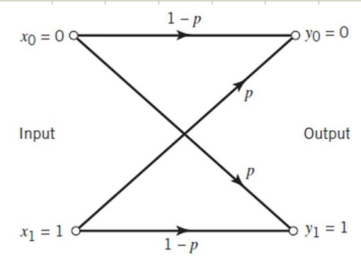
T_c : Channel usage period

Limitations of Channel coding theorem

i) Channel coding theorem doesn't show how to construct good code

ii) Theorem doesn't have precise result for probability of symbol error decoding channel output

Q. Given the following binary symmetric channel & Bernoulli source, Find min value of C



A. Condition $\frac{H(S)}{T_s} \leq \frac{C}{T_c}$

For Bernoulli, $H(S) = 1$, $p(x_0) = p(x_1) = 0.5$

So, $\frac{C}{T_c} \geqslant \frac{1}{T_s} \Rightarrow C \geqslant \frac{T_c}{T_s} = r$

r : code rate of channel encoder

Differential Entropy

→ Consider continuous random variable X with pdf $f_X(x)$

Then differential entropy is $h(X) = \int_{-\infty}^{\infty} f_X(x) \log_2\left(\frac{1}{f_X(x)}\right) dx$

It differs from absolute entropy $H(X)$ by a factor $-\log_2 \Delta x$

Δx : infinitesimal range of random variable X

→ For multi-variate random variable $\vec{X} = [X_1, \dots, X_n]$

$$h(X) = \int_{-\infty}^{\infty} f_X(x) \log_2\left(\frac{1}{f_{\vec{X}}(\vec{x})}\right) d\vec{x}$$

$f_{\vec{X}}(\vec{x})$: joint pdf of $\vec{X}$

Mutual Information for continuous ensembles

→ Mutual information b/w a pair of continuous random variables X and Y is

$$I(X;Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y) \log_2\left(\frac{f_X(x|y)}{f_X(x)}\right) dx\,dy$$

$f_{X,Y}(x,y)$: joint pdf of X and Y

$f_X(x|y)$: conditional pdf of X given $Y=y$

→ Using Bayes' rule, $f_X(x|y) = \frac{f_{X,Y}(x,y)}{f_X(x) f_Y(y)} dx\,dy$

→ Properties of mutual information for continuous ensembles

i) $I(X;Y) = I(Y;X)$

ii) $I(X;Y) \geq 0$

iii) $I(X;Y) = h(X) - h(X|Y)$

iv) $I(Y;X) = h(Y) - h(Y|X)$

→ $h(X|Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y) \log_2\left(\frac{1}{f_X(x|y)}\right) dx\,dy$

$h(Y|X) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y) \log_2\left(\frac{1}{f_Y(y|x)}\right) dx\,dy$

→ Conditional differential entropies of X given Y & Y given X

Uniform Random Variables

→ Differential entropy of URV $X \in [0,a]$

$$h(X) = \int_{-\infty}^{\infty} f_X(x) \log_2\left(\frac{1}{f_X(x)}\right) dx \qquad \left(f_X(x) = \begin{cases} 1/a, & 0 < x < a \\ 0, & \text{otherwise} \end{cases} \right)$$

$$= \int_0^a \frac{1}{a} \log_2 a \, dx = \log_2 a$$

Gaussian Random Variables

→ Differential entropy of GRV X

$$h(x) = \int_{-\infty}^{\infty} f_X(x) \log_2\left(\frac{1}{f_X(x)}\right) dx \qquad \left(f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right)$$

$$= -\log_2 e \int_{-\infty}^{\infty} f_X(x) \ln(f_X(x))\, dx$$

$$= -\log_2 e \left(-\int_{-\infty}^{\infty} \frac{(x-\mu)^2}{2\sigma^2} f_X(x)\,dx - \int_{-\infty}^{\infty} f_X(x) \ln(\sqrt{2\pi}\sigma)\,dx \right)$$

$$= -\log_2 e \left(\frac{-1}{2} - \ln(\sqrt{2\pi}\sigma) \right)$$

$$= \log_2 \sqrt{e} + \log_2 \sqrt{2\pi}\,\sigma$$

$$= \frac{1}{2} \log_2 (2\pi e \sigma^2)$$

Relative Entropy of Continuous Ensembles

→ Relative entropy of a R.V having distributions $f_X(x)$ and $f_Y(x)$,

$$D(f_Y \| f_X) = \int_{-\infty}^{\infty} f_Y(x) \log_2\left(\frac{f_Y(x)}{f_X(x)}\right) dx \geq 0$$

$$\Rightarrow \int_{-\infty}^{\infty} f_Y(x) \log_2\left(\frac{1}{f_Y(x)}\right) dx \leq \int_{-\infty}^{\infty} f_Y(x) \log_2\left(\frac{1}{f_X(x)}\right) dx$$

$$\Rightarrow \qquad h(Y) \leq \int_{-\infty}^{\infty} f_Y(x) \log_2\left(\frac{1}{f_X(x)}\right) dx$$

Bound on differentiable entropy

→ Consider RV's X & Y with mean μ & variance σ^2

→ Suppose X is Gaussian R.V

$$h(Y) \leq \int_{-\infty}^{\infty} f_Y(x) \log_2\left(\frac{1}{f_X(x)}\right) dx$$

$$\leq -\log_e 2 \int_{-\infty}^{\infty} f_Y(x) \ln f_X(x)\, dx$$

$$\leq -\log_e 2 \int_{-\infty}^{\infty} f_Y(x) \left(-\frac{(x-\mu)^2}{2\sigma^2} - \ln(\sqrt{2\pi}\,\sigma) \right) dx \qquad \left(\begin{aligned} &\int_{-\infty}^{\infty} f_Y(x)\,dx = 1 \\ &\int_{-\infty}^{\infty} (x-\mu)^2 f_Y(x)\,dx = \sigma^2 \end{aligned} \right)$$

$$\leq \frac{1}{2} \log_2 (2\pi e \sigma^2)$$

$$\leq h(X)$$

→ Observations :

i) For any finite variance, GRV has largest differential entropy

ii) Entropy of GRV is uniquely determined by its variance

iii) $h(Y) = h(X)$ if Y is also normally distributed random variable

AWGN Channel (Additive Gaussian White Noise)

→ Consider a discrete time memoryless, band-limited, power-limited Gaussian channel

$$\text{(Input) } X_K \longrightarrow \textcircled{\Sigma} \longrightarrow Y_K \text{ (Output)}$$

$$\uparrow$$

Noise N_K

→ $x(t)$ is limited to 'B' Hz & produces '2B' samples per second

→ Samples are transmitted over a band-limited noisy channel every T seconds
↳ 'B' Hz

→ Then total samples in T seconds is $K = 2BT$

→ Properties

i) N_K is zero-mean with PSD of $N_0/2$

ii) Noise is limited to 'B' Hz

iii) Noise power $\sigma^2 = N_0 B$

iv) Channel output is sampled every T seconds.

Sample $Y_K = X_K + N_K$

v) Average power of the samples $E[X_K^2]$ is limited by transmit power P

Information Capacity Law

→ $C = \max\limits_{f_{X_K}(x)} I(X_K ; Y_K) = h(Y_K) - h(Y_K | X_K)$

Subject to $E[X_K^2] \leq P$

→ Since $\underset{\text{Input}}{\underset{\downarrow}{X_K}}$ & $\underset{\text{Noise}}{\underset{\downarrow}{N_K}}$ are independant, $h(Y_K | X_K) = h(N_K)$

So, $I(X_K ; Y_K) = h(Y_K) - h(N_K)$

→ Max value of differential entropy occurs for GRY

Therefore, $\max(I(X_K, Y_K))$ occurs when input samples X_K are also noise-like Gaussian distributed samples with average power P

So, $C = \frac{1}{2} \log_2\left(1 + \frac{P}{\sigma^2}\right)$ bits per channel use

→ Noise power $\sigma^2 = N_0 B$

One channel use for each symbol X_K takes $\frac{K}{T}$ seconds

Then, $C = B \log_2\left(1 + \frac{P}{N_0 B}\right)$ bits per second

↳ Shannon's Information Capacity Law

Q.

- A voice-grade channel of the telephone network has a bandwidth of 3.4 kHz.

a. Calculate the information capacity of the telephone channel for a signal-to-noise ratio of 30 dB.

b. Calculate the minimum signal-to-noise ratio required to support information transmission through the telephone channel at the rate of 9600 bits/s.

A. a) $B = 3.4$ KHz

Signal-to-Noise Ratio (SNR) = 30 dB $\hookrightarrow 10^{30/10} = 1000$

$C = B \log_2 (1 + SNR) = 3400 \times \log_2(1001) = 33.89$ Kbps

b) $C = 9600$ bps

min SNR $= 2^{C/B} - 1 = 2^{\frac{9600}{3400}} - 1 = 7.079 - 1 = 6.079$ or 7.84 dB

Q.

- Alphanumeric data are entered into a computer from a remote terminal through a voice-grade telephone channel. The channel has a bandwidth of 3.4 kHz and output signal-to-noise ratio of 20 dB. The terminal has a total of 128 symbols. Assume that the symbols are equiprobable and the successive transmissions are statistically independent.

a. Calculate the information capacity of the channel.

b. Calculate the maximum symbol rate for which error-free transmission over the channel is possible.

A. a) $B = 3400$ Hz $\qquad$ SNR = 20 dB $= 10^{20/10} = 100$

$P_K = \frac{1}{128}$ for $K = 0, 1, \dots 127$

$C = B \log_2 (1 + SNR)$

$= 3400 \log_2 (1 + 100) = 22.638$ Kbps

b) Max transmission rate for error free transmission $\Rightarrow R_b < C$

$\downarrow$ Bit Rate

$H(x) = \sum_{K=0}^{127} P_K \log_2\left(\frac{1}{P_K}\right) = \sum_{K=0}^{127} \frac{1}{128} \log_2(128) = 128 \times \frac{7}{128} = 7$ bits/symbol

Max Symbols rate $= \frac{C}{H(x)} = \frac{22.638 \times 10^3}{7} = 3.234$ K baud $\hookrightarrow$ (symbols per sec)

So, $R_s < 3.324$ Kbaud

Q.

- A black-and-white television picture may be viewed as consisting of approximately 3×10^5 elements, each of which may occupy one of 10 distinct brightness levels with equal probability. Assume that (1) the rate of transmission is 30 picture frames per second and (2) the signal-to-noise ratio is 30 dB.
- Using the information capacity law, calculate the minimum bandwidth required to support the transmission of the resulting video signal.

A. Given, $P_K = \frac{1}{10}$ for $K = 0, 1, \dots 9$

$H(x) = \sum_{K=0}^{9} P_K \log_2\left(\frac{1}{P_K}\right) = \sum_{K=0}^{9} \frac{1}{10} \log_2(10) = 3.32$ bits/element

Source transmits 30 FPS

So, R_b = No. of frames × No. of elements per frame × No. of bits per element $= 30 \times 3 \times 10^5 \times 3.32 = 29.9$ Mbps

Continuation,

$C = B \log_2 (1 + SNR)$ & $R_b \leq C$

$R_b \leq B \log_2 (1 + SNR) \Rightarrow B \geq \frac{R_b}{\log_2(1 + SNR)}$

SNR = 30 dB $\Rightarrow 1000$

$B_{min} = \frac{29.9 \times 10^6}{\log_2(1001)} = 3$ MHz