According to New Revised Credit System Syllabus

SPPU

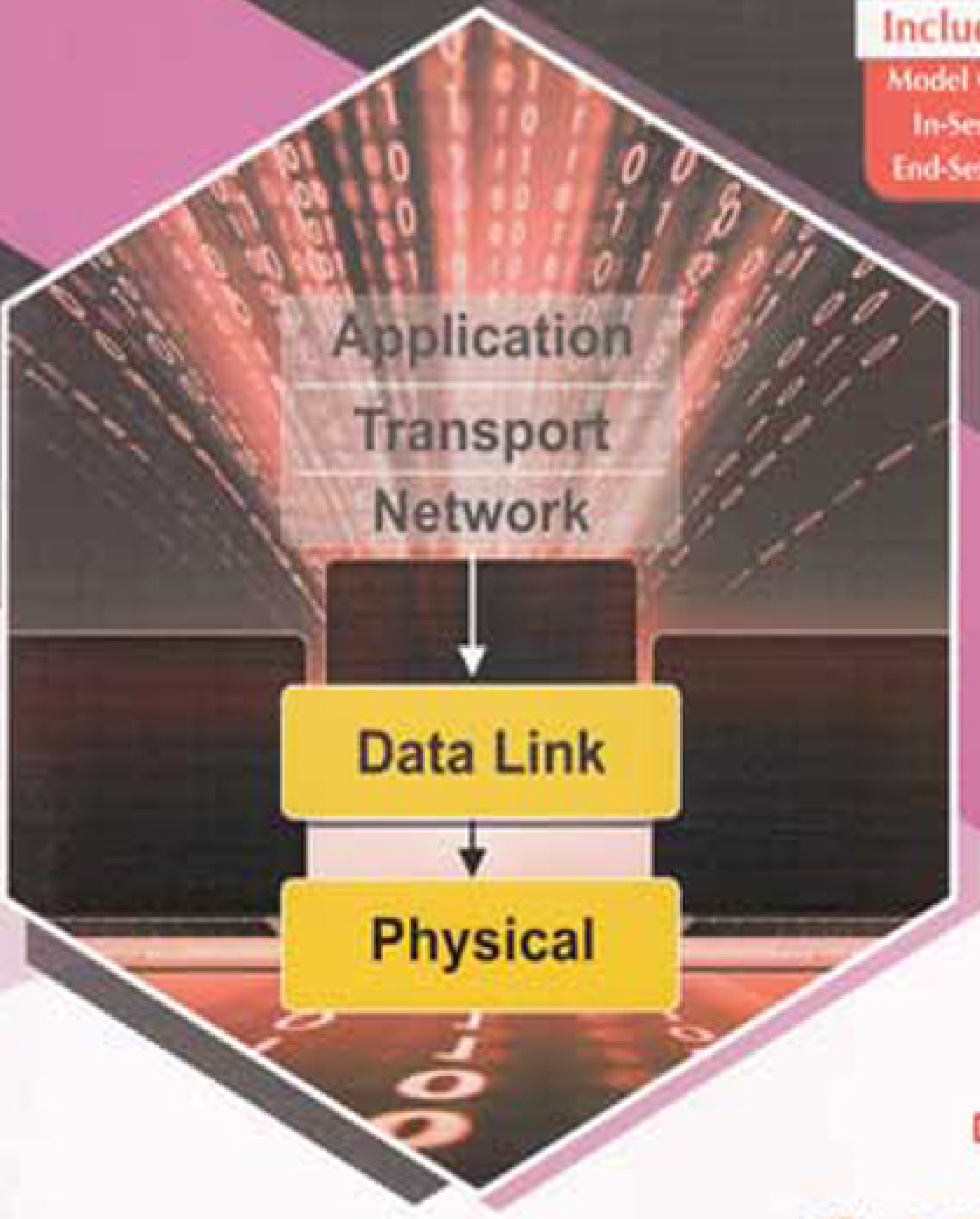
Third Year Degree Course In

ELECTRONICS AND TELECOMMUNICATION ENGG. (Semester - II)

INFORMATION THEORY, CODING AND COMMUNICATION NETWORKS

Includes

Model Question Papers for In-Semster (30 Marks) & End-Sem Exams. (70 Marks)



Dr. G. R. PATIL

NIRALI PRAKASHAN
ADVANCEMENT OF KNOWLEDGE

www.pragationline.com
www.facebook.com/niralibooks

A TEXT BOOK OF

INFORMATION THEORY, CODING AND COMMUNICATION NETWORKS

FOR
SEMESTER – II

THIRD YEAR DEGREE COURSE IN
ELECTRONICS & TELECOMMUNICATION ENGINEERING

Strictly According to New Revised Credit System Syllabus of Savitribai Phule Pune University
(w.e.f. June 2017)

Dr. G. R. PATIL
M.E. (Electronics), Ph. D. (E&TC)
Professor & Head, E & TC Department,
Army Institute of Technology,
Dighi, PUNE

Price ₹ 240.00

NIRALI PRAKASHAN
ADVANCEMENT OF KNOWLEDGE

N4207

ITC & CN (TE SEM. II, E & TC) **ISBN :978-93-87397-56-9**

First Edition : January 2018

Published By :
NIRALI PRAKASHAN
Abhyudaya Pragati, 1312, Shivaji Nagar,
Off J.M. Road, Pune – 411005
Tel - (020) 25512336/37/39, Fax - (020) 25511379
Email : niralipune@pragationline.com

Polyplate

Printed By :
YOGIRAJ PRINTERS AND BINDERS
Survey No. 10/1A, Ghule Industrial Estate
Nanded Gaon Road
Nanded, Pune - 411041
Mobile No. 9404233041/9850046517

☞ DISTRIBUTION CENTRES

PUNE

Nirali Prakashan : 119, Budhwar Peth, Jogeshwari Mandir Lane, Pune 411002, Maharashtra
Tel : (020) 2445 2044, 66022708, Fax : (020) 2445 1538
Email : bookorder@pragationline.com, niralilocal@pragationline.com

Nirali Prakashan : S. No. 28/27, Dhyari, Near Pari Company, Pune 411041
Tel : (020) 24690204 Fax : (020) 24690316
Email : dhyari@pragationline.com, bookorder@pragationline.com

MUMBAI

Nirali Prakashan : 385, S.V.P. Road, Rasdhara Co-op. Hsg. Society Ltd.,
Girgaum, Mumbai 400004, Maharashtra
Tel : (022) 2385 6339 / 2386 9976, Fax : (022) 2386 9976
Email : niralimumbai@pragationline.com

☞ DISTRIBUTION BRANCHES

JALGAON

Nirali Prakashan : 34, V. V. Golani Market, Navi Peth, Jalgaon 425001,
Maharashtra, Tel : (0257) 222 0395, Mob : 94234 91860

KOLHAPUR

Nirali Prakashan : New Mahadvar Road, Kedar Plaza, 1st Floor Opp. IDBI Bank
Kolhapur 416 012, Maharashtra. Mob : 9850046155

NAGPUR

Pratibha Book Distributors : Above Maratha Mandir, Shop No. 3, First Floor,
Rani Jhanshi Square, Sitabuldi, Nagpur 440012, Maharashtra
Tel : (0712) 254 7129

DELHI

Nirali Prakashan : 4593/21, Basement, Aggarwal Lane 15, Ansari Road, Daryaganj
Near Times of India Building, New Delhi 110002
Mob : 08505972553

BENGALURU

Pragati Book House : House No. 1, Sanjeevappa Lane, Avenue Road Cross,
Opp. Rice Church, Bengaluru – 560002.
Tel : (080) 64513344, 64513355,Mob : 9880582331, 9845021552
Email:bharatsavla@yahoo.com

CHENNAI

Pragati Books : 9/1, Montieth Road, Behind Taas Mahal, Egmore,
Chennai 600008 Tamil Nadu, Tel : (044) 6518 3535,
Mob : 94440 01782 / 98450 21552 / 98805 82331,
Email : bharatsavla@yahoo.com

niralipune@pragationline.com | www.pragationline.com
Also find us on www.facebook.com/niralibooks

PREFACE

It gives me great pleasure to present the book **'Information Theory, Coding and Communication Networks'** for the students of Third Year Degree Course in Electronics & Telecommunication Engineering of the Savitribai Phule Pune University. This book is strictly as per the **New Revised Credit System Syllabus 2015** Pattern with effect from the Academic Year 2017-18.

As per New Revised Examination Scheme which has been implemented from this academic year, In-semester assessment carries 30 marks over first three units and End Semester Examination carries 70 marks over entire syllabus out of which first three units will carry 20 marks and units 4, 5, 6 will carry 50 marks. The theory course will have 4 credits.

The book is written such that all the basic concepts are explained in simplified manner. It is presented in a more conceptual manner rather than mathematical, as required by the new examination system. It is my objective to keep the presentation systematic, consistent, intensive and clear through explanatory notes and figures.

Main feature of this book is, Complete Coverage of the New Credit System Syllabus with large number of Worked Solved Examples, Exercises, Model Question Papers of In Sem. and End Sem. Exams.

I am sure that this book will cater to all needs of students for this subject.

I also take this opportunity to express my sincere thanks to Shri. Dineshbhai Furia, Shri. Jignesh Furia, Mrs. Nirali Verma, Shri. M. P. Munde and entire team of Nirali Prakashan namely Mrs. Deepali Lachake (Co-ordinator), who really have taken keen interest and untiring efforts in publishing this text.

The advice and suggestions of our esteemed readers to improve the text are most welcomed, and will be highly appreciated.

Pune **Author**

SYLLABUS

Unit I : Information Theory & Source Coding **(6 Hrs)**

Introduction to information theory, Entropy and its properties, Source coding theorem, Huffman coding, Shannon-Fano coding, The Lempel Ziv algorithm, Run Length Encoding, Discrete memory less channel, Mutual information, Examples of Source coding-Audio and Video Compression.

Unit II : Information Capacity & Channel Coding **(8 Hrs)**

Channel capacity, Channel coding theorem, Differential entropy and mutual Information for continuous ensembles, Information Capacity theorem, Linear Block Codes:Syndrome and error detection, Error detection and correction capability, Standard array and syndrome decoding, Encoding and decoding circuit, Single parity check codes, Repetition codes and dual codes, Hamming code, Golay Code, Interleaved code.

Unit III : Cyclic Codes **(8 Hrs)**

Galois field, Primitive element & Primitive polynomial, Minimal polynomial and generator polynomial, Description of Cyclic Codes, Generator matrix for systematic cyclic code, Encoding for cyclic code, Syndrome decoding of cyclic codes, Circuit implementation of cyclic code.

Unit IV : BCH and Convolutional Codes **(7 Hrs)**

Binary BCH code,Generator polynomial for BCH code, Decoding of BCH code, RS codes, generator polynomial for RS code, Decoding of RS codes, Cyclic Hamming code and Golay code. Introduction of convolution code, State diagram, Tree diagram, Trellis diagram, Sequential decoding and Viterbi decoding

Unit V : Data Communication & Physical Layer **(7 Hrs)**

Data Communications – Networks - Network models – OSI model – Layers in OSI model – TCP / IP protocol suite – Addressing – Guided and Unguided Transmission media.

Unit VI : Data Link Layer **(7Hrs)**

Data link control: Framing – Flow and error control –Protocols for Noiseless and Noisy Channels – HDLC.

CONTENTS

Unit I

UNIT - I

INFORMATION THEORY AND SOURCE CODING

1.1 INTRODUCTION TO INFORMATION THEORY

Information and Communication Technology (ICT) has become backbone of today's economy. Communication systems transmit the information generated by source to some destination. This information can be voice, electronic mail, computer generated data, video, fax, etc. Communication systems, through which this information is passed, are likely to distort this information. There is uncertainty at transmitter, channel and receiver in the communication system. Performance of a communication system must, therefore, be analyzed from the point of view of information transmitted and processed in the system. This will require use of probabilistic techniques because of uncertainty involved.

C. E. Shannon in 1948 published his revolutionary work which laid the foundation of "Information Theory". He stated that we can transmit information through a channel at any rate less than the capacity of channel with arbitrarily small error, whatever may be the disturbances in the channel.

Information theory deals with

(i) Mathematical laws governing systems that communicate and process information.

(ii) Quantitative measures of information and capacity of various systems that transmit, receive or process information.

We have to come out with mathematical models of source which is generating information, channel which is transmitting information and receiver which is receiving information. Measurement of information involved in each stage also needs to be done so that we can design an effective and efficient communication system.

To summarize, information theory is used for mathematical modelling and analysis of communication system, so that we will be able to transmit minimum information and maximize the transmission rate, yet there is reliability in transmission.

1.1.1 Mathematical Model for Information Sources

There are two types of information sources

(i) Analog sources such as voice, video etc.

(ii) Discrete sources such as computer generated data, PCM output etc.

Information produced by any source is random. Consider a discrete information source with m different letters or digits given by $\{x_1, x_2, x_3, \ldots x_m\}$. It emits a sequence of letters selected from the set of letters. To construct mathematical model, we assume that each letter has a probability

$$p(x_j) = p(X = x_j) \quad ; \quad 1 \le j \le m$$

where,

$$\sum_{j=1}^{m} p(x_j) = 1$$

Discrete Memoryless Source (DMS)

If the current letter produced by a source is statistically independent of all past and future outputs, then such source is called Discrete Memoryless Source (DMS).

Discrete Stationary Source (DSS)

If the joint probability of two sequences of length n say $(x_1, x_2, x_3, \ldots x_m\}$ and $\{x_{1+n}, x_{2+n}, \ldots x_{m+n}\}$ are identical for all $n \ge 1$ and for all shifts m, then such source is said to be stationary source.

1.2 UNCERTAINTY AND INFORMATION

The most important feature of any communication system is its uncertainty. Information produced by the source is random in nature. The transmitter, channel and receiver which carry this information, are also statistical in nature. The performance of communication system can be best described by modelling it statistically. Let us first understand how information is measured. Consider the following three statements.

1. Today the sun will set in the West. 2. It will rain somewhere in India tomorrow. 3. Man bites dog.

The amount of information contained in each of above three statements is different. The first statement conveys no information at all, or it is certain to happen. The second statement conveys some information, whereas the last statement carries lot of information as the probability of man biting dog is very very small. It means there is an inverse relationship between probability of an event and amount of information carried by it. Thus, information is an inverse function of probability of occurrence of the event. If we denote $I(x_j)$ as information associated with event x_j, we can write,

$$I(x_j) = f\left[\frac{1}{p(x_j)}\right] \quad \text{... (1.1)}$$

1. $p(x_j) = 0$, $I(x_j)$ should be highest, say ∞.
2. $p(x_j) = 1$, $I(x_j)$ should be lowest, say, 0.
3. If $p(x_j) < p(x_k)$ then $I(x_j) > I(x_k)$
4. If x_j and y_k are statistically independent,

   Then, $$I(x_j, y_k) = I(x_j) + I(y_k)$$

   i.e. $$I(x_j, y_k) = f\left[\frac{1}{p(x_j) \cdot p(y_k)}\right]$$

   Logarithmic function can give such results.

   Hence,

Definition

$$I(x_j) = \log\left[\frac{1}{p(x_j)}\right] = -\log p(x_j) \quad \text{... (1.2)}$$

Thus, for discrete source X producing outputs $x_1, x_2, x_3, \ldots x_m$, information of an event x_j is given by equation (1.2). It is also called self information. Unit of information depends on base of logarithm.

If base is 2, unit is bit. If base is 10, unit is decit or Hartley and if base is e, unit is nat. Most of the time we deal with binary information, hence base 2 is often used and it is default base for log that we will be assuming hence forth. Note that since $0 \le p(x_j) \le 1$, then self information $I(x_j)$ will be non-negative i.e. $I(x_j) \ge 0$.

1.3 ENTROPY AND ITS PROPERTIES

Till now, we were talking about information of single message generated by the source. What about information generated by the source ? It can be total information or average information generated by the source. But only average information can characterize the source. Average information of a source is information generated per individual message. It is also called entropy of the source.

Definition : Entropy of a source is average amount of information generated by the source per message.

Let us consider a source X generating m different messages, $x_1, x_2, x_3, \ldots x_m$ with corresponding probabilities $p(x_1), p(x_2), p(x_3), \ldots p(x_m)$.

It means in a long stream of messages generated by source, the probability of generating message x_i will be $p(x_i)$.

(If $p(x_i) = 0.1$, then 10 out of 100 messages generated by the source will be x_i). Suppose, this source generates L number of messages at a particular time, where $L >> m$.

∴ Number of x_1 messages generated by the source

$$= p(x_1) \times L$$

∴ Total amount of information contained in all x_1's will be

$$= p(x_1) \times L \times \log \frac{1}{p(x_1)}$$

Similarly, number of x_2 messages generated by the source

$$= p(x_2) \times L$$

Total amount of information contained by all x_2's will be

$$= p(x_2) \times L \times \log \frac{1}{p(x_2)}$$

and so on.

Therefore, total amount of information contained in all L messages will be

$$I_{total} = p(x_1) \times L \log \frac{1}{p(x_1)} + p(x_2) \times L \log \frac{1}{p(x_2)} + \ldots\ldots + p(x_m) \times L \log \frac{1}{p(x_m)} \quad \text{... (1.3)}$$

Therefore, average amount of information generated by the source per message will be $\frac{I_{total}}{L}$.

It is called entropy and is denoted by H.

Therefore, entropy of the source X will be

$$H(X) = \frac{I_{total}}{L} \text{ bits/message}$$

$$= p(x_1) \log \frac{1}{p(x_1)} + p(x_2) \log \frac{1}{p(x_2)} + \ldots\ldots + p(x_m) \log \frac{1}{p(x_m)}$$

$$= \sum_{j=1}^{m} p(x_j) \log \frac{1}{p(x_j)} \text{ bits/message}$$

$$= -\sum_{j=1}^{m} p(x_j) \log p(x_j) \text{ bits/message} \qquad \ldots (1.4)$$

Note that the unit of entropy is bits/message. It can also be given as **bits/symbol.**

Example 1.1 : *Find entropy of a source X generating 4 types of messages with probabilities* $\frac{1}{4}, \frac{1}{8}, \frac{1}{8}$ *and* $\frac{1}{2}$.

Solution : Let given source be

$$X = \{x_1, x_2, x_3, x_4\}$$

$$\therefore \quad P(X) = \left\{\frac{1}{4}, \frac{1}{8}, \frac{1}{8}, \frac{1}{2}\right\}$$

Entropy of the source is given by

$$H(X) = \sum_{j=1}^{4} p(x_j) \log_2 \frac{1}{p(x_j)}$$

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

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

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

$$= 1.75 \text{ bits/message}$$

$$\therefore \quad \text{Entropy} = 1.75 \text{ bits/message}$$

If a source is generating only single message i.e. $m = 1$, $p(x_1) = 1$. Entropy of this source

$$H(X) = p(x_1) \log \left[\frac{1}{p(x_1)}\right] = 1 \cdot \log \frac{1}{1}$$

$$= 0 \text{ bits/message}$$

Thus, this source is generating no information at all ! It is because there is no uncertainty in generating message. We are certain that the source is going to generate message x_1 only. Now, consider a source which is generating two messages (Binary source) x_1 and x_2. Let the probabilities be $p(x_1)$ and $p(x_2)$ respectively. The entropy of this source is given by

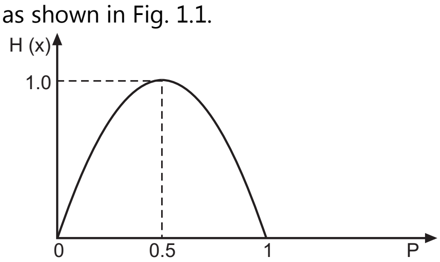
$$H(X) = p(x_1) \log \frac{1}{p(x_1)} + p(x_2) \log p(x_2) \qquad \ldots (1.5)$$

Let $p(x_1) = p$ and $p(x_2) = q = 1 - p$

$$\therefore \quad H(X) = p \log \frac{1}{p} + q \log \frac{1}{q} \qquad \ldots (1.6)$$

$$= p \log \frac{1}{p} + (1 - p) \log \frac{1}{(1 - p)} \qquad \ldots (1.7)$$

If we plot H(X) against p, the graph will be as shown in Fig. 1.1.



Entropy variation against p

Fig. 1.1 : Plot of H(X) for Binary Source

It can be observed from graph that

- H(X) is non-negative.
- H(X) is zero only for p = 0 and p = 1 as there is no uncertainty.
- H(X) is maximum at $p = \frac{1}{2}$.

The value of p, where maximum value of H(X) occurs, can be found by differentiating equation (1.7) w.r.t. p and equating it to zero.

$$H(X) = p \log \frac{1}{p} + (1-p) \log \frac{1}{(1-p)} \quad \text{... (1.8)}$$

$$H(X) = p \frac{\log_e \left(\frac{1}{p}\right)}{\log_e (2)} + (1-p) \frac{\log_e \left(\frac{1}{1-p}\right)}{\log_e (2)}$$

$$\therefore \quad H(X) = \frac{1}{\log_e (2)} \left[p \log_e \left(\frac{1}{p}\right) + (1-p) \log_e \frac{1}{(1-p)} \right] \quad \text{... (1.9)}$$

$$\therefore \quad \frac{d\,H(X)}{dp} = \frac{1}{\log_e (2)} \left[p \times p \times \left(-\frac{1}{p^2}\right) + \log_e \left(\frac{1}{p}\right) \times 1 + (1-p) \times (1-p) \times \frac{-1}{(1-p)^2} \times (-1) + \log_e \left(\frac{1}{1-p}\right) \times (-1) \right]$$

$$= \frac{1}{\log_e (2)} \left[-1 + \log_e \left(\frac{1}{p}\right) + 1 - \log_e \frac{1}{(1-p)} \right]$$

$$= \frac{1}{\log_e (2)} \left[\log_e \left(\frac{1}{p}\right) - \log_e \frac{1}{(1-p)} \right]$$

$\therefore$ Equating to 0 $\Rightarrow$

$$= \frac{1}{\log_e (2)} \left[\log_e \left(\frac{1}{p}\right) - \log_e \frac{1}{(1-p)} \right] = 0$$

$$\therefore \quad \log \frac{1}{p} = \log \frac{1}{1-p}$$

$$\therefore \quad p = 1-p$$

$$\therefore \quad P = \frac{1}{2}$$

It can also be seen that,

$$\frac{d^2H(X)}{d^2p} = -\frac{1}{p} - \frac{1}{1-p} < 0 \quad \text{... (1.10)}$$

Hence, maxima occurs at $p = \frac{1}{2}$

and

$$H(X)_{max} = \frac{1}{2} \times \log \frac{1}{\left(\frac{1}{2}\right)} + \frac{1}{2} \times \log \frac{1}{\left(\frac{1}{2}\right)} = 1 \text{ bit/message}$$

Above result can be extended for a source generating m messages. The entropy of this source will be maximum when all messages are equally likely or equiprobable. The maximum value of entropy will be obtained as below

$$H(x) = \sum_{j=1}^{m} p(x_j) \log \frac{1}{p(x_j)} \quad \text{... (1.11)}$$

$$H(x)_{max} = \sum_{j=1}^{m} \frac{1}{m} \log \left(\frac{1}{1/m}\right) \quad \text{... (1.12)}$$

$$= \log m \text{ bits/message} \quad \text{... (1.13)}$$

When the messages are equally likely, there is maximum uncertainty. Hence, we have maximum value of entropy. Thus, more is uncertainty, more will be the entropy.

Example 1.2: *A memoryless source has alphabet A = [– 5, – 3, – 1, 0, 3, 5] with corresponding probabilities {0.05, 0.1, 0.1, 0.15, 0.05, 0.25, 0.3}. Find the entropy of source. Assume that the source to be quantized according to the rule,*

$$q(-5) = q(-3) = -4$$

$$q(-1) = q(0) = q(1) = 0$$

$$q(3) = q(5) = 4$$

Find the entropy of the quantized source. Comment on the result.