An Introduction To Information Theory Solutions

Jake Parker

December 2023

1 Introduction

1. An alphabet consists of four letters *A, B, C, D* with respective probabilities of transmission $\frac{1}{3}$, $\frac{1}{4}$, $\frac{1}{4}$, $\frac{1}{6}$. Find the average amount of information associated with the transmission of a letter.

Solution:
To find the average amount of information associated with the transmission of a letter we need to find the entropy of the system.

$$\text{Entropy} = -\sum_{i=1}^{n} P(x_i)\log_2 P(x_i)$$
$$\text{Entropy} = -(\tfrac{1}{3}\log_2\tfrac{1}{3} + \tfrac{1}{4}\log_2\tfrac{1}{4} + \tfrac{1}{4}\log_2\tfrac{1}{4} + \tfrac{1}{6}\log_2\tfrac{1}{6})$$
$$\text{Entropy} = 1.9591 \text{ bits}$$

2. An independent, discrete source transmits letters selected from an alphabet consisting of three letters A, B, and C, with respective probabilities

$$p_A = 0.7 \quad p_B = 0.2 \quad p_C = 0.1$$

a) Find the entropy per letter.
b) If consecutive letters are statistically independent and two-symbol words are transmitted, find all the pertinent probabilities for all two-letter words and the entropy of the system of such words.

Solution:
a) To find the entropy per letter multiply the information of each letter by the probability the letter occurs and sum the results.

$$\text{Information} = I_x = -\log_2 p_x$$
$$\text{Entropy per letter} = H_x = I_x * p_x$$
$$H_A = 0.36 \text{ bits} \quad H_B = 0.46 \text{ bits} \quad H_C = 0.33 \text{ bits}$$

$$\Sigma H_x = H_A + H_B + H_C = 1.15 \text{ bits}$$

b) First we need to find the probability of each two-letter word occurring. Assuming each letter is statistically independent of the other, we multiply together the probabilities of each possible two-letter combination to get the probability of each two-letter word.

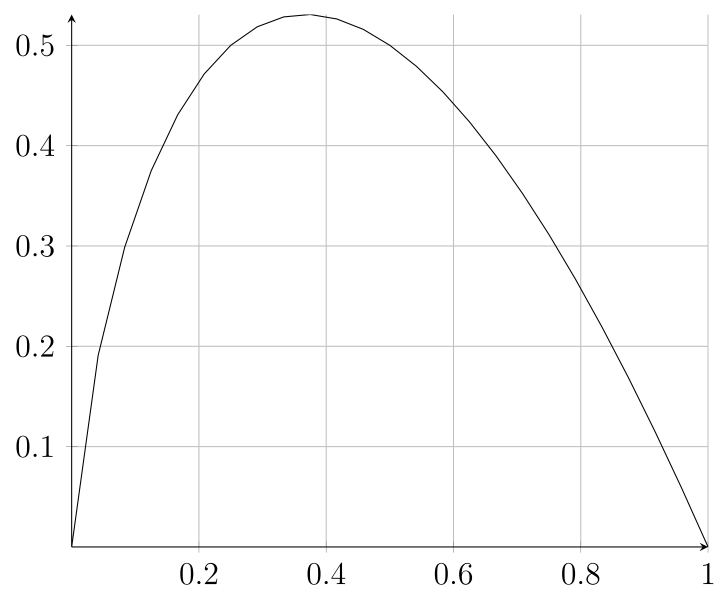
$$p_{AA} = 0.49 \quad p_{AB} = 0.14 \quad p_{AC} = 0.07$$
$$p_{BA} = 0.14 \quad p_{BB} = 0.04 \quad p_{BC} = 0.02$$
$$p_{CA} = 0.07 \quad p_{CB} = 0.02 \quad p_{CC} = 0.01$$

Then we can simply use the Shannon entropy formula to get the entropy of the system.

$$H(X) = -\sum_{i=1}^{n} P(x_i)\log_2 P(x_i)$$
$$H(X) = 2.31 \text{ bits}$$

3. Plot the curve $y = -x\log_2 x$ for $0 \le x \le 1$

Solution:



4. A pair of dice are thrown. We are told that the sum of the faces is 7. What is the average amount of information contained in this message (that is, the entropy associated with the probability scheme of having the sum of the faces equal to 7)?

Solution:

Of the 36 possible rolls between two independent dice, the sum 7 will come up 6 times.

$$(1,6) \quad (2,5) \quad (3,4) \quad (4,3) \quad (5,2) \quad (6,1)$$

This means that we can find average amount of information as follows

$$I = -\log_2 P(x) = -\log_2 \frac{6}{36} = 2.585 \text{ bits}$$

5. An alphabet consists of six symbols A, B, C, D, E, and F which are transmitted with the probabilities indicated below:

A	0	$\frac{1}{2}$
B	01	$\frac{1}{4}$
C	011	$\frac{1}{8}$
D	0111	$\frac{1}{16}$
E	01111	$\frac{1}{32}$
F	011111	$\frac{1}{32}$

a) Find the average amount of information content per letter.

b) If the letters are encoded in a binary system as shown above, find P(1), P(0), and the entropy of the binary source.

Solution:

a) To find the average amount of information content per letter we must use the Shannon entropy formula.

$$\text{Entropy} = H(X) = -\sum_{i=1}^{n} P(x_i)\log_2 P(x_i) = 1.9375 \text{ bits}$$

b) To find P(0) we need to count the 0's per word and multiply that by the frequency of that word showing up. Then we can divide that by the total letters in a word multiplied by the word frequency.

$$P(0) = \frac{\Sigma P(x_i)*N_i}{\Sigma P(x_i)*M_i}$$

$$N_i = \text{the number of 0s in the ith word}$$

$$M_i = \text{the total length of the ith word}$$

$$P(0) = \frac{(\frac{1}{2}*1)+(\frac{1}{4}*1)+(\frac{1}{8}*1)+(\frac{1}{16}*1)+(\frac{1}{32}*1)+(\frac{1}{32}*1)}{(\frac{1}{2}*1)+(\frac{1}{4}*2)+(\frac{1}{8}*3)+(\frac{1}{16}*4)+(\frac{1}{32}*5)+(\frac{1}{32}*6)} = 0.5079$$

Since we already know P(0), we can easily find P(1)

$$P(1) = 1.0 - P(0) = 0.4921$$

Now we can simply find the entropy of a binary case

$$H(X) = -P(0)\log_2(P(0)) - P(1)\log_2(P(1))$$

$$H(X) = -0.5079\log_2(0.5079) - 0.4921\log_2(0.4921)$$

$$H(X) \approx 0.9998 \text{ bits}$$