Information Theory and Coding

Examples for Entropy

م. فؤاد حمادي

2020 - 2019

Example6: a source produce a stream of **twenty** letters (A,B,C,D,E) with probabilities

$P(A) = P(E)$,

$P(B) = P(D)$,

$P(A) = 0.5\ P(B) = 0.25\ P(C)$.

Find

a. The entropy for this source

b. The amount of information each letter convey

c. The amount of information that the total message convey.

Sol

$P(A) = P(E) = 0.1$

$P(B) = P(D) = 0.2$

$P(C) = 0.4$

a. $\mathbf{H = - \Sigma\ pi \log pi}$

$= - (0.1 \log 0.1 + 0.2 \log 0.2 + 0.4 \log 0.4 + 0.2 \log 0.2 + 0.1 \log 0.1)$

$= - (2 \times 0.1 \log 0.1 + 2 \times 0.2 \log 0.2 + 0.4 \log 0.4)$

$= - (2 \times 0.1 \times -3.322 + 2 \times 0.2 \times -2.322 + 0.4 \times -1.322)$

$= 0.66444 + 0.92888 + 0.52888 = 2.1222$

b. $I_A = - \log 0.1 = 3.3222 = I_E$

$I_B = - \log 0.2 = 2.3222 = I_D$

$I_C = - \log 0.4 = 1.3222$

c . $I_{message} = 2\times 0.3222 + 4\times 2.3222 + 8 \times 1.3222 + 4 \times 2.3222 + 2 \times 3.3222$

$= 42.444$

Or

I message = no. of letters $\times\ \bar{I}$

$= 20 * 2.1222$

$= 42.444$

Entropy and Length of the Code

One of the key concepts in coding theory : we want to assign a fewer number of bits to code the more likely events.

0 ≤ Entropy ≤ log2 (M) also 0 ≤ H ≤ L;

This illustrate that the more randomness that exist in the source symbols, the more bits per symbol are required to represent those symbols.

on the other hands, entropy provides us with the theoretical minimum for the average number of bits per symbol (average length of the code) that could be used to encode the same symbol. The closer L is to the entropy, the better the coder.

Code Efficiency and Redundancy

$$\xi_{code} = \frac{H(x)}{L} * 100\% \quad where\ \xi_{code} = \text{code Efficiency}$$

$$R_{code} = \frac{L - H(x)}{L} * 100\% = \left(1 - \frac{H(x)}{L}\right) * 100\%$$

$$= \left(1 - \xi_{code}\right) * 100\% \qquad where\ R_{code} = Code\ Redundancy$$

Source Coding Techniques

1. Fixed Length Coding

In fixed length coding technique all symbols assigned with equal length because the coding don't take the probability in account.

The benefit of the fixed length code is ease of applied (easy in coding and decoding)

Example1: Let x = { $x_1, x_2, \ldots, x_{16}$} where pi= 1/16 for all i , find ζ source code

Sol

$H(x) = \log_2 M$

$= \log_2 16$

$= 4$ bit/symbol (because $p_1 = p_2 = \ldots = p_{16} = 1/M$)

For fixed length code

$L = {}_i = [\ log_2 M\ ] = [\ log_2 16] = [4] = 4$

$\therefore\ \xi\ \ source\ code = H(x) / L * 100\% = 4/4 * 100\% = 100\%$

source symbols	probability	codeword	code length
X_1	P_1	0000	4
X_2	P_2	0001	4
.	.	.	.
.	.	.	.
X_{16}	P_{16}	1111	4

Example2: Let $x = \{ x_1, x_2, x_3, x_4, x_5 \}$ where pi= 1/5 for all i , find

ζ source code

Sol

$H(x) = \log_2 M$

$= \log_2 5 = 2.322$ bit/symbol

For fixed length code

$L = {}_i = [\ log_2 M\ ] = [\ log_2 5] = [2.322] = 3$ bit

$\therefore\ \xi\ \ source\ cod\ e$ = H(x) / L * 100% = 2.322/3 *100% =77%

Symbol	Probability	Code	l_i
x_1	0.2	000	3
x_2	0.2	001	3
x_3	0.2	010	3
x_4	0.2	011	3
x_5	0.2	100	3

Example3: Let x = { $x_1, x_2, \ldots, x_{12}$} where pi= 1/12 for all i ,find ξ source code

Sol

$H(x) = \log_2 M$

$= \log_2 12 = 3.585$ bit/symbol

For fixed length code

$L = {}_i = [\ log_2 M\ ] = [\ log_2 12] = [3.585] = 4$ bit

$\therefore\ \xi\ \ source\ cod\ e = H(x) / L * 100\% = 3.585/4 * 100\% = 89\%$

Symbol	Probability	Code	l_i
X_1	1/12	0000	4
X_2	1/12	0001	4
.	.	.	.
.	.	.	.
X_{12}	1/12	1100	4