Tutorial - 4

Information & Communication

Problem -1 (Independent Random Variables)

Let X, X' be independent random variables with

$X \sim p(x)$ and $X' \sim r(x)$, x, $x' \in \mathcal{X}$.

Then prove that

$$P(X = X') \geq 2^{-H(p) - D(p \| r)}$$

$$P(X = X') \geq 2^{-H(r) - D(r \| p)}$$

Problem-2 (Fixed length source coding)

A source emits a sequence of independent binary digits with probabilities $p(1) = 0.005$ and $p(0) = 0.995$. The digits are taken 100 at a time and a binary codeword is provided for every sequence of 100 digits containing three or fewer 1's.

Assuming that all codewords are the same length, find the minimum length required to provide unique codewords for all sequences with three or fewer 1's. Other sequences need not be assigned any codeword

Problem-2 (Variable Length Source Coding)

Let the range of random variable X be $\{0,1,2,3,4\}$. Consider the two distributions $p(x)$ and $q(x)$ on this random variable. Codes for random variable X

Symbol	$p(x)$	$q(x)$	$C_1(x)$	$C_2(x)$
1	1/2	1/2	0	0
2	1/4	1/8	10	100
3	1/8	1/8	110	101
4	1/16	1/8	1110	110
5	1/16	1/8	1111	111

(a) Calculate $H(p)$, $H(q)$, $D(p\|q)$ and $D(q\|p)$

(b) Check if C_1 and C_2 are prefix-free codes.

(c) Verify that average length of C_1 under p is equal to the entropy $H(p)$. Thus, C_1 is optimal for p. Verify that C_2 is optimal for q.

(d) Now assume that we use C_2 when distribution is p. What is the average length of the codewords? By how much does it exceed entropy $H(p)$?