

Fundamentals of Information Science: Homework 4

March 12, 2025

Problem 1.

Entropy of a sum. Let X and Y be random variables that take on values x_1, x_2, \dots, x_r and y_1, y_2, \dots, y_s , respectively. Let $Z = X + Y$.

(a) Show that $H(Z|X) = H(Y|X)$. Argue that if X, Y are independent, then $H(Y) \leq H(Z)$ and $H(X) \leq H(Z)$. Thus, the addition of independent random variables adds uncertainty.

(b) Give an example of (necessarily dependent) random variables in which $H(X) > H(Z)$ and $H(Y) > H(Z)$.

(c) Under what conditions does $H(Z) = H(X) + H(Y)$?

Problem 2.

AEP and source coding. A discrete memoryless source emits a sequence of statistically independent binary digits with probabilities $p(1) = 0.01$ and $p(0) = 0.99$. 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.

(a) Assuming that all codewords are the same length, find the minimum length required to provide codewords for all sequences with three or fewer 1's.

(b) Calculate the probability of observing a source sequence for which no codeword has been assigned.

(c) Use Chebyshev's inequality (search online if you don't know Chebyshev's inequality) to bound the probability of observing a source sequence for which no codeword has been assigned. Compare this bound with the actual probability computed in part (b).