ECE 563: Homework 2 - Due October 15th, 2024

Problem 1: Properties of mutual information. Let X, Y_1, Y_2 be three RVs.
(a) Given $I(X;Y_1) = I(X;Y_2) = 0$, does it follow that $I(X;Y_1,Y_2) = 0$?
(b) Given $I(X;Y_1) = I(X;Y_2) = 0$, does it follow that $I(Y_1;Y_2) = 0$?

Problem 2: Data Processing Inequality. Let the RVs X, U, Z form a Markov chain, $X \to Y \to Z$.
(a) Show that $H(X|Y) = H(X|Y,Z)$.
(b) Show that $H(X|Y) \leq H(X|Z)$.
(c) Show that $I(X;Y) \geq I(X;Z)$.
(d) Show that $I(X;Z|Y) = 0$.

Problem 3: Submodularity. Show that Shannon's entropy is a submodular function.

Problem 4: Cover and Thomas, Monotonic convergence of the empirical distribution.

Problem 5: Cover and Thomas, Calculation of typical sets.

Problem 6: LLNs. The problem has two parts. It should help you learn more about different types of LLNs.

- Let $x_1, x_2, \ldots \ldots$ and $y_1, y_2, \ldots$ be two sequences of real numbers. The sequence $y_1, y_2, \ldots$ is increasing and diverges (i.e., the limit goes to ∞). Assume now that $\sum_{n=1}^{\infty} \frac{x_n}{y_n}$ converges to a finite limit. Then, $\frac{1}{y_n} \sum_{i=1}^{n} x_i$ converges to 0 as $n \to \infty$. *Hints: you can use the summation by part formula or the Stolz-Cesaro theorem to prove the result - feel free to look up the info I gave you as a hint online (e.g., you can check summation by part and/or Stolz-Cesaro's theorem); you are still not allowed to consult any other source to prove the actual question stated.*
- (Challenging and **optional.** You can look up Kolmogorov's criteria for the strong LLNs just to learn more about his results.) Using the above result, prove that for a sequence of independent but not necessarily identically distributed RVs, $X_1, X_2, \ldots, \ldots$, with zero mean, $\sum_{i=1}^{\infty} \frac{var(X_i)}{i^2} \to 0$ implies that $\frac{\sum_{i=1}^{n} X_i}{n} \to 0$ a.s.

Problem 7: Lossless compression. Consider a Discrete Memoryless Source, X, with an alphabet of 7 symbols $\{A, B, C, D, E, F, G\}$ with probabilities $\{0.05, 0.08, 0.13, 0.09, 0.30, 0.20, 0.15\}$ respectively.
(a) Calculate the entropy of this source.
(b) Ignore the provided probabilities and implement a fixed-length code for this source.
(c) Find a prefix-free variable-length code for this source using the Huffman algorithm. Calculate its average codelength and compare it with the values found in (a) and (b).
(d) Verify that the code you found in (c) meets the Kraft's Inequality.
(e) If we encode source symbols in 10-tuples using a fixed-to-variable length code, what are the lower and upper bounds on the minimum average codelength?
(f) Consider a source output of EFCEAFBE.

- Calculate the sample average of the log pmf of this string of symbols.
- Find the expected value of the log pmf for X.
- Find the probability that the log pmf sample average for this string is within ± 1 of the source entropy if the log pmf variance is 2.
- If we examine n-tuples from the source, what minimum integer value of n is needed if we want the probability of any n-tuple being in the typical set to be greater than 0.99.
- Assume a log pmf variance of 0.25 and $\epsilon = 1$; for the value of n found in the previous part of the problem, how many possible n-tuples are there?
- Calculate the approximate number of n-tuples in the typical set. Calculate the percentage of all n tuples that are in the typical set.
- Calculate the upper bound on the average codelength for a fixed-to-fixed-length code used to encode this typical set.

Problem 8: Lagrange multipliers. Explain the idea behind the Lagrange multipliers method (e.g., why does the formulation of the Lagrange multiplier objective work, what is the role of λ etc) on the example of the following constrained optimization problem: minimize $F(x, y) = 8x^2 - 2y$, where x, y are constrained to satisfy $x^2 + y^2 = 1$.