IERG 6154: Network Information Theory
Homework 1

Due: Jan 25, 2023

1. *Inequalities*: Show that
   (a) $H(X|Z) \leq H(X|Y) + H(Y|Z)$
   (b) $I(X_1, X_2; Y_1, Y_2) \leq I(X_1; Y_1) + I(X_2; Y_2)$ if $p(x_1, x_2, y_1, y_2) = p(x_1, x_2)p(y_1|x_1)p(y_2|x_2)$
   (c) $I(X_1, X_2; Y_1, Y_2) \geq I(X_1; Y_1) + I(X_2; Y_2)$ if $p(x_1, x_2, y_1, y_2) = p(x_1)p(x_2)p(y_1, y_2|x_1, x_2)$
   (d) $h(X+aY) \geq h(X+Y)$ when $Y \sim \mathrm{N}(0,1), a \geq 1$. Assume that X and Y are independent.

2. Prove or give a counterexample to the following inequalities:
   (a) $H(X_1, X_2, X_3) + H(X_1, X_2, X_4) + H(X_2, X_3, X_4) + H(X_1, X_3, X_4) \leq \frac{3}{2}[H(X_1, X_2) + H(X_1, X_4) + H(X_2, X_3) + H(X_3, X_4)]$.
   (b) $H(X_1, X_2, X_3) + H(X_1, X_2, X_4) + H(X_2, X_3, X_4) + H(X_1, X_3, X_4) \leq 3[H(X_1, X_2) + H(X_3, X_4)]$.

3. (Korner-Marton Identity) Show the following equality for any pair of sequences Y^n, Z^n

$$\sum_{i=1}^{n} I(Y^{i-1}; Z_i | Z_{i+1}^n) = \sum_{i=1}^{n} I(Z_{i+1}^n; Y_i | Y^{i-1})$$

   Note: Y^m and Y_n^m denote $Y_1, Y_2, ...Y_m$ and $Y_n, Y_{n+1}, ...Y_m$ respectively. There is an abuse of notation at the edges. Take Y^0 and Z_{n+1}^n as trivial random variables.

   Remark: This equality was discovered by Marton and was first used in a joint paper with Korner (1977). This was conveyed to me by Janos Korner. Unfortunately, it has been mistakenly called Csiszar-sum Lemma by the community after attributing the result to a later paper by Csiszar and Korner.

4. Prove that when $f : \mathbb{R} \to [0, \frac{1}{2}]$ such that $H(f(u))$ is convex and f is twice differentiable, then $H(f(u) * p)$ is convex in u for $p \in [0, 1]$. (Fan Cheng's extension of Mrs. Gerber's lemma). In the above, for $0 \leq x, y \leq 1$ define $x * y = x(1 - y) + y(1 - x)$. As a corollary note that $H(H^{-1}(u) * p)$ is convex in u for $p \in [0, 1]$ where $H^{-1}(u)$ as a mapping from $[0, 1]$ to $[0, \frac{1}{2}]$.

5. (Strong data processing inequalities for relative entropy) Given a channel $W(y|x)$, define

$$\eta_W = \sup_{p,q:q \ll p} \frac{D(Wq \| Wp)}{D(q \| p)}.$$

Show the following:

- When W is $BEC\,(\epsilon)$ then $\eta_W = 1 - \epsilon$
- When W is $BSC\left(\frac{1+\rho}{2}\right)$ then $\eta_W = \rho^2$
- When W is a Z-channel with $W(Y = 0|X = 0) = 1$ and $W(y = 0|x = 1) = z$, then $\eta_W = 1 - z$.

6. Let $X_1, .., X_n$ be i.i.d. finite valued random variables. Let $S_n = X_1 + \cdots + X_n$, $n \geq 1$. Show that

- For all $U \to X_1 + X_2 \to X_1$, we have $\frac{I(U;X_1)}{I(U;X_1+X_2)} \leq \frac{1}{2}$.
- More generally, for $m \leq n$, for all $U \to S_n \to S_m$, we have $\frac{I(U;S_m)}{I(U;S_n)} \leq \frac{m}{n}$.

7. (Compression inequality for mutual information and independent random variables) Let $X_1, .., X_n$ be a collection of mutually independent random variables. Then show that:

- If $\{\beta_T\}_T$ is a *compression* of $\{\alpha_T\}_T$, then for any $P_{U|X_1,..,X_n}$, we have

$$\sum_{T \subseteq [1:n]} \alpha_T I(U; X_T) \leq \sum_{T \subseteq [1:n]} \beta_T I(U; X_T).$$

- A collection of non-negative numbers$\{\beta_T\}_T$ is called a *fractional partition* of $[1 : n]$ if $\sum_{T:i \in T} \beta_T = 1$ for all $1 \leq i \leq n$. Define

$$\gamma_T = \begin{cases} 1 & T = \{1, 2, .., n\} \\ (\sum_T \beta_T) - 1 & T = \emptyset \\ 0 & \text{o.w.} \end{cases}.$$

  Show that γ_T is a compression of β_T.