Homework Set #2

Due 21st October 2022, before 11:59pm.

Submit your solutions to Gradescope with Entry Code: **57DN5B**

Problem 1 (Conditional Mutual Information)

Define $I(X_1; X_2; X_3) = I(X_1; X_2) - I(X_1; X_2|X_3)$.

(a) Prove or disprove $I(X_1; X_2; X_3) \geq 0$.

(b) Show that

$$-\min\{I(X_1; X_2|X_3), I(X_1; X_3|X_2), I(X_2; X_3|X_1)\} \leq I(X_1; X_2; X_3)$$

(c) Show that

$$I(X_1; X_2; X_3) \leq \min\{I(X_1; X_2), I(X_1; X_3), I(X_2; X_3)\}$$

Problem 2 (Secrecy Scenario with Access to Public key)

Let X be plain text, Y be cipher text, and Z be a secret key in a crypto system. Since X can be recovered from Y and Z, we have $H(X\big|Y, Z) = 0$. Unfortunately, there is a publicly known key, represented using the random variable U, which has a dependence with the secret key Z. However, the secret key Z and the public key U are independent of the plain text X. Show that

$$I(X; Y, U) = I(X; Y\big|U) \geq H(X) - H(Z\big|U)$$

Note that if the conditional entropy $H(Z\big|U) < H(X)$ then the cipher text along with the knowledge of the public key, U, reveals something about the plain text.

Problem 3 (Card Game)

Alice and Bob are playing the following card game. Alice can draw a card from her deck X that has three possible cards $X \in \{2, 4, 7\}$ with probabilities $\Pr[X = 2] = \Pr[X = 4] = \Pr[X = 7] = 1/3$. Similarly, Bob can draw a card from his deck Y that has three possible cards $Y \in \{3, 5, 6\}$ with probabilities $\Pr[Y = 3] = \Pr[Y = 5] = 1/5$ and $\Pr[Y = 6] = 3/5$.

Both Alice and Bob start the game ($t = 0$) with zero scores, $S_A^{t=0} = 0$ and $S_B^{t=0} = 0$, respectively. At each time, Alice and Bob draw a card from their decks simultaneously. If the value of Alice's card (X) is greater than the value of Bob's card (Y), then the score of Alice is increased by her card value, while the score of Bob is still the same. On the other hand, the value of Bob's card (Y) is greater than the value of Alice's card (X), then the score of Bob is increased by his card value, while the score of Alice is still the same. For example, suppose the first time $t = 1$, $X = 2$ and $Y = 5$, then the score of Alice will be the same $S_A^{t=1} = 0$, while the score of Bob will be increased $S_B^{t=1} = 5$. Suppose in the second round $t = 2$, $X = 4$ and $Y = 3$, then the score of Alice will be increased to $S_A^{t=2} = 4$, while the score of Bob is $S_B^{t=2} = 5$.

Suppose at the third round $t = 3$, $X = 7$ and $Y = 6$, then we get $S_A^{t=3} = 11$ and $S_B^{t=3} = 5$ and so on.

$$X = \begin{cases} 2 & \text{w.p. } 1/3 \\ 4 & \text{w.p. } 1/3 \\ 7 & \text{w.p. } 1/3 \end{cases}, \quad Y = \begin{cases} 3 & \text{w.p. } 1/5 \\ 5 & \text{w.p. } 1/5 \\ 6 & \text{w.p. } 3/5 \end{cases}$$

(a) How large is the score of Alice after n rounds? Can we find how the *normalized* score $\frac{1}{n}S_A^n$ behaves as $n \to \infty$? Similarly find the same for the score of Bob after n rounds and its *normalized* score $\frac{1}{n}S_B^n$ behavior as $n \to \infty$.
Hint: If X_i, Y_i are variables associated with draws from Alice and Bob (respectively) at draw i (as defined above) then the scores are obtained by examining the random variables Z_A^i, Z_B^i derived from (X_i, Y_i) which represent the increase in score increases for Alice and Bob respectively. For example,

$$Z_A^i = \begin{cases} 4 & \text{with prob. } 1/15 \\ 7 & \text{with prob. } 5/15 \\ 0 & \text{with prob. } 9/15 \end{cases}$$

and you can compute Z_B^i similarly. Then the scores are $S_A^n = \sum_{i=1}^n Z_A^i$, $S_B^n = \sum_{i=1}^n Z_B^i$ where $Z_A^1, \ldots, Z_A^n$ and $Z_B^1, \ldots, Z_B^n$ are random variables defined above.

(b) Suppose Alice is cheating, where she replaces her deck with the following deck:

$$X = \begin{cases} 2 & \text{with prob. } 1/3 \\ 4 & \text{with prob. } 1/3 \\ \alpha & \text{with prob. } 1/3 \end{cases}$$

What is the least value of the parameter α such that Alice's score is greater than Bob's score after n rounds as $n \to \infty$?

Hint: The weak law of large numbers might be useful. The weak law of large number states that for i.i.d. random variables $X_1, \ldots, X_n$, then $\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n X_i \to \mathbb{E}[X]$.
Hint: Define the new variables Z_A, Z_B for this scenario.

Problem 4 (Sufficient Statistic in Whisper Game)

Suppose there are m people playing a Whisper game. The game is played as follows: m people line up side-by-side. Person 1 generates a sequence of n i.i.d. random variables, $Y_1 = (X_1^{(1)}, X_2^{(1)}, \ldots, X_n^{(1)})$ distributed according to Bernoulli(p), i.e.,

$$X_i^{(1)} = \begin{cases} 1, & \text{w.p. } \quad p \\ 0, & \text{w.p} \quad 1-p, \end{cases}$$

where $p \in \{0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}, 1\}$ is not known.

Person 1 whispers this sequence, Y_1, to the next person's ear. Because of the whispering, the next person potentially mistakes the sequence for another sequence, say $Y_2 = (X_1^{(2)}, X_2^{(2)}, \ldots, X_n^{(2)})$, based on the conditional distribution $Q(Y_2 \big| Y_1)$, *i.e.,*

$$Q(Y_2|Y_1) = Q(X_1^{(2)}, X_2^{(2)}, \ldots, X_n^{(2)} \big| X_1^{(1)}, X_2^{(1)}, \ldots, X_n^{(1)}) = \prod_{i=1}^n q(X_i^{(2)} \big| X_i^{(1)})$$

Then the second person whispers her own understanding of the phrase, Y_2, to the third person, who hears Y_3 based on the conditional distribution $Q(Y_3|Y_2)$, *i.e.,*

$$Q(Y_3\big|Y_2) = Q(X_1^{(3)}, X_2^{(3)}, \ldots, X_n^{(3)}\big|X_1^{(2)}, X_2^{(2)}, \ldots, X_n^{(2)}) = \prod_{i=1}^{n} q(X_i^{(3)}\big|X_i^{(2)})$$

and so on until the m^{th} person. Each person only hears the whisper of the previous person. In the end, the m^{th} person broadcasts the estimate of p, i.e. $\hat{p}(Y_m)$, based on the sequence she hears from $(m-1)^{th}$ person, which is $Y_m = (X_1^{(m)}, X_2^{(m)}, \ldots, X_n^{(m)})$. The estimator $\hat{p}(Y_m)$ is given by:

$$\hat{p}(Y_m) = \frac{1}{n}\sum_{i=1}^{n} X_i^{(m)}$$

Note that $\hat{p}(Y_m) \in \{0, \frac{1}{n}, \ldots, \frac{n-1}{n}, 1\}$.

(a) Show that for any distribution on p, we have

$$I(p; \hat{p}(Y_m)) \leq I(p; Y_m)$$

(b) Prove that for $0 \leq k \leq n$,

$$\mathbb{P}\left[Y_m\Big|\hat{p}(Y_m)=\left(\frac{k}{n}\right)\right] = \begin{cases} \frac{1}{\binom{n}{k}}, & \text{if } \sum_{i=1}^{n} X_i^{(m)} = k \\ 0, & \text{otherwise} \end{cases}$$

(c) A statistic $\hat{p}(Y_m)$ is called sufficient if the equality holds in part (a) for any distribution on p. In this part we want to show that $\hat{p}(Y_m)$ is a sufficient statistic using the result in part (b), *i.e.*,

$$I(p; \hat{p}(Y_m)) = I(p; Y_m)$$

(d) Now suppose that each person $i \in \{1, 2, \ldots, m\}$ estimates p using the sequence, Y_i, they hear from the previous person, *i.e.*, their estimator $\hat{p}(Y_i)$ is given by

$$\hat{p}(Y_i) = \frac{1}{n}\sum_{l=1}^{n} X_l^{(i)}$$

Then show that for $j > i$,

$$I(p; \hat{p}(Y_i)) \geq I(p; \hat{p}(Y_j))$$

(e) Show that for $j > i$,

$$I(\hat{p}(Y_m); Y_{m-1}\big|Y_i) \geq I(\hat{p}(Y_m); Y_{m-1}\big|Y_j)$$

Problem 5 (Entropy rate)

Let $X_1, X_2, \ldots$ be a stationary Markov process with the following transition matrix:

$$\Pi = \begin{bmatrix} \frac{1}{4} & \frac{3}{4} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix} \tag{1}$$

(a) Find the stationary distribution of this Markov chain and find $H(X)$.

(b) What is the entropy rate ($\mathcal{H}(X)$) of this random process? and explain the comparison between the entropy rate of this random process with the entropy of the stationary distribution (i.e., compare $H(X)$ with $\mathcal{H}(X)$).

(c) Suppose a new random process $Y_1, Y_2, \ldots$ defined as follows: $Y_1 = X_1$ and $Y_n = X_n \oplus X_{n-1}$ for $n \geq 2$. Find the entropy rate of the new process $\mathcal{H}(Y)$.

(d) Compute the entropy $H(Y_n)$ for any $n \geq 3$.