AUSTRALIAN NATIONAL UNIVERSITY

COMP2610/COMP6261

Information Theory, Semester 2 2022

Assignment 1

Due Date: Monday 29 August 2022, 5:00 pm

Assignment 1 weighting is 10% of the course mark.

Instructions:

Marks:

1. The mark for each qu
are asked to prov
subsequent par

2. **COMP2610 students:**

   - Answer *Questions 1-3* and *Section I of Ques* ed out of 100.
   - You are not expected to answer Question 5 and Section II of Question 4.

3. **COMP6261 students:**
   - Answer *Questions 2*-5. You will be marked out of 100.
   - You are not expected to answer Question 1.

Submission:

1. Submit your assignment together **with a cover page** as a single PDF on Wattle.
2. **Clearly mention whether you are a COMP2610 student or COMP6261 student in the cover page.**
3. The deadline for the assignment is absolute. If you submit after the deadline you get zero marks (100% penalty), unless you are ill, in which case you will need to present a doctor's certificate, or have undergone severe trauma of some kind.
4. All assignments must be done individually. Plagiarism is a university offence and will be dealt with according to university procedures *http://academichonesty.anu.edu.au/UniPolicy.html*.

Question 1 [25 marks total]

****Only COMP 2610 students are expected to attempt this question.**

Suppose only two ANU residences, B (Burton from Garran Hall) and J (John from Burgman College), compete in a series of 5 races. The games play out as follows:

- Each race has exactly one winner (i.e. residences cannot tie).
- In every race, both residences have equal probability for winning.
- Once one residences wins 3 races, no further races are conducted.

Let X denote the random variable whose outcomes $\mathcal{X}$ are all possible strings representing the winners of the races, i.e., $\mathcal{X} = \{BBB, JJJ, BJBB, \dots, BJBJB\}$. Let Y denote a random variable representing the total number of races conducted in the series. Calculate each of the following, showing all your working.

(a) The information in seeing the sequence $BJBJB$. That is, compute $h(X = BJBJB)$

(b) Probability of Burton winning the series, if winning three races represents winning the series

(c) $H(X)$

(d) $H(Y)$

(e) $H(X|Y = 4)$

(f) $H(X|Y)$

(g) $H(Y|X)$. Provide an intuitive explanation of your answe

(h) Ordinarily, the winner of the series is simply the residence that wins the last race. However, if the residence who wins the last race in the series is different to the one who wins the race before the last race (i.e., the second last race), then officials are suspicious of some form of race fixing. In such a series, there is a 50% chance of officials declaring the series invalid, and neither residence is declared winner; the other 50% of the time, the officials decide the series is legitimate, and declare the winner as usual. Let W be a random variable denoting the identity of the winning residence, with possible outcomes $\mathcal{W} = \{A, B, N\}$, for the winner being respectively team A, team B, or neither. Compute $H(W)$.

Question 2 [25 marks total]

****All students are expected to attempt this question.**

(a) Let X be a random variable with m different outcomes $\mathcal{X} = \{1, 2, \cdots, m\}$ and corresponding nonzero probabilities $\{q_1, q_2, \cdots, q_m\}$ arranged in decreasing order. Now define a new random variable Y with m+1 different outcomes $\{\frac{q_1}{5}, \frac{q_2}{5}, \cdots, \frac{q_{m-1}}{5}, \frac{q_m}{5}, \frac{4}{5}\}$. Find the entropy of $H(Y)$ in terms of $H(X)$.

(b) Consider a fair coin flip. What is the mutual information between the top side and the bottom side of the coin?

(c) A 6-sided fair die is rolled. What is the mutual information between the top side and the front face (the side most facing you)?

(d) Show that the entropy of the probability distribution, $(p_1, \cdots, p_i, \cdots, p_j, \cdots, p_m)$, is less than the entropy of the distribution $(p_1, \cdots, \frac{p_i+p_j}{2}, \cdots, \frac{p_i+p_j}{2}, \cdots, p_m)$.

Question 3 [35 marks total]

****All students are expected to attempt this question.**

A standard deck of car , “diamonds”, “clubs”, “spades”) — each with 12 e J, Q, K are called “Jack”, “Queen”, “Kin coloured red; clubs and spades are bla *e cards.*

Each of the 48 cards in a deck is identified by its value d denoted VS. For example, 2♡, J♣, and ♠ are the “two of hearts”, “Jack of cl ctively. The variable C will be used to denote a card’s colour. Let ce card and $f = 0$ otherwise.

1. A card is drawn at random from a thoroughly shuffled deck. Calculate:

   (a) The information in observing a red non-face card, i.e., $h(c = \text{red}, f = 0)$

   (b) The conditional information in observing a King given a face card was drawn, i.e., $h(V = K|f = 1)$.

   (c) The entropies $H(S)$ and $H(V,S)$.

   (d) The mutual information between V and S, i.e., $I(V;S)$.

   (e) The mutual information between the value and colour of a card, i.e., $I(V;C)$.

2. Now consider that 16 cards are removed from a standard deck: All 12 ♠s; the 2♡, 3♡, 4♡, and 5♡.

   (a) Calculate the entropies $H(S)$ and $H(V,S)$. HINT: Express $H(V,S)$ in terms of $H(V|S)$.

   (b) Calculate $I(V;S)$. Explain why it is different to the $I(V;S)$ when a card is drawn at random from a standard of 48 cards (i.e. prior to the removal of 16 cards).

   (c) Calculate $I(V;S|C)$.

Question 4

Let Y be a random variable with possible outcomes $0, 1$, and $p(Y = 1) = 1/2$. Let X be a random variable with possible outcomes $X = a, b, c$. Define

$$\mathbf{p} = (p(X = a|Y = 1), p(X = b|Y = 1), p(X = c|Y = 1))$$

$$\mathbf{q} = (p(X = a|Y = 0), p(X = b|Y = 0), p(X = c|Y = 0)).$$

Suppose that

$$\mathbf{p} = (\frac{3}{5}, \frac{1}{5}, \frac{1}{5})$$

$$\mathbf{q} = (\frac{1}{5}, \frac{3}{5}, \frac{1}{5})$$

Section I [15 marks total]

****All students are expected to attempt this section.**

(a) Using the definition of mutual information, show that for any choice of $\mathbf{p}$, $\mathbf{q}$,

$$I(X;Y) = \frac{D_{\mathrm{KL}}(\mathbf{p}||\mathbf{m}) + D_{\mathrm{KL}}(\mathbf{q}||\mathbf{m})}{2}$$

where $\mathbf{m} = \frac{\mathbf{p}+\mathbf{q}}{2}$

(b) Compute $I(X;$

Section II [15 marks total]

****Only COMP 6261 students are expected to attempt this s**

(c) Let Z be a random variable with outcomes $x \in \mathtt{X}$, such that for all $x \in \mathtt{X}$ and $y \in \{0, 1\}$, $p(Z = x|Y = y) = p(X = x|Y = 1 - y)$. Using part (a) compute $I(Z;Y)$. Explain intuitively the relation of your answer to part (a).

(d) Suppose $\mathbf{p}$ and $\mathbf{q}$ are as in part (a). Using part (b), or otherwise, give an example of a random variable Z with possible outcomes $\mathtt{X}$ satisfying $I(X;Y) > I(Z;Y)$. Explain your answer in terms of the data-processing inequality.

(e) Suppose $\mathbf{p}$ and $\mathbf{q}$ are as in part (a). Give an example of a random variable Z with possible outcomes $x \in \mathtt{X}$ satisfying $I(X;Y) < I(Z;Y)$. Explain your answer in terms of the data-processing inequality.

Question 5 [10 marks total]

****Only COMP 6261 students are expected to attempt this question.**

Suppose X is a real valued random variable with $\mu = E(X) = 0$

(a) Show that for any $t > 0$ and $a \leq X \leq b$,

$$E(e^{tX}) \leq e^{g(t(b-a))},$$

where $g(u) = \log(1 - \gamma + \gamma e^{u}) - \gamma u$, with $\gamma = a/(a-b)$.

(Hint: write X as a convex combination of a and b, where the convex weighting parameter depends upon X. Exploit the convexity of the function $x \to e^{tx}$ and the fact that inequalities are preserved upon taking expectations of both sides, since expectations are integrals.)

(b) By using Taylor's theorem, show that for all $u > 0$,

$$g(u) \leq \frac{u^2}{8}.$$

Furthermore show that

$$E(e^{tX}) \leq e^{\frac{t^2(b-a)^2}{8}}.$$