ASSIGNMENT COVER SHEET

ANU College of Engineering and Computer Science
Australian National University
Canberra ACT 0200 Australia
www.anu.edu.au

+61 2 6125 5254

This coversheet must be attached to the front of your assessment

The assessment is due at 5pm unless otherwise specified in the course outline.

Student ID	u7351505
Student Name	Yifan Luo
Course Code	COMP6261
Course Name	Information Theory
Assignment Item	Assignment 1
Due Date	2022.8.29 5pm
Date Submitted	2022.8.29

I declare that this work:

- ☑ upholds the principles of academic integrity, as defined in the ANU Policy: Code of Practice for Student Academic Integrity;
- ☑ is original, except where collaboration (for example group work) has been authorised in writing by the course convener in the course outline and/or Wattle site;
- ☑ is produced for the purposes of this assessment task and has not been submitted for assessment in any other context, except where authorised in writing by the course convener;
- ☑ gives appropriate acknowledgement of the ideas, scholarship and intellectual property of others insofar as these have been used;
- ☑ in no part involves copying, cheating, collusion, fabrication, plagiarism or recycling.

Signature Yifan Luo

Q.2

(a) $H(X) = \sum_x p(x)\log_2\frac{1}{p(x)} = -q_1\log_2 q_1 - q_2\log_2 q_2 \cdots - q_m\log_2 q_m = -\sum_{i=1}^{m} q_i\log_2 q_i$

$H(Y) = \sum_y p(y)\log_2\frac{1}{p(y)} = -\frac{q_1}{5}\log_2\frac{q_1}{5} - \cdots - \frac{q_m}{5}\log_2\frac{q_m}{5} - \frac{4}{5}\log_2\frac{4}{5} = -\frac{1}{5}\sum_{i=1}^{m} q_i(\log_2 q_i - \log_2 5) - \frac{4}{5}\log_2\frac{4}{5}$

$= -\frac{1}{5}\sum_{i=1}^{m} q_i\log_2 q_i + \frac{1}{5}\sum_{i=1}^{m} q_i\log_2 5 - \frac{4}{5}\log_2\frac{4}{5}$

$= \frac{1}{5}H(X) - \frac{1}{5}\log_2\frac{1}{5} - \frac{4}{5}\log_2\frac{4}{5} = \frac{1}{5}H(X) + \frac{1}{5}\log_2 5 + \frac{4}{5}\log_2\frac{5}{4}$

$= \frac{1}{5}H(x) + \log_2 5 - \frac{8}{5}$

(b) $I(\text{top side}; \text{bottom side}) = H(\text{top side}) - H(\text{top side} \mid \text{bottom side})$

$= \sum_{t\in\{0,1\}} -p(\text{top side}=t)\log_2 p(\text{top side}=t) - \sum_{b\in\{1,0\}} p(\text{bottom side}=b)\, H(\text{top side} \mid \text{bottom side}=b)$

$= -2\times\frac{1}{2}\log_2\frac{1}{2} - 2\times\frac{1}{2}\times 0 = 1 - 0 = 1$

(c) $I(\text{top side}; \text{front side}) = H(\text{top side}) - H(\text{top side} \mid \text{front side})$

$H(\text{top side}) = -\sum_{t\in\{1,\cdots,6\}} P(\text{top side}=t)\log_2 P(\text{top side}=t) = -6\times\frac{1}{6}\log_2\frac{1}{6} = \log_2 6$

$H(\text{top side} \mid \text{front side}) = \sum_{f\in\{1,\cdots,6\}} P(\text{front side}=f)\, H(\text{top side} \mid \text{front side}=f)$

$= 6\times\frac{1}{6}\times 4\times\frac{1}{4}\log 4 = \log_2 4 = 2$

$\therefore I(\text{top side}; \text{front side}) = H(\text{top side}) - H(\text{top side}, \text{front side}) = \log_2 6 - 2 = \log_2 3 - 1$

(d) Let $X \in \{p_1, \cdots, p_i, \cdots, p_j, \cdots, p_m\}$, $Y \in \{p_1, \cdots, \frac{p_i+p_j}{2}, \cdots, \frac{p_i+p_j}{2}, \cdots, p_m\}$, $H(X) = \sum_{a=1}^{m} p_a\log_2\frac{1}{p_a}$

$H(Y) = \underline{\sum_{a=1}^{i-1} p_a\log_2\frac{1}{p_a}} + \frac{p_i+p_j}{2}\log_2\frac{1}{\frac{p_i+p_j}{2}} + \underline{\sum_{b=i+1}^{j-1} p_b\log_2\frac{1}{p_b}} + \frac{p_i+p_j}{2}\log_2\frac{1}{\frac{p_i+p_j}{2}} + \underline{\sum_{c=j+1}^{m} p_c\log_2\frac{1}{p_c}}$ $Y \in \{p_1, \cdots,$

$= H(X) - p_i\log_2\frac{1}{p_i} - p_j\log_2\frac{1}{p_j} + 2\times\frac{p_i+p_j}{2}\log_2\frac{1}{\frac{p_i+p_j}{2}}$

$H(Y) - H(X) = (p_i+p_j)\log_2\frac{2}{p_i+p_j} - p_i\log_2\frac{1}{p_i} - p_j\log_2\frac{1}{p_j} = (p_i+p_j)\log_2\frac{2}{p_i+p_j} - (p_i\log_2\frac{1}{p_i} + p_j\log_2\frac{1}{p_j})$

$= (p_i\log_2\frac{2}{p_i+p_j} - p_i\log_2\frac{1}{p_i}) + (p_j\log_2\frac{2}{p_i+p_j} - p_j\log_2\frac{1}{p_j})$

$= p_i(\log_2\frac{2}{p_i+p_j} - \log_2\frac{1}{p_i}) + p_j(\log_2\frac{2}{p_i+p_j} - \log_2\frac{1}{p_j})$

$= p_i\log_2\frac{\frac{2}{p_i+p_j}}{\frac{1}{p_i}} + p_j\log_2\frac{\frac{2}{p_i+p_j}}{p_j}$

$= p_i\log_2\frac{2p_i}{p_i+p_j} + p_j\log_2\frac{2p_j}{p_i+p_j}$

$\geq p_i(1 - \frac{p_i+p_j}{2p_i}) + p_j(1 - \frac{p_i+p_j}{2p_j})$

$= p_i + p_j - \frac{p_i+p_j}{2} - \frac{p_i+p_j}{2} = 0$

$\therefore H(Y) \geq H(X)$

Q3 1. (a) $h(\text{a red non-face card}) = h(c=\text{red}, f=0) = h(s\in\{\text{heart}, \text{diamond}\}, v\in\{2,3,\cdots,9\})$

$$= \log_2 \frac{1}{p(s,v)} = \log_2 \frac{1}{\frac{16}{48}} = \log_2 \frac{48}{16} = \log_2 3$$

(b) $h(V=k \mid f=1) = \log_2 \frac{1}{p(v=k \mid f=1)}$

$$= \log_2 \frac{1}{\frac{1}{4}} = 2$$

(c) $H(S) = \sum_s p(s) \log_2 \frac{1}{p(s)}$

$$= 4\times \frac{1}{4} \log_2 \frac{1}{\frac{1}{4}} = 2$$

$H(V,S) = \sum_v \sum_s p(v,s) \log_2 \frac{1}{p(v,s)}$

$$= 12\times 4 \times \frac{1}{48} \log_2 \frac{1}{\frac{1}{48}} = \log_2 48$$

(d) $I(V;S) = \sum_{v\in V}\sum_{s\in S} p(v,s) \log_2 \frac{p(v,s)}{p(v)p(s)} = \sum_v\sum_s p(v,s)\log_2 1 = 0$

(e) $I(V;C) = \sum_{v\in V}\sum_{c\in C} p(v,c)\log_2 \frac{p(v,c)}{p(v)p(c)} = \sum_v\sum_c p(v,c)\log_2 1 = 0$

2. (a) $H(S) = \sum_{s\in S} p(s) \log \frac{1}{p(s)} = p(\text{diamond})\log_2 \frac{1}{p(\text{diamond})} + p(\text{clubs})\log_2\frac{1}{p(\text{clubs})}$

$$+ p(\text{hearts})\log_2 \frac{1}{p(\text{heart})}$$

$$= 2\times\frac{12}{32}\log_2\frac{1}{\frac{12}{32}} + \frac{8}{32}\log_2\frac{1}{\frac{8}{32}}$$

$$= \frac{3}{4}\log_2\frac{8}{3} + \frac{1}{4}\log_2 4 = \frac{1}{2} + \frac{3}{4}\log_2\frac{8}{3}$$

$H(V|S) = \sum_s p(s) H(V|S=s) = \sum_s p(s) \sum_v p(v|S=s)\log_2 \frac{1}{p(v|S=s)}$

$$= 2\times\frac{12}{32}\times 12\times\frac{1}{12}\times\log_2\frac{1}{\frac{1}{12}} + \frac{8}{32}\times 8\times\frac{1}{8}\times\log_2\frac{1}{\frac{1}{8}}$$

$$= \frac{3}{4}\times\log_2 12 + \frac{1}{4}\times\log_2 8 = \frac{3}{4}\log_2 12 + \frac{3}{4}$$

$H(V,S) = H(S) + H(V|S) = \frac{1}{2} + \frac{3}{4}\log_2\frac{8}{3} + \frac{3}{4} + \frac{3}{4}\log_2 12 = \frac{5}{4} + \frac{3}{4}(\log_2\frac{8}{3} + \log_2 12)$

$$= \frac{5}{4} + \frac{3}{4}\log_2 32 = \frac{5}{4} + \frac{15}{4} = 5$$

(b) $I(V;S) = H(V) - H(V|S)$

$$= \sum_v p(v)\log_2\frac{1}{p(v)} - \left(\frac{3}{4}\log_2 12 + \frac{3}{4}\right)$$

$$= 4\times\frac{2}{32}\log_2\frac{1}{\frac{2}{32}} + 8\times\frac{3}{32}\log_2\frac{1}{\frac{3}{32}} - \left(\frac{3}{4}\log_2 12 + \frac{3}{4}\right)$$

$$= \frac{1}{4}\times 4 + \frac{3}{4}\times\log_2\frac{32}{3} - \frac{3}{4}\log_2 12 - \frac{3}{4}$$

$$= \frac{1}{4} + \frac{3}{4}\left(\log_2\frac{32}{3} - \log_2 12\right)$$

$$= \frac{1}{4} + \frac{3}{4}\log_2\frac{8}{9}$$

Because after cards are moved, the information of suit will convey information of value. Diamonds and clubs have chance to observing 2, 3, 4, 5, but hearts not.

(c) $I(V;S|C) = H(V|C) - H(V|S,C)$

$$H(V|C) = \sum_c p(c)\, H(V|C=c)$$

$$= \frac{20}{32}\sum_v p(v|C=r)\log_2\frac{1}{p(v|C=r)} + \frac{12}{32}\sum_v p(v|C=b)\log_2\frac{1}{p(v|C=b)}$$

$$= \frac{20}{32}\left(4\times\frac{1}{20}\log_2\frac{1}{\frac{1}{20}} + 8\times\frac{2}{20}\log_2\frac{1}{\frac{2}{20}}\right) + \frac{12}{32}\left(12\times\frac{1}{12}\log_2\frac{1}{\frac{1}{12}}\right)$$

$$= \frac{20}{32}\left(\frac{1}{5}\log_2 20 + \frac{4}{5}\log_2 10\right) + \frac{12}{32}\log_2 12$$

$$= \frac{1}{8}\log_2 20 + \frac{1}{2}\log_2 10 + \frac{3}{8}\log_2 12$$

$$H(V|S,C) = H(V|S) = \frac{3}{4}\log_2 12 + \frac{3}{4}$$

$$I(V;S|C) = \frac{1}{8}\log_2 20 + \frac{1}{2}\log_2 10 - \frac{3}{8}\log_2 12 - \frac{3}{4}$$

$$= -\frac{5}{4} + \frac{1}{8}\log_2 5 + \frac{1}{2}\log_2 10 - \frac{3}{8}\log_2 3$$

$m = p(x)$, $p = p(X|Y=1)$
$q = p(X|Y=0)$

Q4

(a). According to the definition, $p(Y) = \frac{1}{2}$

$p := p(X|Y=1)$, $q := p(x|Y=0)$

$$m := \frac{p+q}{2} = \frac{1}{2}(p(x|Y=1) + p(x|Y=0))$$
$$= p(Y=1)p(x|Y=1) + p(Y=0)p(x|Y=0)$$
$$= p(X)$$

Then, $\frac{D_{KL}(p||m) + D_{KL}(q||m)}{2}$

$$= \frac{1}{2}(D_{KL}(p(x|Y=1)||p(x)) + D_{KL}(p(x|Y=0)||p(x)))$$
$$= \frac{1}{2}D_{KL}(p(x|Y=1)||p(x)) + \frac{1}{2}D_{KL}(p(x|Y=0)||p(x))$$
$$= p(Y=1)\cdot\sum_x p(x|Y=1)\log\frac{p(x|Y=1)}{p(x)} + p(Y=0)\cdot\sum_x p(x|Y=0)\log\frac{p(x|Y=0)}{p(x)}$$
$$= \sum_x\sum_Y p(Y)p(X|Y)\log\frac{p(X|Y)}{p(X)}$$
$$= \sum_x\sum_Y p(X,Y)\log\frac{p(X|Y)p(Y)}{p(X)p(Y)}$$
$$= \sum_x\sum_y p(X,y)\log\frac{p(X,y)}{p(x)p(y)}$$
$$= I(X;Y)$$

(b) In (a), we see that $m = \frac{p+q}{2} = p(x)$, $p(x=a) = \frac{2}{5}$, $p(x=b) = \frac{2}{5}$, $p(x=c) = \frac{1}{5}$

$$I(x;Y) = p(Y=1)\sum_x p(x|Y=1)\log\frac{p(x|Y=1)}{p(x)} + p(Y=0)\sum_x p(x|Y=0)\log\frac{p(x|Y=0)}{p(x)}$$
$$= \frac{1}{2}\cdot\left[\frac{3}{5}\cdot\log\frac{\frac{3}{5}}{\frac{2}{5}} + \frac{1}{5}\cdot\log\frac{\frac{1}{5}}{\frac{2}{5}} + \frac{1}{5}\cdot\log\frac{\frac{1}{5}}{\frac{1}{5}}\right] +$$
$$\frac{1}{2}\cdot\left[\frac{1}{5}\cdot\log\frac{\frac{1}{5}}{\frac{2}{5}} + \frac{3}{5}\log\frac{\frac{3}{5}}{\frac{2}{5}} + \frac{1}{5}\log\frac{\frac{1}{5}}{\frac{1}{5}}\right]$$
$$= \frac{1}{2}\cdot\left[\frac{3}{5}\cdot\log\frac{3}{2} + \frac{1}{5}\log\frac{1}{2} + 0\right] + \frac{1}{2}\cdot\left[\frac{1}{5}\cdot\log\frac{1}{2} + \frac{3}{5}\cdot\log\frac{3}{2} + 0\right]$$
$$= \frac{3}{5}\cdot\log_2\frac{3}{2} + \frac{1}{5}\log_2\frac{1}{2}$$
$$= \frac{3}{5}\log 3 - \frac{4}{5}$$

(c) $p(Z=x \mid Y=y) = p(X=x \mid Y=1-y)$

$p = (p(X=a \mid Y=1) = \frac{3}{5},\ p(X=b \mid Y=1) = \frac{1}{5},\ p(X=c \mid Y=1) = \frac{1}{5})$
$q = (p(X=a \mid Y=0) = \frac{1}{5},\ p(X=b \mid Y=0) = \frac{3}{5},\ p(X=c \mid Y=0) = \frac{1}{5})$

$p' = (p(Z=a \mid Y=1) = p(X=a \mid Y=0) = \frac{1}{5},\ p(Z=b \mid Y=1) = \frac{3}{5},\ p(Z=c \mid Y=1) = \frac{1}{5})$
$q' = (p'(Z=a \mid Y=0) = p(X=a \mid Y=1) = \frac{3}{5},\ p(Z=b \mid Y=0) = \frac{1}{5},\ p(Z=c \mid Y=0) = \frac{1}{5})$

$\therefore p' = q,\ q' = p$.

$I(Z;Y) = \frac{D_{KL}(p' \| m') + D_{KL}(q' \| m')}{2}$, where $m' = \frac{p'+q'}{2} = \frac{q+p}{2} = m$

$= \frac{D_{KL}(q \| m) + D_{KL}(p \| m)}{2}$

$= I(X;Y) = \frac{3}{5} \cdot \log \frac{3}{2} + \frac{1}{5} \log \frac{1}{2}$

$\therefore I(Z;Y) = I(X;Y)$ Intuitively, Y brings the same uncertainty reduction for both X and Z.
and $P(Z \mid Y)$ is the same as $P(X \mid \neg Y)$

(d) Suppose $p(Z=a \mid Y=0) = p(Z=b \mid Y=0) = P(Z=c \mid Y=0) = \frac{1}{3}$
$p(Z=a \mid Y=1) = p(Z=b \mid Y=1) = P(Z=c \mid Y=1) = \frac{1}{3}$
then $P(Z=a) = P(Z=b) = P(Z=c) = \frac{1}{3}$, Z and Y are independent.
Thus, $I(X;Y) > I(Z;Y) = 0$
Intuitively, Y conveies no information of Z, because they are independent.

(e) Suppose $P(Z=a \mid Y=0) = P(Z=b \mid Y=0) = P(Z=c \mid Y=0) = 0$
$p(Z=a \mid Y=1) = p(Z=b \mid Y=1) = P(Z=c \mid Y=1) = \frac{1}{3}$
then $P(Z=a) = P(Z=b) = P(Z=c) = \frac{1}{3}$

$I(Z;Y) = H(Z) - H(Z \mid Y)$
$= 3 \times \frac{1}{3} \log_2 \frac{1}{\frac{1}{3}} - \frac{1}{2} \times 3 \times \frac{1}{3} \log \frac{1}{\frac{1}{3}}$
$= \log_2 3 - \frac{1}{2} \log_2 3$
$= \frac{1}{2} \log_2 3$

$I(X;Y) < I(Z;Y)$

Intuitively, given the knowledge of Y, we know that Z will be 0 if Y equals to 0,
otherwise Z has equal probability of a, b, c. But X is less dependent on Y,
because X is more averagely distributed.

Q.5 (a). Let $a \geq 0$, $b \geq 0$, then represent X as $X = \lambda_1 a + \lambda_2 b$, $\lambda_1 + \lambda_2 = 1$

$\because$ $t > 0$ and $a \leq x \leq b$

$\therefore$ $\lambda_1 = \frac{b-X}{b-a}$, $\lambda_2 = \frac{X-a}{b-a}$

$E(e^{tX})$ is a moment generating function for X.

then $g(t(b-a))$ is the convexion point of e^{tX}

$\because$ $g(u) = \log(1-r+re^u) - ru$, $r = \frac{a}{a-b}$

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

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

(b) Taylor theorem: $f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(k)}(a)}{k!}(x-a)^k + h_k(x)(x-a)^k$

$(\lim_{x \to a} h_k(x) = 0)$

$\because$ $e^{\lambda X}$ is convex,

$\therefore$ $e^{\lambda X} \leq \frac{b-X}{b-a} e^{\lambda a} + \frac{X-a}{b-a} e^{\lambda b}$

$\therefore$ $E(e^{\lambda X}) \leq \frac{b-EX}{b-a} e^{\lambda a} + \frac{EX-a}{b-a} e^{\lambda b} = \frac{b}{b-a} e^{\lambda a} + \frac{-a}{b-a} e^{\lambda b} = e^{L(\lambda(b-a))}$

let $L(x) = \frac{xa}{b-a} + \ln(1 + \frac{a - e^x a}{b-a})$

$\because$ $L(0) = L'(0)$ and $L''(x) \leq \frac{1}{4}$

$\therefore$ By Taylor theorem, $L(x) = L(0) + xL'(0) + \frac{1}{2}h^2 L''(x\theta) \leq \frac{1}{8}x^2$

Thus $E(e^{\lambda X}) \leq e^{\frac{t^2(b-a)^2}{8}}$